Appendix A

Complements to Topology and Measure Theory

We review here the facts about topology and measure theory that are used in the main body. Those that are covered in every graduate course on integration are stated without proof. Some facts that might be considered as going beyond the very basics are proved, or at least a reference is given. The presentation is not linear – the two indexes will help the reader navigate.

A.1 Notations and Conventions

Convention A.1.1 *The reals are denoted by* $\mathbb{R}$, *the complex numbers by* $\mathbb{C}$, *the rationals by* $\mathbb{Q}$. $\mathbb{R}^d_*$ *is* ***punctured d-space*** $\mathbb{R}^d \setminus \{0\}$.

A real number a *will be called* ***positive*** *if* $a \geq 0$, *and* $\mathbb{R}_+$ *denotes the set of positive reals. Similarly, a real-valued function* f *is positive if* $f(x) \geq 0$ *for all points* x *in its domain* $\operatorname{dom}(f)$. *If we want to emphasize that* f *is* ***strictly positive****:* $f(x) > 0 \quad \forall\, x \in \operatorname{dom}(f)$, *we shall say so. It is clear what the words "****negative****" and "****strictly negative****" mean. If* $\mathcal{F}$ *is a collection of functions, then* $\mathcal{F}_+$ *will denote the positive functions in* $\mathcal{F}$, *etc. Note that a positive function may be zero on a large set, in fact, everywhere. The statements "*b ***exceeds*** a*," "*b ***is bigger than*** a*," and "*a ***is less than*** b*" all mean "*$a \leq b$*;" modified by the word "strictly" they mean "*$a < b$*."*

A.1.2 The Extended Real Line $\overline{\mathbb{R}}$ is the real line augmented by the two symbols $-\infty$ and $\infty = +\infty$:

$$\overline{\mathbb{R}} = \{-\infty\} \cup \mathbb{R} \cup \{+\infty\} .$$

We adopt the usual conventions concerning the arithmetic and order structure of the extended reals $\overline{\mathbb{R}}$:

$$-\infty < r < +\infty \quad \forall\, r \in \mathbb{R}\, ; \quad |\pm\infty| = +\infty\, ;$$

$$-\infty \wedge r = -\infty\, , \ \ \infty \vee r = \infty \quad \forall\, r \in \mathbb{R}\, ;$$

$$-\infty + r = -\infty\, , \ +\infty + r = +\infty \quad \forall\, r \in \mathbb{R}\, ;$$

$$r \cdot \pm\infty = \begin{cases} \pm\infty & \text{for } r > 0\, , \\ 0 & \text{for } r = 0\, , \\ \mp\infty & \text{for } r < 0\, ; \end{cases} \qquad \pm\infty^p = \begin{cases} \pm\infty & \text{for } p > 0\, , \\ 1 & \text{for } p = 0\, , \\ 0 & \text{for } p < 0\, . \end{cases}$$

The symbols $\infty - \infty$ and $0/0$ are not defined; there is no way to do this without confounding the order or the previous conventions.

A function whose values may include $\pm\infty$ is often called a ***numerical function***. The ***extended reals*** $\overline{\mathbb{R}}$ form a complete metric space under the ***arctan metric***

$$\overline{\rho}(r,s) \stackrel{\text{def}}{=} \big|\arctan(r) - \arctan(s)\big| \,, \qquad r,s \in \overline{\mathbb{R}} \,.$$

Here $\arctan(\pm\infty) \stackrel{\text{def}}{=} \pm\,\pi/2$. $\overline{\mathbb{R}}$ is compact in the topology $\overline{\tau}$ of $\overline{\rho}$. The natural injection $\mathbb{R} \hookrightarrow \overline{\mathbb{R}}$ is a ***homeomorphism***; that is to say, $\overline{\tau}$ agrees on $\mathbb{R}$ with the usual topology.

$a \vee b$ $(a \wedge b)$ is the larger (smaller) of a and b. If f,g are numerical functions, then $f \vee g$ $(f \wedge g)$ denote the pointwise maximum (minimum) of f,g. When a set $S \subset \overline{\mathbb{R}}$ is unbounded above we write $\sup S = \infty$, and $\inf S = -\infty$ when it is not bounded below. It is convenient to define the infimum of the empty set $\emptyset$ to be $+\infty$ and to write $\sup\emptyset = -\infty$.

A.1.3 Different length measurements of vectors and sequences come in handy in different places. For $0 < p < \infty$ the ***ℓ^p-length*** of $x = (x^1, \ldots, x^n) \in \mathbb{R}^n$ or of a sequence $(x^1, x^2, \ldots)$ is written variously as

$$|\,x\,|_p = \|\,x\,\|_{\ell^p} \stackrel{\text{def}}{=} \Big(\textstyle\sum_\nu |x^\nu|^p\Big)^{1/p}\,, \quad \text{while} \quad |\,\boldsymbol{z}\,|_\infty = \|\,\boldsymbol{z}\,\|_{\ell^\infty} \stackrel{\text{def}}{=} \sup_\eta |\,z^\eta\,|$$

denotes the ℓ^∞-length of a d-tuple $\boldsymbol{z} = (z^1, z^2, \ldots, z^d)$ or a sequence $\boldsymbol{z} = (z^0, z^1, \ldots)$. The vector space of all scalar sequences is a Fréchet space (which see) under the topology of pointwise convergence and is denoted by ℓ^0. The sequences x having $|\,x\,|_p < \infty$ form a Banach space under $|\ \ |_p$, $1 \le p < \infty$, which is denoted by ℓ^p. For $0 < p < q$ we have

$$|\boldsymbol{z}|_q \le |\boldsymbol{z}|_p \,; \quad \text{and} \quad |\boldsymbol{z}|_p \le d^{1/(q-p)} \cdot |\boldsymbol{z}|_q \tag{A.1.1}$$

on sequences $\boldsymbol{z}$ of finite length d. $|\ \ |$ stands not only for the ordinary absolute value on $\mathbb{R}$ or $\mathbb{C}$ but for any of the norms $|\ \ |_p$ on $\mathbb{R}^n$ when $p \in [0,\infty]$ need not be displayed.

Next some notation and conventions concerning sets and functions, which will simplify the typography considerably:

Notation A.1.4 (Knuth [57]) *A statement enclosed in rectangular brackets denotes the set of points where it is true. For instance, the symbol* $[f = 1]$ *is short for* $\{x \in \operatorname{dom}(f) : f(x) = 1\}$. *Similarly,* $[f > r]$ *is the set of points* x *where* $f(x) > r$, $[f_n \not\to]$ *is the set of points where the sequence* (f_n) *fails to converge, etc.*

Convention A.1.5 (Knuth [57]) *Occasionally we shall use the same name or symbol for a set* A *and its indicator function:* A *is also the function that returns* 1 *when its argument lies in* A, 0 *otherwise. For instance,* $[f > r]$ *denotes not only the set of points where* f *strictly exceeds* r *but also the function that returns* 1 *where* $f > r$ *and* 0 *elsewhere.*

Remark A.1.6 The indicator function of A is written 1_A by most mathematicians, $\imath_A$, χ_A, or I_A or even $1A$ by others, and A by a select few. There is a considerable typographical advantage in writing it as A: $[T_n^k < r]$ or $\left[U_{\mathcal{S}}^{[a,b]} \geq n\right]$ are rather easier on the eye than $1_{[T_n^k<r]}$ or $1_{\left[U_{\mathcal{S}}^{[a,b]} \geq n\right]}$, respectively. When functions z are written $s \mapsto z_s$, as is common in stochastic analysis, the indicator of the interval $(a, b]$ has under this convention at s the value $(a, b]_s$ rather than $1_{(a,b]s}$.

In deference to prevailing custom we shall use this swell convention only sparingly, however, and write 1_A when possible. We do invite the aficionado of the 1_A-notation to compare how much eye strain and verbiage is saved on the occasions when we employ Knuth's nifty convention.

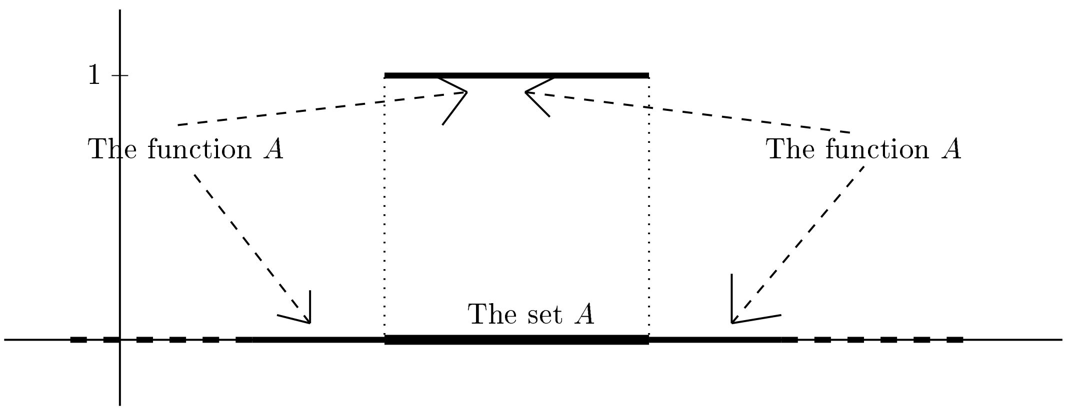


Figure A.14 A set and its indicator function have the same name

Exercise A.1.7 Here is a little practice to get used to Knuth's convention: (i) Let f be an ***idempotent function***: $f^2 = f$. Then $f = [f = 1] = [f \neq 0]$. (ii) Let (f_n) be a sequence of functions. Then the sequence $\left([f_n \to] \cdot f_n\right)$ converges everywhere. (iii) $\int (0, 1](x) \cdot x^2 \, dx = 1/3$. (iv) For $f_n(x) \stackrel{\text{def}}{=} \sin(nx)$ compute $[f_n \not\to]$. (v) Let $A_1, A_2, \ldots$ be sets. Then $A_1^c = 1 - A_1$, $\bigcup_n A_n = \sup_n A_n = \bigvee_n A_n$, $\bigcap_n A_n = \inf_n A_n = \bigwedge_n A_n$. (vi) Every real-valued function is the pointwise limit of the simple step functions

$$f_n = \sum_{k=-4^n}^{4^n} k2^{-n} \cdot \left[k2^{-n} \leq f < (k+1)2^{-n}\right].$$

(vii) The sets in an algebra of functions form an algebra of sets. (viii) A family of idempotent functions is a σ-algebra if and only if it is closed under $f \mapsto 1 - f$ and under finite infima and countable suprema. (ix) Let $f : X \to Y$ be a map and $S \subset Y$. Then $f^{-1}(S) = S \circ f \subset X$.

A.2 Topological Miscellanea

The Theorem of Stone–Weierstraß

All measures appearing in nature owe their σ-additivity to the following very simple result or some variant of it; it is also used in the proof of theorem A.2.2.

Lemma A.2.1 (Dini's Theorem) *Let $\boldsymbol{B}$ be a topological space and Φ a collection of positive continuous functions on $\boldsymbol{B}$ that vanish at infinity.*[1] *Assume Φ is decreasingly directed*[2,3] *and the pointwise infimum of Φ is zero. Then $\Phi \to 0$ uniformly; that is to say, for every $\epsilon > 0$ there is a $\psi \in \Phi$ with $\psi(x) \le \epsilon$ for all $x \in \boldsymbol{B}$ and therefore $\phi \le \epsilon$ uniformly for all $\phi \in \Phi$ with $\phi \le \psi$.*

Proof. The sets $[\phi \ge \epsilon]$, $\phi \in \Phi$, are compact and have void intersection. There are finitely many of them, say $[\phi_i \ge \epsilon]$, $i = 1, \ldots, n$, whose intersection is void (exercise A.2.12). There exists a $\psi \in \Phi$ smaller than[3] $\phi_1 \wedge \cdots \wedge \phi_n$. If $\phi \in \Phi$ is smaller than ψ, then $|\phi| = \phi < \epsilon$ everywhere on $\boldsymbol{B}$. ∎

Consider a vector space $\mathcal{E}$ of real-valued functions on some set $\boldsymbol{B}$. It is an ***algebra*** if with any two functions ϕ, ψ it contains their pointwise product $\phi\psi$. For this it suffices that it contain with any function ϕ its square ϕ^2. Indeed, by ***polarization*** then $\phi\psi = 1/2\big((\phi+\psi)^2 - \phi^2 - \psi^2\big) \in \mathcal{E}$. $\mathcal{E}$ is a ***vector lattice*** if with any two functions ϕ, ψ it contains their pointwise maximum $\phi \vee \psi$ and their pointwise minimum $\phi \wedge \psi$. For this it suffices that it contain with any function ϕ its absolute value $|\phi|$. Indeed, $\phi \vee \psi = 1/2\big(|\phi - \psi| + (\phi + \psi)\big)$, and $\phi \wedge \psi = (\phi + \psi) - (\phi \vee \psi)$. $\mathcal{E}$ is ***closed under chopping*** if with any function ϕ it contains the ***chopped function*** $\phi \wedge 1$. It then contains $f \wedge q = q(f/q \wedge 1)$ for any strictly positive scalar q. A ***lattice algebra*** is a vector space of functions that is both an algebra and a lattice under pointwise operations.

Theorem A.2.2 (Stone–Weierstraß) *Let $\mathcal{E}$ be an algebra* OR *a vector lattice closed under chopping, of bounded real-valued functions on some set $\boldsymbol{B}$. We denote by Z the set $\{x \in \boldsymbol{B} : \phi(x) = 0 \quad \forall\, \phi \in \mathcal{E}\}$ of common zeroes of $\mathcal{E}$, and identify a function of $\mathcal{E}$ in the obvious fashion with its restriction to $\boldsymbol{B}_0 \overset{\text{def}}{=} \boldsymbol{B} \setminus Z$.*

(i) The uniform closure $\overline{\mathcal{E}}$ of $\mathcal{E}$ is both an algebra AND *a vector lattice closed under chopping. Furthermore, if $\Phi : \mathbb{R} \to \mathbb{R}$ is continuous with $\Phi(0) = 0$, then $\Phi \circ \overline{\phi} \in \overline{\mathcal{E}}$ for any $\overline{\phi} \in \overline{\mathcal{E}}$.*

---

[1] A set is ***relatively compact*** if its closure is compact. ϕ vanishes at ∞ if its carrier $[|\phi| \ge \epsilon]$ is relatively compact for every $\epsilon > 0$. The collection of continuous functions vanishing at infinity is denoted by $C_0(\boldsymbol{B})$ and is given the topology of uniform convergence. $C_0(\boldsymbol{B})$ is identified in the obvious way with the collection of continuous functions on the one-point compactification $\boldsymbol{B}^\Delta \overset{\text{def}}{=} \boldsymbol{B} \cup \{\Delta\}$ (see page 374) that vanish at Δ.

[2] That is to say, for any two $\phi_1, \phi_2 \in \Phi$ there is a $\phi \in \Phi$ less than both ϕ_1 and ϕ_2. Φ is ***increasingly directed*** if for any two $\phi_1, \phi_2 \in \Phi$ there is a $\phi \in \Phi$ with $\phi \ge \phi_1 \vee \phi_2$.

[3] See convention A.1.1 on page 363 about language concerning order relations.

(ii) There exist a locally compact Hausdorff space $\widehat{\boldsymbol{B}}$ *and a map* $j : \boldsymbol{B}_0 \to \widehat{\boldsymbol{B}}$ *with dense image such that* $\widehat{\phi} \mapsto \widehat{\phi} \circ j$ *is an algebraic and order isomorphism of* $C_0(\widehat{\boldsymbol{B}})$ *with* $\overline{\mathcal{E}} \simeq \overline{\mathcal{E}}_{|\boldsymbol{B}_o}$. *We call* $\widehat{\boldsymbol{B}}$ *the* ***spectrum*** *of* $\mathcal{E}$ *and* $j : \boldsymbol{B}_0 \to \widehat{\boldsymbol{B}}$ *the* ***local*** $\mathcal{E}$***-compactification*** *of* $\boldsymbol{B}$. $\widehat{\boldsymbol{B}}$ *is compact if and only if* $\mathcal{E}$ *contains a function that is bounded away from zero.*[4] *If* $\mathcal{E}$ ***separates the points***[5] *of* $\boldsymbol{B}_0$, *then* j *is injective. If* $\mathcal{E}$ *is countably generated,*[6] *then* $\widehat{\boldsymbol{B}}$ *is separable and metrizable.*

(iii) Suppose that there is a locally compact Hausdorff topology τ *on* $\boldsymbol{B}$ *and* $\mathcal{E} \subset C_0(\boldsymbol{B}, \tau)$, *and assume that* $\mathcal{E}$ *separates the points of* $\boldsymbol{B}_0 \overset{\text{def}}{=} \boldsymbol{B} \backslash Z$. *Then* $\overline{\mathcal{E}}$ *equals the algebra of all continuous functions that vanish at infinity and on* Z.[7]

Proof. (i) There are several steps. (a) If $\mathcal{E}$ is an algebra OR a vector lattice closed under chopping, then its uniform closure $\overline{\mathcal{E}}$ is clearly an algebra OR a vector lattice closed under chopping, respectively.

(b) Assume that $\mathcal{E}$ is an algebra and let us show that then $\overline{\mathcal{E}}$ is a vector lattice closed under chopping. To this end define polynomials $p_n(t)$ on $[-1, 1]$ inductively by $p_0 = 0$, $p_{n+1}(t) = 1/2\big(t^2 + 2p_n(t) - \big(p_n(t)\big)^2\big)$. Then $p_n(t)$ is a polynomial in t^2 with zero constant term. Two easy manipulations result in

$$2\big(|t| - p_{n+1}(t)\big) = \big(2 - |t|\big)|t| - \big(2 - p_n(t)\big)p_n(t)$$

and $$2\big(p_{n+1}(t) - p_n(t)\big) = t^2 - \big(p_n(t)\big)^2 .$$

Now $(2 - x)x = 2x - x^2$ is increasing on $[0, 1]$. If, by induction hypothesis, $0 \le p_n(t) \le |t|$ for $|t| \le 1$, then $p_{n+1}(t)$ will satisfy the same inequality; as it is true for p_0, it holds for all the p_n. The second equation shows that $p_n(t)$ increases with n for $t \in [-1, 1]$. As this sequence is also bounded, it has a limit $p(t) \ge 0$. $p(t)$ must satisfy $0 = t^2 - (p(t))^2$ and thus equals $|t|$. Due to Dini's theorem A.2.1, $|t| - p_n(t)$ decreases uniformly on $[-1, 1]$ to 0. Given a $\overline{\phi} \in \overline{\mathcal{E}}$, set $M = \|\overline{\phi}\|_\infty \vee 1$. Then $P_n(t) \overset{\text{def}}{=} Mp_n(t/M)$ converges to $|t|$ uniformly on $[-M, M]$, and consequently $|f| = \lim P_n(f)$ belongs to $\overline{\overline{\mathcal{E}}} = \overline{\mathcal{E}}$. To see that $\overline{\mathcal{E}}$ is closed under chopping consider the polynomials $Q'_n(t) \overset{\text{def}}{=} 1/2\big(t + 1 - P_n(t-1)\big)$. They converge uniformly on $[-M, M]$ to $1/2\big(t + 1 - |t - 1|\big) = t \wedge 1$. So do the polynomials $Q_n(t) = Q'_n(t) - Q'_n(0)$, which have the virtue of vanishing at zero, so that $Q_n \circ \overline{\phi} \in \overline{\mathcal{E}}$. Therefore $\overline{\phi} \wedge 1 = \lim Q_n \circ \overline{\phi}$ belongs to $\overline{\overline{\mathcal{E}}} = \overline{\mathcal{E}}$.

(c) Next assume that $\mathcal{E}$ is a vector lattice closed under chopping, and let us show that then $\overline{\mathcal{E}}$ is an algebra. Given $\overline{\phi} \in \overline{\mathcal{E}}$ and $\epsilon \in (0, 1)$, again set

---

[4] ϕ is ***bounded away from zero*** if $\inf\{|\phi(x)| : x \in \boldsymbol{B}\} > 0$.

[5] That is to say, for any $x \ne y$ in $\boldsymbol{B}_0$ there is a $\phi \in \mathcal{E}$ with $\phi(x) \ne \phi(y)$.

[6] That is to say, there is a countable set $\mathcal{E}_0 \subset \mathcal{E}$ such that $\mathcal{E}$ is contained in the smallest uniformly closed algebra containing $\mathcal{E}_0$.

[7] If $Z = \emptyset$, this means $\overline{\mathcal{E}} = C_0(\boldsymbol{B})$; if in addition τ is compact, then this means $\overline{\mathcal{E}} = C(\boldsymbol{B})$.

$M = \|\overline{\phi}\|_\infty + 1$. For $k \in \mathbb{Z} \cap [-M/\epsilon, M/\epsilon]$ let $\ell_k(t) = 2k\epsilon t - k^2\epsilon^2$ denote the tangent to the function $t \mapsto t^2$ at $t = k\epsilon$. Since $\ell_k \vee 0 = (2k\epsilon t - k^2\epsilon^2) \vee 0 = 2k\epsilon t - (2k\epsilon t \wedge k^2\epsilon^2)$, we have

$$\begin{aligned}\Phi_\epsilon(t) &\stackrel{\text{def}}{=} \bigvee \left\{2k\epsilon t - k^2\epsilon^2 : q\; k \in \mathbb{Z},\, |k| < M/\epsilon\right\} \\ &= \bigvee \left\{2k\epsilon t - (2k\epsilon t \wedge k^2\epsilon^2) :\; k \in \mathbb{Z},\, |k| < M/\epsilon\right\} .\end{aligned}$$

Now clearly $t^2 - \epsilon \le \Phi_\epsilon(t) \le t^2$ on $[-M, M]$, and the second line above shows that $\Phi_e \circ \overline{\phi} \in \overline{\mathcal{E}}$. We conclude that $\overline{\phi}^2 = \lim_{\epsilon\to 0} \Phi_e \circ \overline{\phi} \in \overline{\overline{\mathcal{E}}} = \overline{\mathcal{E}}$.

We turn to (iii), assuming to start with that τ is compact. Let $\overline{\mathcal{E}} \oplus \mathbb{R} \stackrel{\text{def}}{=} \{\phi + r :\; \phi \in \overline{\mathcal{E}},\, r \in \mathbb{R}\}$. This is an algebra AND a vector lattice[8] over $\mathbb{R}$ of bounded τ-continuous functions. It is uniformly closed and contains the constants. Consider a continuous function f that is constant on Z, and let $\epsilon > 0$ be given. For any two different points $s, t \in \boldsymbol{B}$, not both in Z, there is a function $\overline{\psi}_{s,t}$ in $\overline{\mathcal{E}}$ with $\overline{\psi}_{s,t}(s) \ne \overline{\psi}_{s,t}(t)$. Set

$$\overline{\phi}_{s,t}(\tau) = f(s) + \left(\frac{f(t) - f(s)}{\overline{\psi}_{s,t}(t) - \overline{\psi}_{s,t}(s)}\right) \cdot (\overline{\psi}_{s,t}(\tau) - \overline{\psi}_{s,t}(s)) .$$

If $s = t$ or $s, t \in Z$, set $\overline{\phi}_{s,t}(\tau) = f(t)$. Then $\overline{\phi}_{s,t}$ belongs to $\overline{\mathcal{E}} \oplus \mathbb{R}$ and takes at s and t the same values as f. Fix $t \in \boldsymbol{B}$ and consider the sets $U^t_s = [\overline{\phi}_{s,t} > f - \epsilon]$. They are open, and they cover $\boldsymbol{B}$ as s ranges over $\boldsymbol{B}$; indeed, the point $s \in \boldsymbol{B}$ belongs to U^t_s. Since $\boldsymbol{B}$ is compact, there is a finite subcover $\{U^t_{s_i} :\; 1 \le i \le n\}$. Set $\overline{\phi}^t = \bigvee_{i=1}^n \overline{\phi}_{s_i,t}$. This function belongs to $\overline{\mathcal{E}} \oplus \mathbb{R}$, is everywhere bigger than $f - \epsilon$, and coincides with f at t. Next consider the open cover $\{[\overline{\phi}^t < f + \epsilon] :\; t \in \boldsymbol{B}\}$. It has a finite subcover $\{[\overline{\phi}^{t_i} < f + \epsilon] :\; 1 \le i \le k\}$, and the function $\phi \stackrel{\text{def}}{=} \bigwedge_{i=1}^k \overline{\phi}^{t_i} \in \overline{\mathcal{E}} \oplus \mathbb{R}$ is clearly uniformly as close as ϵ to f. In other words, there is a sequence $\overline{\phi}_n + r_n \in \overline{\mathcal{E}} \oplus \mathbb{R}$ that converges uniformly to f. Now if Z is non-void and f vanishes on Z, then $r_n \to 0$ and $\overline{\phi}_n \in \overline{\mathcal{E}}$ converges uniformly to f. If $Z = \emptyset$, then there is, for every $s \in \boldsymbol{B}$, a $\overline{\phi}_s \in \overline{\mathcal{E}}$ with $\overline{\phi}_s(s) > 1$. By compactness there will be finitely many of the $\overline{\phi}_s$, say $\overline{\phi}_{s_1}, \ldots, \overline{\phi}_{s_n}$, with $\overline{\phi} \stackrel{\text{def}}{=} \bigvee_i \overline{\phi}_{s_i} > 1$. Then $1 = \overline{\phi} \wedge 1 \in \overline{\mathcal{E}}$ and consequently $\overline{\mathcal{E}} \oplus \mathbb{R} = \overline{\mathcal{E}}$. In both cases $f \in \overline{\overline{\mathcal{E}}} = \overline{\mathcal{E}}$. If τ is not compact, we view $\overline{\mathcal{E}} \oplus \mathbb{R}$ as a uniformly closed algebra of bounded continuous functions on the one-point compactification $\boldsymbol{B}^\Delta = \boldsymbol{B} \,\dot\cup\, \{\Delta\}$ and an $f \in C_0(\boldsymbol{B})$ that vanishes on Z as a continuous bounded function on $\boldsymbol{B}^\Delta$ that vanishes on $Z \cup \{\Delta\}$, the common zeroes of $\mathcal{E}$ on $\boldsymbol{B}^\Delta$, and apply the above: if $\overline{\mathcal{E}} \oplus \mathbb{R} \ni \phi_n + r_n \to f$ uniformly on $\boldsymbol{B}^\Delta$, then $r_n \to 0$, and $f \in \overline{\overline{\mathcal{E}}} = \overline{\mathcal{E}}$.

[8] To see that $\overline{\mathcal{E}} \oplus \mathbb{R}$ is closed under pointwise infima write $(\overline{\phi} + r) \wedge (\overline{\psi} + s) = (\overline{\phi} - \overline{\psi}) \wedge (s - r) + \overline{\psi} + r$. Since without loss of generality $r \le s$, the right-hand side belongs to $\overline{\mathcal{E}} \oplus \mathbb{R}$.

(d) Of (i) only the last claim remains to be proved. Now thanks to (iii) there is a sequence of polynomials q_n that vanish at zero and converge uniformly on the compact set $[-\|\overline{\phi}\|_\infty, \|\overline{\phi}\|_\infty]$ to Φ. Then

$$\Phi \circ \overline{\phi} = \lim q_n \circ \overline{\phi} \in \overline{\overline{\mathcal{E}}} = \overline{\mathcal{E}} .$$

(ii) Let $\mathcal{E}_0$ be a subset of $\mathcal{E}$ that generates $\mathcal{E}$ in the sense that $\mathcal{E}$ is contained in the smallest uniformly closed algebra containing $\mathcal{E}_0$. Set

$$\Pi = \prod_{\psi \in \mathcal{E}_0} \left[-\|\psi\|_u, +\|\psi\|_u \right] .$$

This product of compact intervals is a compact Hausdorff space in the product topology (exercise A.2.13), metrizable if $\mathcal{E}_0$ is countable. Its typical element is an "$\mathcal{E}_0$-tuple" $(\xi_\psi)_{\psi \in \mathcal{E}_0}$ with $\xi_\psi \in [-\|\psi\|_u, +\|\psi\|_u]$. There is a natural map $j : \boldsymbol{B} \to \Pi$ given by $x \mapsto (\psi(x))_{\psi \in \mathcal{E}_0}$. Let $\overline{\boldsymbol{B}}$ denote the closure of $j(\boldsymbol{B})$ in Π, the ***$\mathcal{E}$-completion*** of $\boldsymbol{B}$ (see lemma A.2.16). The finite linear combinations of finite products of coordinate functions $\widehat{\phi} : (\xi_\psi)_{\psi \in \mathcal{E}_0} \mapsto \xi_\phi$, $\phi \in \mathcal{E}_0$, form an algebra $\mathcal{A} \subset C(\overline{\boldsymbol{B}})$ that separates the points. Now set $\overline{Z} \stackrel{\text{def}}{=} \{\widehat{z} \in \overline{\boldsymbol{B}} : \widehat{\phi}(\widehat{z}) = 0 \quad \forall\, \widehat{\phi} \in \mathcal{A}\}$. This set is either empty or contains one point, $(0, 0, \ldots)$, and j maps $\boldsymbol{B}_0 \stackrel{\text{def}}{=} \boldsymbol{B} \setminus Z$ into $\widehat{\boldsymbol{B}} \stackrel{\text{def}}{=} \overline{\boldsymbol{B}} \setminus \overline{Z}$. View $\mathcal{A}$ as a subalgebra of $C_0(\widehat{\boldsymbol{B}})$ that separates the points of $\boldsymbol{B}$. The linear multiplicative map $\widehat{\phi} \mapsto \widehat{\phi} \circ j$ evidently takes $\mathcal{A}$ to the smallest algebra containing $\mathcal{E}_0$ and preserves the uniform norm. It extends therefore to a linear isometry of $\overline{\mathcal{A}}$ – which by (iii) coincides with $C_0(\widehat{\boldsymbol{B}})$ – with $\overline{\mathcal{E}}$; it is evidently linear and multiplicative and preserves the order. Finally, if $\phi \in \mathcal{E}$ separates the points $x, y \in \boldsymbol{B}$, then the function $\widehat{\phi} \in \overline{\mathcal{A}}$ that has $\phi = \widehat{\phi} \circ j$ separates $j(x), j(y)$, so when $\mathcal{E}$ separates the points then j is injective. ∎

Exercise A.2.3 Let $\boldsymbol{A}$ be any subset of $\boldsymbol{B}$. (i) A function f can be approximated uniformly on $\boldsymbol{A}$ by functions in $\mathcal{E}$ if and only if it is the restriction to $\boldsymbol{A}$ of a function in $\overline{\mathcal{E}}$. (ii) If $f_1, f_2 : \boldsymbol{B} \to \overline{\mathbb{R}}$ can be approximated uniformly on $\boldsymbol{A}$ by functions in $\mathcal{E}$ (in the arctan metric $\overline{\rho}$; see item A.1.2), then $\overline{\rho}(f_1, f_2) : b \mapsto \overline{\rho}\big(f_1(b), f_2(b)\big)$ is the restriction to $\boldsymbol{A}$ of a function in $\overline{\mathcal{E}}$.

All spaces of elementary integrands that we meet in this book are self-confined in the following sense.

Definition A.2.4 *A subset $S \subset \boldsymbol{B}$ is called **$\mathcal{E}$-confined** if there is a function $\phi \in \mathcal{E}$ that is greater than 1 on S: $\phi \geq 1_S$. A function $f : \boldsymbol{B} \to \mathbb{R}$ is $\mathcal{E}$-confined if its carrier*[9] *$[f \neq 0]$ is $\mathcal{E}$-confined; the collection of $\mathcal{E}$-confined functions in $\mathcal{E}$ is denoted by $\mathcal{E}_{00}$. A sequence of functions f_n on $\boldsymbol{B}$ is $\mathcal{E}$-confined if the f_n all vanish outside the same $\mathcal{E}$-confined set; and $\mathcal{E}$ is **self-confined** if all of its members are $\mathcal{E}$-confined, i.e., if $\mathcal{E} = \mathcal{E}_{00}$. A function f is the **$\mathcal{E}$-confined uniform limit** of the sequence (f_n) if (f_n) is*

[9] The ***carrier of a function*** ϕ is the set $[\phi \neq 0]$.

$\mathcal{E}$-confined and converges uniformly to f.

The typical examples of self-confined lattice algebras are the step functions over a ring of sets and the space $C_{00}(\boldsymbol{B})$ of continuous functions with compact support on $\boldsymbol{B}$. The product $\mathcal{E}_1 \otimes \mathcal{E}_2$ of two self-confined algebras or vector lattices closed under chopping is clearly self-confined.

A.2.5 The notion of a confined uniform limit is a topological notion: for every $\mathcal{E}$-confined set A let $\mathcal{F}_A$ denote the algebra of bounded functions confined by A. Its natural topology is the topology of uniform convergence. The natural topology on the vector space $\mathcal{F}_{\mathcal{E}}$ of bounded $\mathcal{E}$-confined functions, union of the $\mathcal{F}_A$, the ***topology of $\mathcal{E}$-confined uniform convergence*** is the finest topology on bounded $\mathcal{E}$-confined functions that agrees on every $\mathcal{F}_A$ with the topology of uniform convergence. It makes the bounded $\mathcal{E}$-confined functions, the union of the $\mathcal{F}_A$, into a topological vector space. Now let $\mathcal{I}$ be a linear map from $\mathcal{F}_{\mathcal{E}}$ to a topological vector space and show that the following are equivalent: (i) $\mathcal{I}$ is continuous in this topology; (ii) the restriction of $\mathcal{I}$ to any of the $\mathcal{F}_A$ is continuous; (iii) $\mathcal{I}$ maps order-bounded subsets of $\mathcal{F}_{\mathcal{E}}$ to bounded subsets of the target space.

Exercise A.2.6 Show: if $\mathcal{E}$ is a self-confined algebra or vector lattice closed under chopping, then a uniform limit $\overline{\phi} \in \overline{\mathcal{E}}$ is $\mathcal{E}$-confined if and only if it is the uniform limit of a $\mathcal{E}$-confined sequence in $\mathcal{E}$; we then say "$\overline{\phi}$ is the ***confined uniform limit***" of a sequence in $\mathcal{E}$. Therefore the "confined uniform closure $\overline{\mathcal{E}}_{00}$ of $\mathcal{E}$" is a self-confined algebra and a vector lattice closed under chopping.

The next two corollaries to Weierstraß' theorem employ the local $\mathcal{E}$-compactification $j : \boldsymbol{B}_0 \to \widehat{\boldsymbol{B}}$ to establish results that are crucial for the integration theory of integrators and random measures (see proposition 3.3.2 and lemma 3.10.2). In order to ease their statements and the arguments to prove them we introduce the following notation: for every $X \in \overline{\mathcal{E}}$ the unique continuous function $\widehat{X}$ on $\widehat{\boldsymbol{B}}$ that has $\widehat{X} \circ j = X$ will be called the ***Gelfand transform*** of X; next, given any functional $\mathcal{I}$ on $\mathcal{E}$ we define $\widehat{\mathcal{I}}$ on $\widehat{\mathcal{E}}$ by $\widehat{\mathcal{I}}(\widehat{X}) = \mathcal{I}(\widehat{X} \circ j)$ and call it the ***Gelfand transform*** of $\mathcal{I}$.

For simplicity's sake we assume in the remainder of this subsection that $\mathcal{E}$ is a self-confined algebra and a vector lattice closed under chopping, of bounded functions on some set $\boldsymbol{B}$.

Corollary A.2.7 *Let (L,τ) be a topological vector space and $\tau_0 \subset \tau$ a weaker Hausdorff topology on L. If $\mathcal{I} : \mathcal{E} \to L$ is a linear map whose Gelfand transform $\widehat{\mathcal{I}}$ has an extension satisfying the Dominated Convergence Theorem, and if $\mathcal{I}$ is σ-continuous in the topology τ_0, then it is in fact σ-additive in the topology τ.*

Proof. Let $\mathcal{E} \ni X_n \downarrow 0$. Then the sequence $(\widehat{X}_n)$ of Gelfand transforms decreases on $\widehat{\boldsymbol{B}}$ and has a pointwise infimum $\widehat{K} : \widehat{\boldsymbol{B}} \to \mathbb{R}$. By the DCT, the sequence $\big(\widehat{\mathcal{I}}(\widehat{X}_n)\big)$ has a τ-limit f in L, the value of the extension at $\widehat{K}$. Clearly $f = \tau\!-\lim \widehat{\mathcal{I}}(\widehat{X}_n) = \tau\!-\lim \mathcal{I}(X_n) = \tau_0\!-\lim \mathcal{I}(X_n) = 0$. Since

τ_0 is Hausdorff, $f = 0$. The σ-continuity of $\mathcal{I}$ is established, and exercise 3.1.5 on page 90 produces the σ-additivity. This argument repeats that of proposition 3.3.2 on page 108. ∎

Corollary A.2.8 *Let $\boldsymbol{H}$ be a locally compact space equipped with the algebra $\mathcal{H} \overset{\text{def}}{=} C_{00}(\boldsymbol{H})$ of continuous functions of compact support. The cartesian product $\check{\boldsymbol{B}} \overset{\text{def}}{=} \boldsymbol{H} \times \boldsymbol{B}$ is equipped with the algebra $\check{\mathcal{E}} \overset{\text{def}}{=} \mathcal{H} \otimes \mathcal{E}$ of functions*

$$(\eta, \varpi) \mapsto \sum_i H_i(\eta) X_i(\varpi)\ , \qquad H_i \in \mathcal{H}, X_i \in \mathcal{E},\ \textit{the sum finite.}$$

*Suppose θ is a real-valued linear functional on $\check{\mathcal{E}}$ that maps order-bounded sets of $\check{\mathcal{E}}$ to bounded sets of reals and that is **marginally $\boldsymbol{\sigma}$-additive** on $\mathcal{E}$; that is to say, the measure $X \mapsto \theta(H \otimes X)$ on $\mathcal{E}$ is σ-additive for every $H \in \mathcal{H}$. Then θ is in fact σ-additive.*[10]

Proof. First observe easily that $\mathcal{E} \ni X_n \downarrow 0$ implies $\theta(\check{H} \cdot X_n) \to 0$ for every $\check{H} \in \check{\mathcal{E}}$. Another way of saying this is that for every $\check{H} \in \check{\mathcal{E}}$ the measure $X \mapsto \theta^{\check{H}}(X) \overset{\text{def}}{=} \theta(\check{H} \cdot X)$ on $\mathcal{E}$ is σ-additive.

From this let us deduce that the variation $¦\theta¦$ has the same property. To this end let $\jmath : \boldsymbol{B}_0 \to \widehat{\boldsymbol{B}}$ denote the local $\mathcal{E}$-compactification of $\boldsymbol{B}$; the local $\mathcal{H}$-compactification of $\boldsymbol{H}$ clearly is the identity map $id : \boldsymbol{H} \to \boldsymbol{H}$. The spectrum of $\check{\mathcal{E}}$ is $\widehat{\check{\boldsymbol{B}}} = \boldsymbol{H} \times \widehat{\boldsymbol{B}}$ with local $\check{\mathcal{E}}$-compactification $\check{\jmath} \overset{\text{def}}{=} id \otimes \jmath$. The Gelfand transform $\widehat{\theta}$ is a σ-additive measure on $\widehat{\check{\mathcal{E}}}$ of finite variation $¦\widehat{\theta}¦ = \widehat{¦\theta¦}$; in fact, $¦\widehat{\theta}¦$ is a positive Radon measure on $\widehat{\check{\boldsymbol{B}}}$. There exists a locally $¦\widehat{\theta}¦$-integrable function $\widehat{\check{\Gamma}}$ with $\widehat{\check{\Gamma}}^2 = 1$ and $\widehat{\theta} = \widehat{\check{\Gamma}} \cdot ¦\widehat{\theta}¦$ on $\widehat{\check{\boldsymbol{B}}}$, to wit, the Radon–Nikodym derivative $d\widehat{\theta}/d¦\widehat{\theta}¦$. With these notations in place pick an $H \in \mathcal{H}_+$ with compact carrier K and let (X_n) be a sequence in $\mathcal{E}$ that decreases pointwise to 0. There is no loss of generality in assuming that both $X_1 < 1$ and $H < 1$. Given an $\epsilon > 0$, let $\widehat{E}$ be the closure in $\widehat{\boldsymbol{B}}$ of $\jmath([X_1 > 0])$ and find an $\widehat{\check{X}} \in \widehat{\check{\mathcal{E}}}$ with

$$\int_{K \times E} \big|\widehat{\check{\Gamma}} - \widehat{\check{X}}\big|\ d¦\widehat{\theta}¦ < \epsilon\ .$$

Then
$$\begin{aligned}
¦\widehat{\theta}¦(H \otimes \widehat{X}_n) &= \int_{K \times \widehat{E}} H(\eta) \widehat{X}_n(\widehat{\varpi}) \widehat{\check{\Gamma}}(\eta, \widehat{\varpi})\ \widehat{\theta}(d\eta, d\widehat{\varpi}) \\
&\le \int_{K \times \widehat{E}} H(\eta) \widehat{X}_n(\widehat{\varpi}) \widehat{\check{X}}(\eta, \widehat{\varpi})\ \widehat{\theta}(d\eta, d\widehat{\varpi}) + \epsilon \\
&= \int_{\boldsymbol{H} \times \widehat{\boldsymbol{B}}} H(\eta) \widehat{X}_n(\widehat{\varpi}) \widehat{\check{X}}(\eta, \widehat{\varpi})\ \widehat{\theta}(d\eta, d\widehat{\varpi}) + \epsilon \\
&= \theta^{H\check{X}}(X_n) + \epsilon
\end{aligned}$$

[10] Actually, it suffices to assume that $\boldsymbol{H}$ is Suslin, that the vector lattice $\mathcal{H} \subset \mathcal{B}^*(\boldsymbol{H})$ generates $\mathcal{B}^*(\boldsymbol{H})$, and that θ is also marginally σ-additive on $\mathcal{H}$ – see [94].

has limit less than ϵ by the very first observation above. Therefore

$$\not|\theta\not|(H\otimes X_n) = \not|\widehat{\theta}\not|(H\otimes \widehat{X}_n) \downarrow 0 \,.$$

Now let $\check{\mathcal{E}} \ni \check{X}_n \downarrow 0$. There are a compact set $K \subset \boldsymbol{H}$ so that $\check{X}_1(\eta,\varpi) = 0$ whenever $\eta \notin K$, and an $H \in \mathcal{H}$ equal to 1 on K. The functions $X_n : \varpi \mapsto \max_{\eta\in\boldsymbol{H}} \check{X}(\eta,\varpi)$ belong to $\mathcal{E}$, thanks to the compactness of K, and decrease pointwise to zero on $\boldsymbol{B}$ as $n\to\infty$. Since $\check{X}_n \le H\otimes X_n$, $\not|\theta\not|(\check{X}_n) \le \not|\theta\not|(H\otimes X_n) \xrightarrow[n\to\infty]{} 0$: $\not|\theta\not|$ and with it θ is indeed σ-additive. ∎

Exercise A.2.9 Let $\theta : \mathcal{E} \to \mathbb{R}$ be a linear functional of finite variation. Then its Gelfand transform $\widehat{\theta} : \widehat{\mathcal{E}} \to \mathbb{R}$ is σ-additive due to Dini's theorem A.2.1 and has the usual integral extension featuring the Dominated Convergence Theorem (see pages 395–398). Show: θ is σ-additive if and only if $\int \widehat{k}\, d\widehat{\theta} = 0$ for every function $\widehat{k}$ on the spectrum $\widehat{\boldsymbol{B}}$ that is the pointwise infimum of a sequence in $\widehat{\mathcal{E}}$ and vanishes on $j(\boldsymbol{B})$.

Exercise A.2.10 Consider a linear map $\mathcal{I}$ on $\mathcal{E}$ with values in a space $L^p(\mu)$, $1 \le p < \infty$, that maps order intervals of $\mathcal{E}$ to bounded subsets of L^p. Show: if $\mathcal{I}$ is weakly σ-additive, then it is σ-additive in in the norm topology L^p.

Weierstraß' original proof of his approximation theorem applies to functions on $\mathbb{R}^d$ and employs the heat kernel γ_{tI} (exercise A.3.48 on page 420). It yields in its less general setting an approximation in a finer topology than the uniform one. We give a sketch of the result and its proof, since it is used in Itô's theorem, for instance. Consider an open subset D of $\mathbb{R}^d$. For $0 \le k \in \mathbb{N}$ denote by $C^k(D)$ the algebra of real-valued functions on D that have continuous partial derivatives of orders $1,\ldots,k$. The natural topology of $C^k(D)$ is the topology of uniform convergence on compact subsets of D, of functions and all of their partials up to and including order k. In order to describe this topology with seminorms let $\mathfrak{D}^k$ denote the collection of all partial derivative operators of order not exceeding k; $\mathfrak{D}^k$ contains by convention in particular the zeroeth-order partial derivative $\Phi \mapsto \Phi$. Then set, for any compact subset $K \subset D$ and $\Phi \in C^k(D)$,

$$\|\Phi\|_{k,K} \stackrel{\text{def}}{=} \sup_{x\in K} \sup_{\partial\in\mathfrak{D}^k} |\partial\Phi(x)| \,.$$

These seminorms, one for every compact $K \subset D$, actually make for a metrizable topology: there is a sequence K_n of compact sets with $K_n \subset \mathring{K}_{n+1}$ whose interiors $\mathring{K}_n$ exhaust D; and

$$\rho(\Phi,\Psi) \stackrel{\text{def}}{=} \sum_n 2^{-n}\bigl(1 \wedge \|\Phi-\Psi\|_{n,K_n}\bigr)$$

is a metric defining the natural topology of $C^k(D)$, which is clearly much finer than the topology of uniform convergence on compacta.

Proposition A.2.11 *The polynomials are dense in $C^k(D)$ in this topology.*

Here is a terse sketch of the proof. Let K be a compact subset of D. There exists a compact $K' \subset D$ whose interior $\mathring{K}'$ contains K. Given $\Phi \in C^k(D)$, denote by Φ_σ the convolution of the heat kernel γ_{tI} with the product[11] $\mathring{K}' \cdot \Phi$. Since Φ and its partials are bounded on $\mathring{K}'$, the integral defining the convolution exists and defines a real-analytic function Φ_t. Some easy but space-consuming estimates show that all partials of Φ_t converge uniformly on K to the corresponding partials of Φ as $t \downarrow 0$: the real-analytic functions are dense in $C^k(D)$. Then of course so are the polynomials.

Topologies, Filters, Uniformities

A topology on a space S is a collection $\mathfrak{t}$ of subsets that contains the whole space S and the empty set $\emptyset$ and that is closed under taking finite intersections and arbitrary unions. The sets of $\mathfrak{t}$ are called the ***open sets*** or $\mathfrak{t}$-***open sets***. Their complements are the ***closed sets***. Every subset $A \subseteq S$ contains a largest open set, denoted by $\mathring{A}$ and called the $\mathfrak{t}$-***interior*** of A; and every $A \subseteq S$ is contained in a smallest closed set, denoted by $\bar{A}$ and called the $\mathfrak{t}$-***closure*** of. A subset $A \subset S$ is given the ***induced topology*** $\mathfrak{t}_A \stackrel{\text{def}}{=} \{A \cap U : U \in \mathfrak{t}\}$. For details see [56] and [35]. A ***filter*** on S is a collection $\mathfrak{F}$ of non-void subsets of S that is closed under taking finite intersections and arbitrary supersets. The ***tail filter*** of a sequence (x_n) is the collection of all sets that contain a whole tail $\{x_n : n \geq N\}$ of the sequence. The ***neighborhood filter*** $\mathfrak{V}(x)$ of a point $x \in S$ for the topology $\mathfrak{t}$ is the filter of all subsets that contain a $\mathfrak{t}$-open set containing x. The filter $\mathfrak{F}$ ***converges*** to x if $\mathfrak{F}$ ***refines*** $\mathfrak{V}(x)$, that is to say if $\mathfrak{F} \supset \mathfrak{V}(x)$. Clearly a sequence converges if and only if its tail filter does. By Zorn's lemma, every filter is contained in (refined by) an ***ultrafilter***, that is to say, in a filter that has no proper refinement.

Let $(S, \mathfrak{t}_S)$ and $(T, \mathfrak{t}_T)$ be topological spaces. A map $f : S \to T$ is ***continuous*** if the inverse image of every set in $\mathfrak{t}_T$ belongs to $\mathfrak{t}_S$. This is the case if and only if $\mathfrak{V}(x)$ refines $f^{-1}(\mathfrak{V}(f(x))$ at all $x \in S$.

The topology $\mathfrak{t}$ is ***Hausdorff*** if any two distinct points $x, x' \in S$ have non-intersecting neighborhoods V, V', respectively. It is completely regular if given $x \in E$ and $C \subset E$ closed one can find a continuous function that is zero on C and non-zero at x.

If the closure $\bar{U}$ is the whole ambient set S, then U is called $\mathfrak{t}$-***dense***. The topology $\mathfrak{t}$ is ***separable*** if S contains a countable $\mathfrak{t}$-dense set.

Exercise A.2.12 A filter $\mathfrak{U}$ on S is an ultrafilter if and only if for every $A \subseteq S$ either A or its complement A^c belongs to $\mathfrak{U}$. The following are equivalent: (i) every cover of S by open sets has a finite subcover; (ii) every collection of closed subsets with void intersection contains a finite subcollection whose intersection is void; (iii) every ultrafilter in S converges. In this case the topology is called ***compact***.

---

[11] $\mathring{K}'$ denotes both the set $\mathring{K}'$ and its indicator function – see convention A.1.5 on page 364.

Exercise A.2.13 (Tychonoff's Theorem) Let $E_\alpha, \mathfrak{t}_\alpha$, $\alpha \in A$, be topological spaces. The product topology $\mathfrak{t}$ on $E = \prod E_\alpha$ is the coarsest topology with respect to which all of the projections onto the factors E_α are continuous. The projection of an ultrafilter on E onto any of the factors is an ultrafilter there. Use this to prove Tychonoff's theorem: if the $\mathfrak{t}_\alpha$ are all compact, then so is $\mathfrak{t}$.

Exercise A.2.14 If $f : S \to T$ is continuous and $A \subset S$ is compact (in the induced topology, of course), then the forward image $f(A)$ is compact.

A topological space $(S, \mathfrak{t})$ is ***locally compact*** if every point has a basis of compact neighborhoods, that is to say, if every neighborhood of every point contains a compact neighborhood of that point. The ***one-point compactification*** S^Δ of $(S, \mathfrak{t})$ is obtained by adjoining one point, often denoted by Δ and called the ***point at infinity*** or the ***grave***, and declaring its neighborhood system to consist of the complements of the compact subsets of S. If S is already compact, then Δ is evidently an isolated point of $S^\Delta = S \dot\cup \{\Delta\}$.

A ***pseudometric*** on a set E is a function $d : E \times E \to \mathbb{R}_+$ that has $d(x,x) = 0$; is symmetric: $d(x,y) = d(y,x)$; and obeys the ***triangle inequality***: $d(x,z) \le d(x,y) + d(y,z)$. If $d(x,y) = 0$ implies that $x = y$, then d is a ***metric***. Let $\mathfrak{u}$ be a collection of pseudometrics on E. Another pseudometric d' is ***uniformly continuous*** with respect to $\mathfrak{u}$ if for every $\epsilon > 0$ there are $d_1, \ldots, d_k \in \mathfrak{u}$ and $\delta > 0$ such that

$$d_1(x,y) < \delta, \ldots, d_k(x,y) < \delta \implies d'(x,y) < \epsilon , \qquad \forall\, x,y \in E .$$

The ***saturation*** of $\mathfrak{u}$ consists of all pseudometrics that are uniformly continuous with respect to $\mathfrak{u}$. It contains in particular the pointwise sum and maximum of any two pseudometrics in $\mathfrak{u}$, and any positive scalar multiple of any pseudometric in $\mathfrak{u}$. A ***uniformity*** on E is simply a collection $\mathfrak{u}$ of pseudometrics that is saturated; a ***basis*** of $\mathfrak{u}$ is any subcollection $\mathfrak{u}_0 \subset \mathfrak{u}$ whose saturation equals $\mathfrak{u}$. The topology of $\mathfrak{u}$ is the topology $\mathfrak{t}_\mathfrak{u}$ generated by the open "pseudoballs" $B_{d,\epsilon}(x_0) \overset{\text{def}}{=} \{x \in E : d(x,x_0) < \epsilon\}$, $d \in \mathfrak{u}$, $\epsilon > 0$.

A map $f : E \to E'$ between uniform spaces $(E, \mathfrak{u})$ and $(E', \mathfrak{u}')$ is ***uniformly continuous*** if the pseudometric $(x,y) \mapsto d'\big(f(x), f(y)\big)$ belongs to $\mathfrak{u}$, for every $d' \in \mathfrak{u}'$. The composition of two uniformly continuous functions is obviously uniformly continuous again. The restrictions of the pseudometrics in $\mathfrak{u}$ to a fixed subset A of S clearly generate a uniformity on A, the ***induced uniformity***. A function on S is ***uniformly continuous on A*** if its restriction to A is uniformly continuous in this induced uniformity.

The filter $\mathfrak{F}$ on E is ***Cauchy*** if it contains arbitrarily small sets; that is to say, for every pseudometric $d \in \mathfrak{u}$ and every $\epsilon > 0$ there is an $F \in \mathfrak{F}$ with d-diam$(F) \overset{\text{def}}{=} \sup\{d(x,y) : x,y \in F\} < \epsilon$. The uniform space $(E, \mathfrak{u})$ is ***complete*** if every Cauchy filter $\mathfrak{F}$ converges. Every uniform space $(E, \mathfrak{u})$ has a ***Hausdorff completion***. This is a complete uniform space $(\overline{E}, \overline{\mathfrak{u}})$ whose topology $\mathfrak{t}_{\overline{\mathfrak{u}}}$ is Hausdorff, together with a uniformly continuous map $j : E \to \overline{E}$ such that the following holds: whenever $f : E \to Y$ is a uniformly continuous map into a Hausdorff complete uniform space Y, then there exists a unique uniformly continuous map $\overline{f} : \overline{E} \to Y$ such that $f = \overline{f} \circ j$. If a topology $\mathfrak{t}$ can be generated by some uniformity $\mathfrak{u}$, then it is ***uniformizable***; if $\mathfrak{u}$ has a basis consisting of a singleton d, then $\mathfrak{t}$ is ***pseudometrizable*** and ***metrizable*** if d is a metric; if $\mathfrak{u}$ and d can be chosen complete, then $\mathfrak{t}$ is ***completely (pseudo)metrizable***.

Exercise A.2.15 A Cauchy filter $\mathfrak{F}$ that has a convergent refinement converges. Therefore, if the topology of the uniformity $\mathfrak{u}$ is compact, then $\mathfrak{u}$ is complete.

A *compact* topology is generated by a unique uniformity: it is uniformizable in a unique way; if its topology has a countable basis, then it is completely pseudometrizable and completely metrizable if and only if it is also Hausdorff. A continuous function on a compact space and with values in a uniform space is uniformly continuous.

In this book two types of uniformity play a role. First there is the case that $\mathfrak{u}$ has a basis consisting of a single element d, usually a metric. The second instance is this: suppose $\mathcal{E}$ is a collection of real-valued functions on E. The ***$\mathcal{E}$-uniformity*** on E is the saturation of the collection of pseudometrics d_ϕ defined by

$$d_\phi(x,y) = \big|\phi(x) - \phi(y)\big| , \qquad \phi \in \mathcal{E} ,\ x,y \in E .$$

It is also called the uniformity generated by $\mathcal{E}$ and is denoted by $\mathfrak{u}[\mathcal{E}]$. We leave to the reader the following facts:

Lemma A.2.16 *Assume that $\mathcal{E}$ consists of bounded functions on some set E.*

(i) The uniformity generated by $\mathcal{E}$ coincides with the uniformity generated by the smallest uniformly closed algebra containing $\mathcal{E}$ and the constants. (ii) If $\mathcal{E}$ contains a countable uniformly dense set, then $\mathfrak{u}[\mathcal{E}]$ is pseudometrizable: it has a basis consisting of a single pseudometric d. If in addition $\overline{\mathcal{E}}$ separates the points of E, then d is a metric and $\mathfrak{t}_{\mathfrak{u}[\mathcal{E}]}$ is Hausdorff.

(iii) The Hausdorff completion of $(E, \mathfrak{u}[\mathcal{E}])$ is compact; it is the space $\overline{E}$ of the proof of theorem A.2.2 equipped with the uniformity generated by its continuous functions. If $\overline{\mathcal{E}}$ contains the constants, it equals $\widehat{E}$; otherwise it is the one-point compactification of $\widehat{E}$.

(iv) Let $A \subset E$ and let $f : A \to E'$ be a uniformly continuous[12] map to a complete uniform space $(E', \mathfrak{u}')$. Then $f(A)$ is relatively compact in $(E', \mathfrak{t}_{\mathfrak{u}'})$. Suppose $\mathcal{E}$ is an algebra or a vector lattice closed under chopping; then a real-valued function on A is uniformly continuous[12] if and only if it can be approximated uniformly on A by functions in $\mathcal{E} \oplus \mathbb{R}$, and an $\overline{\mathbb{R}}$-valued function is uniformly continuous if and only if it is the uniform limit (under the arctan metric $\overline{\rho}$!) of functions in $\mathcal{E} \oplus \mathbb{R}$.

[12] The uniformity of A is of course the one induced from $\mathfrak{u}[\mathcal{E}]$: it has the basis of pseudometrics $(x,y) \mapsto d_\phi(x,y) = \big|\phi(x) - \phi(y)\big|$, $d_\phi \in \mathfrak{u}[\mathcal{E}]$, $x, y \in A$, and is therefore the uniformity generated by the restrictions of the $\phi \in \mathcal{E}$ to A. The uniformity on $\mathbb{R}$ is of course given by the usual metric $\rho(r,s) \stackrel{\text{def}}{=} |r - s|$, the uniformity of the extended reals by the arctan metric $\overline{\rho}(r,s)$ – see item A.1.2.

Exercise A.2.17 A subset of a uniform space is called ***precompact*** if its image in the completion is relatively compact. A precompact subset of a complete uniform space is relatively compact.

Exercise A.2.18 Let (D,d) be a metric space. The distance of a point $x \in D$ from a set $F \subset D$ is $d(x,F) \stackrel{\text{def}}{=} \inf\{d(x,x') : x' \in F\}$. The ϵ-neighborhood of F is the set of all points whose distance from F is strictly less than ϵ; it evidently equals the union of all ϵ-balls with centers in F. A subset $K \subset D$ is called ***totally bounded*** if for every $\epsilon > 0$ there is a finite set $F_\epsilon \subset D$ whose ϵ-neighborhood contains K. Show that a subset $K \subseteq D$ is precompact if and only if it is totally bounded.

Semicontinuity

Let E be a topological space. The collection of bounded continuous real-valued functions on E is denoted by $C_b(E)$. It is a lattice algebra containing the constants. A real-valued function f on E is ***lower semicontinuous*** at $x \in E$ if $\liminf_{y\to x} f(y) \ge f(x)$; it is called ***upper semicontinuous*** at $x \in E$ if $\limsup_{y\to x} f(y) \le f(x)$. f is simply lower (upper) semicontinuous if it is lower (upper) semicontinuous at every point of E. For example, an open set is a lower semicontinuous function, and a closed set is an upper semicontinuous function.[13]

Lemma A.2.19 *Assume that the topological space E is completely regular.*
(a) For a bounded function f the following are equivalent:

- *(i) f is lower (upper) semicontinuous;*
- *(ii) f is the pointwise supremum of the continuous functions $\phi \le f$ (f is the pointwise infimum of the continuous functions $\phi \ge f$);*
- *(iii) $-f$ is upper (lower) semicontinuous;*
- *(iv) for every $r \in \mathbb{R}$ the set $[f > r]$ (the set $[f < r]$) is open.*

(b) Let $\mathcal{A}$ be a vector lattice of bounded continuous functions that contains the constants and generates the topology.[14] *Then:*

- *(i) If $U \subset E$ is open and $K \subset U$ compact, then there is a function $\phi \in \mathcal{A}$ with values in $[0,1]$ that equals 1 on K and vanishes outside U.*
- *(ii) Every bounded lower semicontinuous function h is the pointwise supremum of an increasingly directed subfamily $\mathcal{A}^h$ of $\mathcal{A}$.*

Proof. We leave (a) to the reader. (b) The sets of the form $[\phi > r]$, $\phi \in \mathcal{A}$, $r > 0$, clearly form a subbasis of the topology generated by $\mathcal{A}$. Since $[\phi > r] = [(\phi/r - 1) \vee 0 > 0]$, so do the sets of the form $[\phi > 0]$, $0 \le \phi \in \mathcal{A}$. A finite intersection of such sets is again of this form: $\bigcap_i [\phi_i > 0]$ equals $[\bigvee_i \phi_i > 0]$.

---

[13] S denotes both the set S and its indicator function – see convention A.1.5 on page 364.
[14] The ***topology generated by a collection Γ of functions*** is the coarsest topology with respect to which every $\gamma \in \Gamma$ is continuous. A net (x_α) converges to x in this topology if and only if $\gamma(x_\alpha) \to \gamma(x)$ for all $\gamma \in \Gamma$. Γ is said to define the given topology τ if the topology it generates coincides with τ; if τ is metrizable, this is the same as saying that a sequence x_n converges to x if and only if $\gamma(x_n) \to \gamma(x)$ for all $\gamma \in \Gamma$.

The sets of the form $[\phi > 0]$, $\phi \in \mathcal{A}_+$, thus form a basis of the topology generated by $\mathcal{A}$.

(i) Since K is compact, there is a finite collection $\{\phi_i\} \subset \mathcal{A}_+$ such that $K \subset \bigcup_i [\phi_i > 0] \subset U$. Then $\psi \stackrel{\text{def}}{=} \bigvee \phi_i$ vanishes outside U and is strictly positive on K. Let $r > 0$ be its minimum on K. The function $\phi \stackrel{\text{def}}{=} (\psi/r) \wedge 1$ of $\mathcal{A}_+$ meets the description of (i).

(ii) We start with the case that the lower semicontinuous function h is positive. For every $q > 0$ and $x \in [h > q]$ let $\phi_x^q \in \mathcal{A}$ be as provided by (i): $\phi_x^q(x) = 1$, and $\phi_x^q(x') = 0$ where $h(x') \le q$. Clearly $q \cdot \phi_x^q < h$. The finite suprema of the functions $q \cdot \phi_x^q \in \mathcal{A}$ form an increasingly directed collection $\mathcal{A}^h \subset \mathcal{A}$ whose pointwise supremum evidently is h. If h is not positive, we apply the foregoing to $h + \| h \|_\infty$. ∎

Separable Metric Spaces

Recall that a topological space E is ***metrizable*** if there exists a metric d that defines the topology in the sense that the neighborhood filter $\mathcal{V}(x)$ of every point $x \in E$ has a basis of ***d-balls*** $B_r(x) \stackrel{\text{def}}{=} \{x' : d(x, x') < r\}$ – then there are in general many metrics doing this. The next two results facilitate the measure theory on separable and metrizable spaces.

Lemma A.2.20 *Assume that E is separable and metrizable.*

(i) There exists a countably generated[6] *uniformly closed lattice algebra $\mathcal{U}[E]$ of bounded uniformly continuous functions that contains the constants, generates the topology,*[14] *and has in addition the property that every bounded lower semicontinuous function is the pointwise supremum of an increasing sequence in $\mathcal{U}[E]$, and every bounded upper semicontinuous function is the pointwise infimum of a decreasing sequence in $\mathcal{U}[E]$.*

(ii) Any increasingly (decreasingly) directed[2] *subset Φ of $C_b(E)$ contains a sequence that has the same pointwise supremum (infimum) as Φ.*

Proof. (i) Let d be a metric for E and $D = \{x_1, x_2, \ldots\}$ a countable dense subset. The collection Γ of bounded uniformly continuous functions $\gamma_{k,n} : x \mapsto kd(x, x_n) \wedge 1$, $x \in E$, $k, n \in \mathbb{N}$, is countable and generates the topology; indeed the open balls $[\gamma_{k,n} < 1/2]$ evidently form a basis of the topology. Let $\mathcal{A}$ denote the collection of finite $\mathbb{Q}$-linear combinations of 1 and finite products of functions in Γ. This is a countable algebra over $\mathbb{Q}$ containing the scalars whose uniform closure $\mathcal{U}[E]$ is both an algebra and a vector lattice (theorem A.2.2).

Let h be a lower semicontinuous function. Lemma A.2.19 provides an increasingly directed family $\mathcal{U}^h \subset \mathcal{U}[E]$ whose pointwise supremum is h; that it can be chosen countable follows from (ii).

(ii) Assume Φ is increasingly directed and has bounded pointwise supremum h. For every $\phi \in \Phi$, $x \in E$, and $n \in \mathbb{N}$ let $\psi_{\phi,x,n}$ be an element of $\mathcal{A}$ with $\psi_{\phi,x,n} \le \phi$ and $\psi_{\phi,x,n}(x) > \phi(x) - 1/n$. The collection $\mathcal{A}^h$ of these

$\psi_{\phi,x,n}$ is at most countable: $\mathcal{A}^h = \{\psi_1, \psi_2, \ldots\}$, and its pointwise supremum is h. For every n select a $\phi'_n \in \Phi$ with $\psi_n \le \phi'_n$. Then set $\phi_1 = \phi'_1$, and when $\phi_1, \ldots, \phi_n \in \Phi$ have been defined let ϕ_{n+1} be an element of Φ that exceeds $\phi_1, \ldots, \phi_n, \phi'_{n+1}$. Clearly $\phi_n \uparrow h$. ∎

Lemma A.2.21 *(a) Let $\mathcal{X}, \mathcal{Y}$ be metric spaces, $\mathcal{Y}$ compact, and suppose that $K \subset \mathcal{X} \times \mathcal{Y}$ is σ-compact and non-void. Then there is a Borel cross section; that is to say, there is a Borel map $\gamma : \mathcal{X} \to \mathcal{Y}$ "whose graph lies in K when it can:" when $x \in \pi_{\mathcal{X}}(K)$ then $\big(x, \gamma(x)\big) \in K$ – see figure A.15.*

(b) Let $\mathcal{X}, \mathcal{Y}$ be separable metric spaces, $\mathcal{X}$ locally compact and $\mathcal{Y}$ compact, and suppose $G : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ is a continuous function. There exists a Borel function $\gamma : \mathcal{X} \to \mathcal{Y}$ such that for all $x \in \mathcal{X}$

$$\sup\Big\{G(x,y) : y \in \mathcal{Y}\Big\} = G\big(x, \gamma(x)\big) .$$

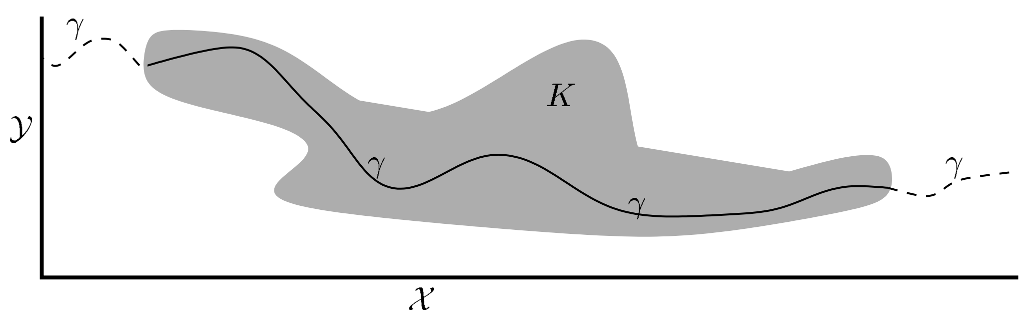


Figure A.15 The Cross Section Lemma

Proof. (a) To start with, consider the case that $\mathcal{Y}$ is the unit interval I and that K is compact. Then $\gamma^K(x) \stackrel{\text{def}}{=} \inf\{t : (x,t) \in K\} \wedge 1$ defines a lower semicontinuous function from $\mathcal{X}$ to I with $\big(x, \gamma(x)\big) \in K$ when $x \in \pi_{\mathcal{X}}(K)$. If K is σ-compact, then there is an increasing sequence $\big(K_n\big)$ of compacta with union K. The cross sections γ^{K_n} give rise to the decreasing and ultimately constant sequence (γ_n) defined inductively by $\gamma_1 \stackrel{\text{def}}{=} \gamma^{K_1}$,

$$\gamma_{n+1} \stackrel{\text{def}}{=} \begin{cases} \gamma_n & \text{on } [\gamma_n < 1], \\ \gamma^{K_{n+1}} & \text{on } [\gamma_n = 1]. \end{cases}$$

Clearly $\gamma \stackrel{\text{def}}{=} \inf \gamma_n$ is Borel, and $\big(x, \gamma(x)\big) \in K$ when $x \in \pi_{\mathcal{X}}(K)$. If $\mathcal{Y}$ is not the unit interval, then we use the universality A.2.22 of the Cantor set $C \subset I$: it provides a continuous surjection $\phi : C \to \mathcal{Y}$. Then $K' \stackrel{\text{def}}{=} (\phi \times id_{\mathcal{X}})^{-1}(K)$ is a σ-compact subset of $I \times \mathcal{X}$, there is a Borel function $\gamma^{K'} : \mathcal{X} \to C$ whose restriction to $\pi_{\mathcal{X}}(K') = \pi_{\mathcal{X}}(K)$ has its graph in K', and $\gamma \stackrel{\text{def}}{=} \phi \circ \gamma^{K'}$ is the desired Borel cross section.

(b) Set $\sigma(x) \stackrel{\text{def}}{=} \sup\big\{G(x,y) : y \in \mathcal{Y}\big\}$. Because of the compactness of $\mathcal{Y}$, σ is a continuous function on $\mathcal{X}$ and $K \stackrel{\text{def}}{=} \{(x,y) : G(x,y) = \sigma(x)\}$ is a σ-compact subset of $\mathcal{X} \times \mathcal{Y}$ with $\mathcal{X}$-projection $\mathcal{X}$. Part (a) furnishes γ. ∎

Exercise A.2.22 (Universality of the Cantor Set) For every compact metric space $\mathcal{Y}$ there exists a continuous map from the Cantor set onto $\mathcal{Y}$.

Exercise A.2.23 Let F be a Hausdorff space and E a subset whose induced topology can be defined by a complete metric ρ. Then E is a $\mathcal{G}_\delta$-set; that is to say, there is a sequence of open subsets of F whose intersection is E.

Exercise A.2.24 Let (P, d) be a separable complete metric space. There exists a compact metric space $\widehat{P}$ and a homeomorphism j of P onto a subset of $\widehat{P}$. j can be chosen so that $j(P)$ is a dense $\mathcal{G}_\delta$-set and a $\mathcal{K}_{\sigma\delta}$-set of $\widehat{P}$.

Topological Vector Spaces

A real vector space $\mathcal{V}$ together with a topology on it is a ***topological vector space*** if the linear and topological structures are compatible in this sense: the maps $(f, g) \mapsto f + g$ from $\mathcal{V} \times \mathcal{V}$ to $\mathcal{V}$ and $(r, f) \mapsto r \cdot f$ from $\mathbb{R} \times \mathcal{V}$ to $\mathcal{V}$ are continuous.

A subset B of the topological vector space $\mathcal{V}$ is ***bounded*** if it is ***absorbed*** by any neighborhood V of zero; this means that there exists a scalar λ so that $B \subset \lambda V \stackrel{\text{def}}{=} \{\lambda v : v \in V\}$.

The main examples of topological vector spaces concerning us in this book are the spaces $\mathcal{L}^p$ and L^p for $0 \le p \le \infty$ and the spaces $C_0(E)$ and $C(E)$ of continuous functions. We recall now a few common notions that should help the reader navigate their topologies.

A set $V \subset \mathcal{V}$ is ***convex*** if for any two scalars λ_1, λ_2 with absolute value less than 1 and sum 1 and for any two points $v_1, v_2 \in V$ we have $\lambda_1 v_1 + \lambda_2 v_2 \in V$. A topological vector space $\mathcal{V}$ is ***locally convex*** if the neighborhood filter at zero (and then at any point) has a basis of convex sets. The examples above all have this feature, except the spaces $\mathcal{L}^p$ and L^p when $0 \le p < 1$.

Theorem A.2.25 *Let $\mathcal{V}$ be a locally convex topological vector space.*
(i) Let $A, B \subset \mathcal{V}$ be convex, non–void, and disjoint, A closed and B either open or compact. There exist a continuous linear functional $x^* : \mathcal{V} \to \mathbb{R}$ and a number c so that $x^*(a) \le c$ for all $a \in A$ and $x^*(b) > c$ for all $b \in B$.
(ii) **(Hahn–Banach)** *A linear functional defined and continuous on a linear subspace of $\mathcal{V}$ has an extension to a continuous linear functional on all of $\mathcal{V}$.*
(iii) A convex subset of $\mathcal{V}$ is closed if and only if it is weakly closed.
(iv) **(Alaoglu)** *An equicontinuous set of linear functionals on $\mathcal{V}$ is relatively weak*–compact. (See the Answers for these terms and a proof.)*

A.2.26 Gauges It is easy to see that a topological vector space admits a collection Γ of gauges $\lceil\!\lceil \ \rceil\!\rceil : \mathcal{V} \to \mathbb{R}_+$ that define the topology in the sense that $f_n \to f$ if and only if $\lceil\!\lceil f - f_n \rceil\!\rceil \to 0$ for all $\lceil\!\lceil \ \rceil\!\rceil \in \Gamma$. This is the same as saying that the "balls"

$$B_\epsilon(0) \stackrel{\text{def}}{=} \{f : \lceil\!\lceil f \rceil\!\rceil < \epsilon\}, \qquad \lceil\!\lceil \ \rceil\!\rceil \in \Gamma, \ \epsilon > 0,$$

form a basis of the neighborhood system at 0 and implies that $\lceil\!\lceil rf \rceil\!\rceil \xrightarrow[r\to 0]{} 0$ for all $f \in \mathcal{V}$ and all $\lceil\!\lceil \ \rceil\!\rceil \in \Gamma$. There are always many such gauges. Namely, let $\{V_n\}$ be a decreasing sequence of neighborhoods of 0 with $V_0 = \mathcal{V}$. Then

$$\lceil\!\lceil f \rceil\!\rceil \stackrel{\text{def}}{=} \Big(\inf\{n : \ f \in V_n\}\Big)^{-1}$$

will be a gauge. If the V_n form basis of neighborhoods at zero, then Γ can be taken to be the singleton $\{\lceil\!\lceil \ \rceil\!\rceil\}$ above. With a little more effort it can be shown that there are continuous gauges defining the topology that are ***subadditive***: $\lceil\!\lceil f+g \rceil\!\rceil \le \lceil\!\lceil f \rceil\!\rceil + \lceil\!\lceil g \rceil\!\rceil$. For such a gauge, $dist(f,g) \stackrel{\text{def}}{=} \lceil\!\lceil f-g \rceil\!\rceil$ defines a translation-invariant pseudometric, a metric if and only if $\mathcal{V}$ is Hausdorff. From now on the word ***gauge*** will mean a continuous subadditive gauge.

A locally convex topological vector space whose topology can be defined by a complete metric is a ***Fréchet space***. Here are two examples that recur throughout the text:

Examples A.2.27 (i) Suppose E is a locally compact separable metric space and F is a Fréchet space with translation-invariant metric ρ (visualize $\mathbb{R}$). Let $C_F(E)$ denote the vector space of all continuous functions from E to F. The ***topology of uniform convergence on compacta*** on $C_F(E)$ is given by the following collection of gauges, one for every compact set $K \subset E$,

$$\lceil\!\lceil \phi \rceil\!\rceil_K \stackrel{\text{def}}{=} \sup\big\{\rho\big(\phi(x)\big) : x \in K\big\}\ ,\ \phi : E \to F\ . \tag{A.2.1}$$

It is Fréchet. Indeed, a cover by compacta K_n with $K_n \subset \mathring{K}_{n+1}$ gives rise to the gauge

$$\phi \mapsto \textstyle\sum_n \lceil\!\lceil \phi \rceil\!\rceil_{K_n} \wedge 2^{-n}\ , \tag{A.2.2}$$

which in turn gives rise to a complete metric for the topology of $C_F(E)$. If F is separable, then so is $C_F(E)$.

(ii) Suppose that $E = \mathbb{R}_+$, but consider the space $\mathscr{D}_F$, the ***path space***, of functions $\phi : \mathbb{R}_+ \to F$ that are right-continuous and have a left limit at every instant $t \in \mathbb{R}_+$. Inasmuch as such a càdlàg path is bounded on every bounded interval, the supremum in (A.2.1) is finite, and (A.2.2) again describes the Fréchet topology of uniform convergence on compacta. But now this topology is not separable in general, even when F is as simple as $\mathbb{R}$. The indicator functions $\phi_t \stackrel{\text{def}}{=} 1_{[0,t)}$, $0 < t < 1$, have $\lceil\!\lceil \phi_s - \phi_t \rceil\!\rceil_{[0,1]} = 1$, yet they are uncountable in number. ∎

With every convex neighborhood V of zero there comes the ***Minkowski functional*** $\|f\| \stackrel{\text{def}}{=} \inf\{|r| : f/r \in V\}$. This continuous gauge is both subadditive and ***absolute-homogeneous***: $\|r \cdot f\| = |r| \cdot \|f\|$ for $f \in \mathcal{V}$ and $r \in \mathbb{R}$. An absolute-homogeneous subadditive gauge is a ***seminorm***. If $\mathcal{V}$ is locally convex, then their collection defines the topology. Prime examples of spaces whose topology is defined by a single seminorm are the spaces $\mathcal{L}^p$ and L^p for $1 \le p \le \infty$, and $C_0(E)$.

Exercise A.2.28 Suppose that $\mathcal{V}$ has a countable basis at 0. Then $B \subset \mathcal{V}$ is bounded if and only if for one, and then every, continuous gauge $\lceil\!\lceil\ \rceil\!\rceil$ on $\mathcal{V}$ that defines the topology

$$\sup\{\lceil\!\lceil \lambda \cdot f \rceil\!\rceil : f \in B\} \xrightarrow[\lambda \to 0]{} 0 .$$

Exercise A.2.29 Let $\mathcal{V}$ be a topological vector space with a countable base at 0, and $\lceil\!\lceil\ \rceil\!\rceil$ and $\lceil\!\lceil\ \rceil\!\rceil'$ two gauges on $\mathcal{V}$ that define the topology – they need not be subadditive nor continuous except at 0. There exists an increasing right-continuous function $\Phi : \mathbb{R}_+ \to \overline{\mathbb{R}}_+$ with $\Phi(r) \xrightarrow[r \to 0]{} 0$ such that $\lceil\!\lceil f \rceil\!\rceil' \le \Phi(\lceil\!\lceil f \rceil\!\rceil)$ for all $f \in \mathcal{V}$.

A.2.30 Quasinormed Spaces In some contexts it is more convenient to use the homogeneity of the $\|\ \|_{L^p}$ on L^p rather than the subadditivity of the $\lceil\!\lceil\ \rceil\!\rceil_{L^p}$. In order to treat Banach spaces and spaces L^p simultaneously one uses the notion of a ***quasinorm*** on a vector space E. This is a function $\|\ \| : E \to \mathbb{R}_+$ such that

$$\|x\| = 0 \iff x = 0 \quad \text{and} \quad \|r\cdot x\| = |r| \cdot \|x\| \qquad \forall\, r \in \mathbb{R}, x \in E .$$

A topological vector space is ***quasinormed*** if it is equipped with a quasinorm $\|\ \|$ that defines the topology, i.e., such that $x_n \xrightarrow[n\to\infty]{} x$ if and only if $\|x_n - x\| \xrightarrow[n\to\infty]{} 0$. If $(E, \|\ \|_E)$ and $(F, \|\ \|_F)$ are quasinormed topological vector spaces and $u : E \to F$ is a continuous linear map between them, then the ***size of u*** is naturally measured by the number

$$\|u\| = \|u\|_{L(E,F)} \stackrel{\text{def}}{=} \sup\big\{ \|u(x)\|_F : x \in E, \|x\|_E \le 1 \big\} .$$

A subadditive quasinorm clearly is a seminorm; so is an absolute-homogeneous gauge.

Exercise A.2.31 Let $\mathcal{V}$ be a vector space equipped with a seminorm $\|\ \|$. The set $N \stackrel{\text{def}}{=} \{x \in \mathcal{V} : \|x\| = 0\}$ is a vector subspace and coincides with the closure of $\{0\}$. On the quotient $\dot{\mathcal{V}} \stackrel{\text{def}}{=} \mathcal{V}/N$ set $\|\dot{x}\| \stackrel{\text{def}}{=} \|x\|$. This does not depend on the representative x in the equivalence class $\dot{x} \in \dot{\mathcal{V}}$ and makes $(\dot{\mathcal{V}}, \|\ \|)$ a normed space. The transition from $(\mathcal{V}, \|\ \|)$ to $(\dot{\mathcal{V}}, \|\ \|)$ is such a standard operation that it is sometimes not mentioned, that $\mathcal{V}$ and $\dot{\mathcal{V}}$ are identified, and that reference is made to "the norm" $\|\ \|$ on $\mathcal{V}$.

A.2.32 Weak Topologies Let $\mathcal{V}$ be a vector space and $\mathcal{M}$ a collection of linear functionals $\mu : \mathcal{V} \to \mathbb{R}$. This gives rise to two topologies. One is the topology $\sigma(\mathcal{V}, \mathcal{M})$ on $\mathcal{V}$, the coarsest topology with respect to which every functional $\mu \in \mathcal{M}$ is continuous; it makes $\mathcal{V}$ into a locally convex topological vector space. The other is $\sigma(\mathcal{M}, \mathcal{V})$, the topology on $\mathcal{M}$ of pointwise convergence on $\mathcal{V}$. For an example assume that $\mathcal{V}$ is already a topological vector space under some topology τ and $\mathcal{M}$ consists of all τ-continuous linear functionals on $\mathcal{V}$, a vector space usually called the ***dual of*** $\mathcal{V}$ and denoted by $\mathcal{V}^*$. Then $\sigma(\mathcal{V}, \mathcal{V}^*)$ is called by analysts the ***weak topology*** on $\mathcal{V}$ and $\sigma(\mathcal{V}^*, \mathcal{V})$ the ***weak*** * ***topology*** on $\mathcal{V}^*$. When $\mathcal{V} = C_0(E)$ and $\mathcal{M} = \mathfrak{P}^* \subset C_0(E)^*$ probabilists like to call the latter the topology of weak convergence – as though life weren't confusing enough already!

Exercise A.2.33 If $\mathcal{V}$ is given the topology $\sigma(\mathcal{V},\mathcal{M})$, then the dual of $\mathcal{V}$ coincides with the vector space generated by $\mathcal{M}$.

The Minimax Theorem, Lemmas of Gronwall and Kolmogoroff

Lemma A.2.34 (Ky–Fan) *Let K be a compact convex subset of a topological vector space and $\mathcal{H}$ a family of upper semicontinuous concave numerical functions on K. Assume that the functions of $\mathcal{H}$ do not take the value $+\infty$ and that any convex combination of any two functions in $\mathcal{H}$ majorizes another function of $\mathcal{H}$. If every function $h \in \mathcal{H}$ is nonnegative at some point $k_h \in K$, then there is a common point $k \in K$ at which all of the functions $h \in \mathcal{H}$ take a nonnegative value.*

Proof. We argue by contradiction and assume that the conclusion fails. Then the convex compact sets $[h \geq 0]$, $h \in \mathcal{H}$, have void intersection, and there will be finitely many $h \in \mathcal{H}$, say $h_1, \ldots, h_N$, with

$$\bigcap_{n=1}^{N} [h_n \geq 0] = \emptyset \, . \tag{A.2.3}$$

Let the collection $\{h_1, \ldots, h_N\}$ be chosen so that N is minimal. Since $[h \geq 0] \neq \emptyset$ for every $h \in \mathcal{H}$, we must have $N \geq 2$. The compact convex set

$$K' \overset{\text{def}}{=} \bigcap_{n=3}^{N} [h_n \geq 0] \subset K$$

is contained in $[h_1 < 0] \cup [h_2 < 0]$ (if $N = 2$ it equals K). Both h_1 and h_2 take nonnegative values on K'; indeed, if h_1 did not, then h_2 could be struck from the collection, and vice versa, in contradiction to the minimality of N.

Let us see how to proceed in a very simple situation: suppose K is the unit interval $I = [0,1]$ and $\mathcal{H}$ consists of affine functions. Then K' is a closed subinterval I' of I, and h_1 and h_2 take their positive maxima at one of the endpoints of it, evidently not in the same one. In particular, I' is not degenerate. Since the open sets $[h_1 < 0]$ and $[h_2 < 0]$ together cover the interval I', but neither does by itself, there is a point $\xi \in I'$ at which both h_1 and h_2 are strictly negative; ξ evidently lies in the interior of I'. Let $\eta = \max\bigl\{h_1(\xi), h_2(\xi)\bigr\}$. Any convex combination $h' = r_1 h_1 + r_2 h_2$ of h_1, h_2 will at ξ have a value less than $\eta < 0$. It is clearly possible to choose $r_1, r_2 \geq 0$ with sum 1 so that h' has at the left endpoint of I' a value in $(\eta/2, 0)$. The affine function h' is then evidently strictly less than zero on all of I'. There exists by assumption a function $h \in \mathcal{H}$ with $h \leq h'$; it can replace the pair $\{h_1, h_2\}$ in equation (A.2.3), which is in contradiction to the minimality of N. The desired result is established in the simple case that K is the unit interval and $\mathcal{H}$ consists of affine functions.

First note that the set $[h_i > -\infty]$ is convex, as the increasing union of the convex sets $\bigcup_{k\in\mathbb{N}}[h_i \geq -k]$, $i = 1, 2$. Thus

$$K'_0 \stackrel{\text{def}}{=} K' \cap [h_1 > -\infty] \cap [h_2 > -\infty]$$

is convex. Next observe that there is an $\epsilon > 0$ such that the open set $[h_1 + \epsilon < 0] \cup [h_2 + \epsilon < 0]$ still covers K'. For every $k \in K'_0$ consider the affine function

$$a_k : t \mapsto - \ \Big(t \cdot \big(h_1(k) + \epsilon\big) \ + \ (1-t) \cdot \big(h_2(k) + \epsilon\big)\Big) \ ,$$

i.e., $$a_k(t) \stackrel{\text{def}}{=} - \ \Big(t \cdot h_1(k) + (1-t) \cdot h_2(k)\Big) - \epsilon \ ,$$

on the unit interval I. Every one of them is nonnegative at some point of I; for instance, if $k \in [h_1 + \epsilon < 0]$, then $\lim_{t\to 1} a_k(t) = -\big(h_1(k) + \epsilon\big) > 0$. An easy calculation using the concavity of h_i shows that a convex combination $ra_k + (1-r)a_{k'}$ majorizes $a_{rk+(1-r)k'}$. We can apply the first part of the proof and conclude that there exists a $\tau \in I$ at which every one of the functions a_k is nonnegative. This reads

$$h'(k) \stackrel{\text{def}}{=} \tau \cdot h_1(k) + (1-\tau) \cdot h_2(k) \leq -\epsilon < 0 \qquad\qquad k \in K'_0 \ .$$

Now τ is not the right endpoint 1; if it were, then we would have $h_1 < -\epsilon$ on K'_0, and a suitable convex combination $rh_1 + (1-r)h_2$ would majorize a function $h \in \mathcal{H}$ that is strictly negative on K; this then could replace the pair $\{h_1, h_2\}$ in equation (A.2.3). By the same token $\tau \neq 0$. But then h' is strictly negative on all of K and there is an $h \in \mathcal{H}$ majorized by h', which can then replace the pair $\{h_1, h_2\}$. In all cases we arrive at a contradiction to the minimality of N. ∎

Lemma A.2.35 (Gronwall's Lemma) *Let $\phi : [0,\infty] \to [0,\infty)$ be an increasing function satisfying*

$$\phi(t) \leq A(t) + \int_0^t \phi(s)\, \eta(s)\, ds \qquad\qquad t \geq 0 \ ,$$

where $\eta : [0,\infty) \to \mathbb{R}$ is positive and Borel, and $A : [0,\infty) \to [0,\infty)$ is increasing. Then

$$\phi(t) \leq A(t) \cdot \exp\Big(\int_0^t \eta(s)\, ds\Big) \ , \qquad\qquad t \geq 0 \ .$$

Proof. To start with, assume that ϕ is right-continuous and A constant.

Set $$H(t) \stackrel{\text{def}}{=} \exp\Big(\int_0^t \eta(s)\, ds\Big) \ , \quad \text{fix an } \ \epsilon > 0 \ ,$$

and set $$t_0 \stackrel{\text{def}}{=} \inf\big\{s : \ \phi(s) \geq \big(A + \epsilon\big) \cdot H(s)\big\} \ .$$

Then $$\begin{aligned} \phi(t_0) &\leq A + (A+\epsilon) \int_0^{t_0} H(s)\, \eta(s) ds \\ &= A + (A+\epsilon)\big(H(t_0) - H(0)\big) = (A+\epsilon)H(t_0) - \epsilon \\ &< (A+\epsilon)H(t_0) \ . \end{aligned}$$

Since this is a strict inequality and ϕ is right-continuous, $\phi(t_0') \le (A+\epsilon)H(t_0')$ for some $t_0' > t_0$. Thus $t_0 = \infty$, and $\phi(t) \le (A+\epsilon)H(t)$ for all $t \ge 0$. Since $\epsilon > 0$ was arbitrary, $\phi(t) \le AH(t)$. In the general case fix a t and set

$$\psi(s) = \inf\bigl\{\phi(\tau \wedge t) : \tau > s\bigr\} .$$

ψ is right-continuous, equals ϕ at all but countably many points of $[0,t]$, and satisfies $\psi(\tau) \le A(t) + \int_0^\tau \psi(s)\,\eta(s)\,ds$ for $\tau \ge 0$. The first part of the proof applies and yields $\phi(t) \le \psi(t) \le A(t) \cdot H(t)$. ∎

Exercise A.2.36 Let $x : [0,\infty] \to [0,\infty)$ be an increasing function satisfying

$$x_\mu \le C + \max_{\rho=p,q} \Bigl(\int_0^\mu (A + Bx_\lambda)^\rho \, d\lambda\Bigr)^{1/\rho} , \qquad \mu \ge 0 ,$$

for some $1 \le p \le q < \infty$ and some constants $A, B > 0$. Then there exist constants $\alpha \le 2(A/B + C)$ and $\beta \le \max_{\rho=p,q}(2B)^\rho/\rho$ such that $x_\lambda \le \alpha e^{\beta\lambda}$ for all $\lambda > 0$.

Lemma A.2.37 (Kolmogorov) *Let U be an open subset of $\mathbb{R}^d$, and let*

$$\{X_u \;:\; u \in U\}$$

be a family of functions,[15] *all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and having values in the same complete metric space (E, ρ). Assume that $\omega \mapsto \rho(X_u(\omega), X_v(\omega))$ is measurable for any two $u, v \in U$ and that there exist constants $p, \beta > 0$ and $C < \infty$ so that*

$$\mathbb{E}\left[\rho(X_u, X_v)^p\right] \le C \cdot |u - v|^{d+\beta} \quad \text{for } \; u, v \in U . \tag{A.2.4}$$

Then there exists a family $\{X_u' \;:\; u \in U\}$ of the same description which in addition has the following properties: (i) $X_.'$ is a modification of $X_.$, meaning that $\mathbb{P}[X_u \ne X_u'] = 0$ for every $u \in U$; and (ii) for every single $\omega \in \Omega$ the map $u \mapsto X_u(\omega)$ from U to E is continuous. In fact there exists, for every $\alpha > 0$, a subset $\Omega_\alpha \in \mathcal{F}$ with $\mathbb{P}[\Omega_\alpha] > 1 - \alpha$ such that the family

$$\{u \mapsto X_u'(\omega) \;:\; \omega \in \Omega_\alpha\}$$

of E-valued functions is equicontinuous on U and uniformly equicontinuous on every compact subset K of U; that is to say, for every $\epsilon > 0$ there is a $\delta > 0$ independent of $\omega \in \Omega_\alpha$ such that $|u - v| < \delta$ implies $\rho(X_u'(\omega), X_v'(\omega)) < \epsilon$ for all $u, v \in K$ and all $\omega \in \Omega_\alpha$. (In fact, $\delta = \delta_{K;\alpha,p,C,\beta}(\epsilon)$ depends only on the indicated quantities.)

[15] Not necessarily measurable for the Borels of (E, ρ).

Exercise A.2.38 (Ascoli–Arzelà) Let $K \subseteq U$, $\epsilon \mapsto \delta(\epsilon)$ an increasing function from $(0,1)$ to $(0,1)$, and $C \subseteq E$ compact. The collection $\mathcal{K}(\delta(\cdot))$ of paths $x : K \to C$ satisfying

$$|u-v| \le \delta(\epsilon) \implies \rho(x_u, x_v) \le \epsilon$$

is compact in the topology of uniform convergence of paths; conversely, a compact set of continuous paths is uniformly equicontinuous and the union of their ranges is relatively compact. Therefore, if E happens to be compact, then the set of paths $\{X'_{\boldsymbol{\cdot}}(\omega) : \omega \in \Omega_\alpha\}$ of lemma A.2.37 is relatively compact in the topology of uniform convergence on compacta.

Proof of A.2.37. Instead of the customary euclidean norm $|\ |_2$ we may and shall employ the sup-norm $|\ |_\infty$. For $n \in \mathbb{N}$ let U_n be the collection of vectors in U whose coordinates are of the form $k2^{-n}$, with $k \in \mathbb{Z}$ and $|k2^{-n}| < n$. Then set $U_\infty = \bigcup_n U_n$. This is the set of ***dyadic-rational points*** in U and is clearly in U. To start with we investigate the random variables[15] X_u, $u \in U_\infty$.

Let $0 < \lambda < \beta/p$.[16] If $u, v \in U_n$ are nearest neighbors, that is to say if $|u-v| = 2^{-n}$, then Chebyscheff's inequality and (A.2.4) give

$$\begin{aligned}\mathbb{P}\Big(\big[\rho(X_u, X_v) > 2^{-\lambda n}\big]\Big) &\le 2^{p\lambda n} \cdot \mathbb{E}\left[\rho(X_u, X_v)^p\right] \\ &\le C \cdot 2^{p\lambda n} \cdot 2^{-n(d+\beta)} = C \cdot 2^{(p\lambda-\beta-d)n} .\end{aligned}$$

Now a point $u \in U_n$ has less than 3^d nearest neighbors v in U_n, and there are less than $(2n2^n)^d$ points in U_n. Consequently

$$\mathbb{P}\Big[\bigcup_{\substack{u,v \in U_n \\ |u-v| = 2^{-n}}} \big[\rho(X_u, X_v) > 2^{-\lambda n}\big]\Big]$$

$$\le C \cdot 2^{(p\lambda-\beta-d)n} \cdot (6n)^d \cdot 2^{nd} = C \cdot (6n)^d 2^{-(\beta-p\lambda)n} .$$

Since $2^{-(\beta-p\lambda)} < 1$, these numbers are summable over n. Given $\alpha > 0$, we can find an integer N_α depending only[16] on C, β, p such that the set

$$\mathcal{N}_\alpha = \bigcup_{n \ge N_\alpha} \bigcup_{\substack{u,v \in U_n \\ |u-v| = 2^{-n}}} \big[\rho(X_u, X_v) > 2^{-\lambda n}\big]$$

has $\mathbb{P}[\mathcal{N}_\alpha] < \alpha$. Its complement $\Omega_\alpha = \Omega \setminus \mathcal{N}_\alpha$ has measure

$$\mathbb{P}[\Omega_\alpha] > 1 - \alpha .$$

A point $\omega \in \Omega_\alpha$ has the property that whenever $n > N_\alpha$ and $u, v \in U_n$ have distance $|u-v| = 2^{-n}$ then

$$\rho\big(X_u(\omega), X_v(\omega)\big) \le 2^{-\lambda n} . \tag{$*$}$$

[16] For instance, $\lambda = \beta/2p$.

Let K be a compact subset of U, and let us start on the last claim by showing that $\{u \mapsto X_u(\omega) : \omega \in \Omega_\alpha\}$ *is uniformly equicontinuous on* $U_\infty \cap K$. To this end let $\epsilon > 0$ be given. There is an $n_0 > N_\alpha$ such that

$$2^{-\lambda n_0} < \epsilon \cdot (1 - 2^{-\lambda}) \cdot 2^{\lambda - 1} .$$

Note that this number depends only on α, ϵ, and the constants of inequality (A.2.4). Next let n_1 be so large that 2^{-n_1} is smaller than the distance of K from the complement of U, and let n_2 be so large that K is contained in the centered ball (the shape of a box) of diameter (side) $2n_2$. We respond to ϵ by setting

$$n = n_0 \vee n_1 \vee n_2 \quad \text{and} \quad \delta = 2^{-n} .$$

Clearly δ was manufactured from $\epsilon, K\alpha, p, C, \beta$ alone. We shall show that $|u - v| < \delta$ implies $\rho\big(X_u(\omega), X_v(\omega)\big) \le \epsilon$ for all $\omega \in \Omega_\alpha$ and $u, v \in K \cap U_\infty$. Now if $u, v \in U_\infty$, then there is a "mesh-size" $m \ge n$ such that both u and v belong to U_m. Write $u = u_m$ and $v = v_m$. There exist $u_{m-1}, v_{m-1} \in U_{m-1}$

$$\text{with} \qquad |u_m - u_{m-1}|_\infty \le 2^{-m} \ , \ |v_m - v_{m-1}| \le 2^{-m} \ ,$$

$$\text{and} \qquad |u_{m-1} - v_{m-1}| \le |u_m - v_m| \ .$$

Namely, if $u = (k_1 2^{-m}, \ldots, k_d 2^{-m})$ and $v = (\ell_1 2^{-m}, \ldots, \ell_d 2^{-m})$, say, we add or subtract 1 from an odd k_δ according as $k_\delta - \ell_\delta$ is strictly positive or negative; and if k_δ is even or if $k_\delta - \ell_\delta = 0$, we do nothing. Then we go through the same procedure with v. Since $\delta \le 2^{n_1}$, the (box-shaped) balls with radius 2^{-n} about u_m, v_m lie entirely inside U, and then so do the points u_{m-1}, v_{m-1}. Since $\delta \le 2^{-n_2}$, they actually belong to U_{m-1}. By the same token there exist $u_{m-2}, v_{m-2} \in U_{m-2}$

$$\text{with} \qquad |u_{m-1} - u_{m-2}| \le 2^{-m-1} \ , \ |v_{m-1} - v_{m-2}| \le 2^{-m-1} \ ,$$

$$\text{and} \qquad |u_{m-2} - v_{m-2}| \le |u_{m-1} - v_{m-1}| \ .$$

Continue on. Clearly $u_n = v_n$. In view of $(*)$ we have, for $\omega \in \Omega_\alpha$,

$$\begin{aligned}
\rho\big(X_u(\omega), X_v(\omega)\big) &\le \rho\big(X_u(\omega), X_{u_{m-1}}(\omega)\big) + \ldots + \rho\big(X_{u_{n+1}}(\omega), X_{u_n}(\omega)\big) \\
&\quad + \rho\big(X_{u_n}(\omega), X_{v_n}(\omega)\big) \\
&\quad + \rho\big(X_{v_n}(\omega), X_{v_{n+1}}(\omega)\big) + \ldots + \rho\big(X_{v_{m-1}}(\omega), X_v(\omega)\big) \\
&\le 2^{-m\lambda} + 2^{-(m-1)\lambda} + \ldots + 2^{-(n+1)\lambda} \\
&\quad + 0 \\
&\quad + 2^{-(n+1)\lambda} + \ldots + 2^{-(m-1)\lambda} + 2^{-m\lambda} \\
&\le 2 \sum_{i=n+1}^{\infty} (2^{-\lambda})^i = 2 \cdot 2^{-\lambda n} \cdot \big(2^{-\lambda}/(1 - 2^{-\lambda})\big) \le \epsilon \ .
\end{aligned}$$

To summarize: the family $\{u \mapsto X_u(\omega) : \omega \in \Omega_\alpha\}$ of E-valued functions is uniformly equicontinuous on every relatively compact subset K of U_∞.

Now set $\Omega_0 = \bigcup_n \Omega_{1/n}$. For every $\omega \in \Omega_0$, the map $u \mapsto X_u(\omega)$ is uniformly continuous on relatively compact subsets of U_∞ and thus has a unique continuous extension to all of U. Namely, for arbitrary $u \in U$ we set

$$X'_u(\omega) \stackrel{\text{def}}{=} \lim \big\{X_q(\omega) : U_\infty \ni q \to u\big\} , \qquad \omega \in \Omega_0 .$$

This limit exists, since $\big\{X_q(\omega) : U_\infty \ni q \to u\big\}$ is Cauchy and E is complete. In the points ω of the negligible set $\mathcal{N} = \Omega_0^c = \bigcap_\alpha \mathcal{N}_\alpha$ we set X'_u equal to some fixed point $x_0 \in E$. From inequality (A.2.4) it is plain that $X'_u = X_u$ almost surely. The resulting selection meets the description; it is, for instance, an easy exercise to check that the δ given above as a response to K and ϵ serves as well to show the uniform equicontinuity of the family $\big\{u \mapsto X'_u(\omega) : \omega \in \Omega_\alpha\big\}$ of functions on K. ∎

Exercise A.2.39 The proof above shows that there is a negligible set $\mathcal{N}$ such that, for every $\omega \notin \mathcal{N}$, $q \mapsto X_q(\omega)$ is uniformly continuous on every bounded set of dyadic rationals in U.

Exercise A.2.40 Assume that the set U, while possibly not open, is contained in the closure of its interior. Assume further that the family $\{X_u : u \in U\}$ satisfies merely, for some fixed $p > 0, \beta > 0$:

$$\limsup_{U \ni v, v' \to u} \frac{\mathbb{E}\left[\rho(X_v, X_{v'})^p\right]}{\left|v - v'\right|^{d+\beta}} < \infty \qquad \forall\, u \in U .$$

Again a modification can be found that is continuous in $u \in U$ for all $\omega \in \Omega$.

Exercise A.2.41 Any two continuous modifications X'_u, X''_u are indistinguishable in the sense that the set $\{\omega : \exists u \in U \text{ with } X'_u(\omega) \neq X''_u(\omega)\}$ is negligible.

Lemma A.2.42 (Taylor's Formula) *Suppose $D \subset \mathbb{R}^d$ is open and convex and $\Phi : D \to \mathbb{R}$ is n-times continuously differentiable. Then*[17]

$$\begin{aligned}\Phi(\boldsymbol{z} + \boldsymbol{\Delta}) - \Phi(\boldsymbol{z}) &= \Phi_{;\eta}(\boldsymbol{z}) \cdot \Delta^\eta + \int_0^1 (1-\lambda)\, \Phi_{;\eta\theta}\big(\boldsymbol{z} + \lambda\boldsymbol{\Delta}\big)\, d\lambda \cdot \Delta^\eta \Delta^\theta \\ &= \sum_{\nu=1}^{n-1} \frac{1}{\nu!} \Phi_{;\eta_1 \ldots \eta_\nu}(\boldsymbol{z}) \cdot \Delta^{\eta_1} \cdots \Delta^{\eta_\nu} \\ &\quad + \int_0^1 \frac{(1-\lambda)^{n-1}}{(n-1)!}\, \Phi_{;\eta_1 \ldots \eta_n}\big(\boldsymbol{z} + \lambda\boldsymbol{\Delta}\big)\, d\lambda \cdot \Delta^{\eta_1} \cdots \Delta^{\eta_n}\end{aligned}$$

for any two points $\boldsymbol{z}, \boldsymbol{z} + \boldsymbol{\Delta} \in D$.

---

[17] Subscripts after semicolons denote partial derivatives, e.g.,

$$\Phi_{;\eta} \stackrel{\text{def}}{=} \frac{\partial \Phi}{\partial x^\eta} \quad \text{and} \quad \Phi_{;\eta\theta} \stackrel{\text{def}}{=} \frac{\partial^2 \Phi}{\partial x^\eta \partial x^\theta} .$$

Summation over repeated indices in opposite positions is implied by Einstein's convention, which is adopted throughout.

A.2.43 Let $1 \le p < \infty$, and set $n = [p]$ and $\epsilon = p - n$. Then for $z, \delta \in \mathbb{R}$

$$|z+\delta|^p = |z|^p + \sum_{\nu=1}^{n-1} \binom{p}{\nu} |z|^{p-\nu} (\operatorname{sgn} z)^\nu \cdot \delta^\nu$$
$$+ \int_0^1 n(1-\lambda)^{n-1} \binom{p}{n} |z+\lambda\delta|^\epsilon \big(\operatorname{sgn}(z+\lambda\delta)\big)^n \, d\lambda \cdot \delta^n ,$$

where $$\binom{p}{\nu} \stackrel{\text{def}}{=} \frac{p(p-1)\cdots(p-\nu+1)}{\nu!} \quad \text{and} \quad \operatorname{sgn} z \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{if } z = 0 \\ -1 & \text{if } z < 0 . \end{cases}$$

Differentiation

Definition A.2.44 (Big O and Little o) *Let N, D, s be real-valued functions depending on the same arguments $u, v, \ldots$. One says*

$$\text{“}N = O(D) \text{ as } s \to 0\text{” if } \quad \lim_{\delta \to 0} \sup \left\{ \frac{N(u,v,\ldots)}{D(u,v,\ldots)} : s(u,v,\ldots) \le \delta \right\} < \infty ,$$

$$\text{“}N = o(D) \text{ as } s \to 0\text{” if } \quad \lim_{\delta \to 0} \sup \left\{ \frac{N(u,v,\ldots)}{D(u,v,\ldots)} : s(u,v,\ldots) \le \delta \right\} = 0 .$$

If $D = s$, one simply says “$N = O(D)$” or “$N = o(D)$,” respectively.

This nifty convention eases many arguments, including the usual definition of differentiability, also called ***Fréchet differentiability***:

Definition A.2.45 *Let F be a map from an open subset U of a seminormed space E to another seminormed space S. F is **differentiable at a point** $u \in U$ if there exists a bounded*[18] *linear operator $DF[u] : E \to S$, written $\eta \mapsto DF[u]{\cdot}\eta$ and called the **derivative of F at u**, such that the **remainder** RF, defined by*

$$F(v) - F(u) = DF[u]{\cdot}(v-u) + RF[v;u] ,$$

has $$\|RF[v;u]\|_S = o(\|v-u\|_E) \quad \text{as} \quad v \to u .$$

*If F is differentiable at all points of U, it is called differentiable on U or simply **differentiable**; if in that case $u \mapsto DF[u]$ is continuous in the operator norm, then F is **continuously differentiable**; if in that case $\|RF[v;u]\|_S = o(\|v-u\|_E)$,*[19] *$F$ is **uniformly differentiable**.*

*Next let $\mathfrak{F}$ be a whole family of maps from U to S, all differentiable at $u \in U$. Then $\mathfrak{F}$ is **equidifferentiable** at u if $\sup\{\|RF[v;u]\|_S : F \in \mathfrak{F}\}$ is*

---

[18] This means that the operator norm $\| DF[u] \|_{E \to S} \stackrel{\text{def}}{=} \sup\{\|DF[u] \cdot \eta\|_S : \|\eta\|_E \le 1\}$ is finite.

[19] That is to say $\sup\big\{\|RF[v;u]\|_S / \|v-u\|_E : \|v-u\|_E \le \delta\big\} \xrightarrow[\delta \to 0]{} 0$, which explains the word “uniformly.”

$o(\|v-u\|_E)$ as $v \to u$, and ***uniformly equidifferentiable*** *if the previous supremum is $o(\|v-u\|_E)$ as $\|v-u\|_E \to 0$.*

Exercise A.2.46 (i) Establish the usual rules of differentiation. (ii) If F is differentiable at u, then $\|F(v)-F(u)\|_S = O(\|v-u\|_E)$ as $v \to u$. (iii) Suppose now that U is open and convex and F is differentiable on U. Then F is Lipschitz with constant L if and only if $\|DF[u]\|_{E\to S}$ is bounded; and in that case

$$L = \sup{}_u \|DF[u]\|_{E\to S} .$$

(iv) If F is continuously differentiable on U, then it is uniformly differentiable on every relatively compact subset of U; furthermore, there is this representation of the remainder:

$$RF[v;u] = \int_0^1 \big(DF[u+\lambda(v-u)] - DF[u]\big)\cdot(v-u)\, d\lambda .$$

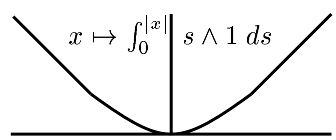
Exercise A.2.47 For differentiable $f : \mathbb{R} \to \mathbb{R}$, $Df[x]$ is multiplication by $f'(x)$.

Now suppose $F = F[u,x]$ is a differentiable function of two variables, $u \in U$ and $x \in V \subset X$, X being another seminormed space. This means of course that F is differentiable on $U \times V \subset E \times X$. Then $DF[u,x]$ has the form

$$\begin{aligned} DF[u,x] \cdot \begin{pmatrix} \eta \\ \xi \end{pmatrix} &= \Big(D_1F[u,x], D_2F[u,x]\Big)\cdot \begin{pmatrix} \eta \\ \xi \end{pmatrix} \\ &= D_1F[u,x]\cdot\eta + D_2F[u,x]\cdot\xi , \quad \eta \in E, \xi \in X , \end{aligned}$$

where $D_1F[u,x]$ is the ***partial in the u-direction*** and $D_2F[u,x]$ the ***partial in the x-direction***. In particular, when the arguments u,x are real we often write $F_{;1} = F_{;u} \stackrel{\text{def}}{=} D_1F$, $F_{;2} = F_{;x} \stackrel{\text{def}}{=} D_2F$, etc.

Example A.2.48 — of Trouble Consider a differentiable function f on the line of not more than linear growth, for example $f(x) \stackrel{\text{def}}{=} \int_0^{|x|} s \wedge 1\, ds$. One hopes that composition with f, which takes ϕ to $F[\phi] \stackrel{\text{def}}{=} f \circ \phi$, might define a Fréchet differentiable map F from $L^p(\mathbb{P})$ to itself. Alas, it does not. Namely, if $DF[0]$ exists, it must equal multiplication by $f'(0)$, which in the example above equals zero – but then $RF(0,\phi) = F[\phi] - F[0] - DF[0]\cdot\phi = F[\phi] = f \circ \phi$ does not go to zero faster in $L^p(\mathbb{P})$-mean $\| \ \|_p$ than does $\|\phi - 0\|_p$ – simply take ϕ through a sequence of indicator functions converging to zero in $L^p(\mathbb{P})$-mean.



F is, however, differentiable, even uniformly so, as a map from $L^p(\mathbb{P})$ to $L^{p^\circ}(\mathbb{P})$ for any p° *strictly smaller* than p, whenever the derivative f' is continuous and bounded. Indeed, by Taylor's formula of order one (see lemma A.2.42 on page 387)

$$\begin{aligned} F[\psi] = F[\phi] &+ f'(\phi)\cdot(\psi-\phi) \\ &+ \int_0^1 \Big[f'\big(\phi+\sigma(\psi-\phi)\big) - f'\big(\phi\big)\Big]\, d\sigma \cdot (\psi-\phi) , \end{aligned}$$

whence, with Hölder's inequality and $1/p^\circ = 1/p + 1/r$ defining r,

$$\|RF[\psi;\phi]\|_{p^\circ} \le \Big\| \int_0^1 \Big[f'\big(\phi + \sigma(\psi - \phi)\big) - f'\big(\phi\big)\Big]\, d\sigma \Big\|_r \cdot \|\psi - \phi\|_p \,.$$

The first factor tends to zero as $\|\psi - \phi\|_p \to 0$, due to theorem A.8.6, so that $\|RF[\psi;\phi]\|_{p^\circ} = o\big(\|\psi - \phi\|_p\big)$. Thus F is uniformly differentiable as a map from $L^p(\mathbb{P})$ to $L^{p^\circ}(\mathbb{P})$, with[20] $DF[\phi] = f' \circ \phi$.

Note the following phenomenon: the derivative $\xi \mapsto DF[\phi]\cdot\xi$ is actually a continuous linear map from $L^p(\mathbb{P})$ to itself whose operator norm is bounded independently of ϕ by $\|f'\| \stackrel{\text{def}}{=} \sup_x |f'(x)|$. It is just that the remainder $RF[\psi;\phi]$ is $o(\|\psi - \phi\|_p)$ only if it is measured with the weaker seminorm $\|\ \|_{p^\circ}$. The example gives rise to the following notion:

Definition A.2.49 *(i) Let $(S, \|\ \|_S)$ be a seminormed space and $\|\ \|_S^\circ \le \|\ \|_S$ a weaker seminorm. A map F from an open subset $\mathcal{U}$ of a seminormed space $(E, \|\ \|_E)$ to S is $\|\ \|_S^\circ$-weakly differentiable at $u \in \mathcal{U}$ if there exists a bounded[18] linear map $DF[u] : E \to S$ such that*

$$F[v] = F[u] + DF[u]\cdot(v - u) + RF[u;v] \qquad \forall\, v \in \mathcal{U}\,,$$

with $\quad \|RF[u;v]\|_S^\circ = o\big(\|v - u\|_E\big)$ *as* $v \to u$ *, i.e.,* $\dfrac{\|RF[u;v]\|_S^\circ}{\|v - u\|_E} \xrightarrow[v \to u]{} 0\,.$

(ii) Suppose that S comes equipped with a family $\mathfrak{N}^\circ$ of seminorms $\|\ \|_S^\circ \le \|\ \|_S$ such that $\|x\|_S = \sup\{\|x\|_S^\circ : \|\ \|_S^\circ \in \mathfrak{N}^\circ\}$ $\forall\, x \in S$. If F is $\|\ \|_S^\circ$-weakly differentiable at $u \in \mathcal{U}$ for every $\|\ \|_S^\circ \in \mathfrak{N}^\circ$, then we call F ***weakly differentiable at*** *u. If F is weakly differentiable at every $u \in \mathcal{U}$, it is simply called weakly differentiable; if, moreover, the decay of the remainder is independent of $u, v \in U$:*

$$\sup\Big\{ \frac{\|RF[u;v]\|_S^\circ}{\delta} : u, v \in U\,,\ \|v - u\|_E < \delta \Big\} \xrightarrow[\delta \to 0]{} 0$$

for every $\|\ \|_S^\circ \in \mathfrak{N}^\circ$, then F is ***uniformly weakly differentiable on*** *U.*

Here is a reprise of the calculus for this notion:

Exercise A.2.50 (a) The linear operator $DF[u]$ of (i) if extant is unique, and $F \to DF[u]$ is linear. To say that F is weakly differentiable means that F is, for every $\|\ \|_S^\circ \in \mathfrak{N}^\circ$, Fréchet differentiable as a map from $(E, \|\ \|_E)$ to $(S, \|\ \|_S^\circ)$ and has a derivative that is continuous as a linear operator from $(E, \|\ \|_E)$ to $(S, \|\ \|_S)$. (b) Formulate and prove the product rule and the chain rule for weak differentiability. (c) Show that if F is $\|\ \|_S^\circ$-weakly differentiable, then for all $u, v \in E$

$$\|F[v] - F[u]\|_S^\circ \le \sup\big\{ \|DF[u]\| : u \in E \big\} \cdot \|v - u\|_E \,.$$

[20] $DF[\phi]\cdot\xi = f' \circ \phi \cdot \xi$. In other words, $DF[\phi]$ is multiplication by $f' \circ \phi$.

A.3 Measure and Integration

σ-Algebras

A ***measurable space*** is a set F equipped with a σ-algebra $\mathcal{F}$ of subsets of F. A ***random variable*** is a map f whose domain is a measurable space $(F,\mathcal{F})$ and which takes values in another measurable space $(G,\mathcal{G})$. It is understood that a random variable f is ***measurable***: the inverse image $f^{-1}(G_0)$ of every set $G_0 \in \mathcal{G}$ belongs to $\mathcal{F}$. If there is need to specify which σ-algebra on the domain is meant, we say "f is ***measurable on*** $\mathcal{F}$" and write $f \in \mathcal{F}$. If we want to specify both σ-algebras involved, we say "f is $\mathcal{F}/\mathcal{G}$-measurable" and write $f \in \mathcal{F}/\mathcal{G}$. If $G = \mathbb{R}$ or $G = \mathbb{R}^n$, then it is understood that $\mathcal{G}$ is the σ-algebra of Borel sets (see below). A random variable is ***simple*** if it takes only finitely many different values.

The intersection of any collection of σ-algebras is a σ-algebra. Given some property P of σ-algebras, we may therefore talk about the ***σ-algebra generated by*** P: it is the intersection of all σ-algebras having P. We assume here that there is at least one σ-algebra having P, so that the collection whose intersection is taken is not empty – the σ-algebra of all subsets will usually do. Given a collection Φ of functions on F with values in measurable spaces, the ***σ-algebra generated by Φ*** is the smallest σ-algebra on which every function $\phi \in \Phi$ is measurable. For instance, if F is a topological space, there are the σ-algebra $\mathcal{B}^*(F)$ of ***Baire sets*** and the σ-algebra $\mathcal{B}^\bullet(F)$ of ***Borel sets***. The former is the smallest σ-algebra on which all continuous real-valued functions are measurable, and the latter is the generally larger σ-algebra generated by the open sets.[13] Functions measurable on $\mathcal{B}^*(F)$ or $\mathcal{B}^\bullet(F)$ are called ***Baire functions*** or ***Borel functions***, respectively.

Exercise A.3.1 (i) On a metrizable space the Baire and Borel σ-algebras coincide, and so the Baire functions and the Borel functions agree. In particular, on $\mathbb{R}^n$ and on the path spaces $\mathscr{C}^n$ or the Skorohod spaces $\mathscr{D}^n$, $n = 1, 2, \ldots$, the Baire functions and the Borel functions coincide.
(ii) Consider a measurable space $(F,\mathcal{F})$ and a topological space G equipped with its Baire σ-algebra $\mathcal{B}^*(G)$. If a sequence (f_n) of $\mathcal{F}/\mathcal{B}^*(G)$-measurable maps converges pointwise on F to a map $f : F \to G$, then f is again $\mathcal{F}/\mathcal{G}$-measurable.
(iii) The conclusion generally fails if $\mathcal{B}^*(G)$ is replaced by the Borel σ-algebra $\mathcal{B}^\bullet(G)$.
(iv) An finitely generated algebra $\mathcal{A}$ of sets is generated by its finite collection of ***atoms***; these are the sets in $\mathcal{A}$ that have no proper non–void subset belonging to $\mathcal{A}$.

Sequential Closure

Inasmuch as the permanence property under pointwise limits of sequences exhibited in exercise A.3.1 (ii) is the main merit of the notions of σ-algebra and $\mathcal{F}/\mathcal{G}$-measurability, it deserves a bit of study of its own:

A collection $\mathcal{B}$ of functions defined on some set E and having values in a topological space is called ***sequentially closed*** if the limit of any pointwise convergent sequence in $\mathcal{B}$ belongs to $\mathcal{B}$ as well. In most of the applications of this notion the functions in $\mathcal{B}$ are considered ***numerical***, i.e., they are allowed to take values in the extended reals $\overline{\mathbb{R}}$. For example, the collection of $\mathcal{F}/\mathcal{G}$-measurable random variables above, the collection of $[\![\ ]\!]^*$-measurable processes, and the collection of $[\![\ ]\!]^*$-measurable *sets* each are sequentially closed.

The intersection of any family of sequentially closed collections of functions on E plainly is sequentially closed. If $\mathcal{E}$ is any collection of functions, then there is thus a smallest sequentially closed collection $\mathcal{E}^\sigma$ of functions containing $\mathcal{E}$, to wit, the intersection of *all* sequentially closed collections containing $\mathcal{E}$. $\mathcal{E}^\sigma$ can be constructed by transfinite induction as follows. Set $\mathcal{E}_0 \stackrel{\text{def}}{=} \mathcal{E}$. Suppose that $\mathcal{E}_\alpha$ has been defined for all ordinals $\alpha < \beta$. If β is the successor of α, then define $\mathcal{E}_\beta$ to be the set of all functions that are limits of a sequence in $\mathcal{E}_\alpha$; if β is not a successor, then set $\mathcal{E}_\beta \stackrel{\text{def}}{=} \bigcup_{\alpha<\beta} \mathcal{E}_\alpha$. Then $\mathcal{E}_\beta = \mathcal{E}^\sigma$ for all β that exceed the first uncountable ordinal $\aleph_1$.

It is reasonable to call $\mathcal{E}^\sigma$ the ***sequential closure*** or ***sequential span*** of $\mathcal{E}$. If $\mathcal{E}$ is considered as a collection of numerical (real-valued) functions, and if this point must be emphasized, we shall denote the sequential closure by $\mathcal{E}^\sigma_{\overline{\mathbb{R}}}$ ($\mathcal{E}^\sigma_{\mathbb{R}}$). It will generally be clear from the context which is meant.

Exercise A.3.2 (i) Every $f \in \mathcal{E}^\sigma$ is contained in the sequential closure of a countable subcollection of $\mathcal{E}$. (ii) $\mathcal{E}$ is called ***σ-finite*** if it contains a countable subcollection whose pointwise supremum is everywhere strictly positive. Show: if $\mathcal{E}$ is a ring of sets or a vector lattice closed under chopping or an algebra of bounded functions, then $\mathcal{E}$ is σ-finite if and only if $1 \in \mathcal{E}^\sigma$. (iii) The collection of real-valued functions in $\mathcal{E}^\sigma_{\overline{\mathbb{R}}}$ coincides with $\mathcal{E}^\sigma_{\mathbb{R}}$.

Lemma A.3.3 *(i) If $\mathcal{E}$ is a vector space, algebra, or vector lattice of real-valued functions, then so is its sequential closure $\mathcal{E}^\sigma_{\mathbb{R}}$. (ii) If $\mathcal{E}$ is an algebra of bounded functions or a vector lattice closed under chopping, then $\mathcal{E}^\sigma_{\mathbb{R}}$ is both. Furthermore, if $\mathcal{E}$ is σ-finite, then the collection $\mathcal{E}^\sigma_e$ of sets in $\mathcal{E}^\sigma$ is the σ-algebra generated by $\mathcal{E}$, and $\mathcal{E}^\sigma$ consists precisely of the functions measurable on $\mathcal{E}^\sigma_e$.*

Proof. (i) Let $*$ stand for $+, -, \cdot, \vee, \wedge$, etc. Suppose $\mathcal{E}$ is closed under $*$.

The collection $$\mathcal{E}^* \stackrel{\text{def}}{=} \big\{ f : \ f * \phi \in \mathcal{E}^\sigma \quad \forall\, \phi \in \mathcal{E} \big\} \tag{$*$}$$
then contains $\mathcal{E}$, and it is sequentially closed. Thus it contains $\mathcal{E}^\sigma$. This shows that

the collection $$\mathcal{E}^{**} \stackrel{\text{def}}{=} \big\{ g : \ f * g \in \mathcal{E}^\sigma \quad \forall\, f \in \mathcal{E}^\sigma \big\} \tag{$**$}$$

contains $\mathcal{E}$. This collection is evidently also sequentially closed, so it contains $\mathcal{E}^\sigma$. That is to say, $\mathcal{E}^\sigma$ is closed under $*$ as well.

(ii) The constant 1 belongs to $\mathcal{E}^\sigma$ (exercise A.3.2). Therefore $\mathcal{E}^\sigma_e$ is not merely a σ-ring but a σ-algebra. Let $f \in \mathcal{E}^\sigma$. The set [13] $[f > 0] = \lim_{n\to\infty} 0 \vee \big((n \cdot f) \wedge 1\big)$, being the limit of an increasing bounded sequence, belongs to $\mathcal{E}^\sigma_e$. We conclude that for every $r \in \mathbb{R}$ and $f \in \mathcal{E}^\sigma$ the set [13]

$[f > r] = [f - r > 0]$ belongs to $\mathcal{E}_e^\sigma$: f is measurable on $\mathcal{E}_e^\sigma$. Conversely, if f is measurable on $\mathcal{E}_e^\sigma$, then it is the limit of the functions

$$\sum_{|\nu| \le 2^n} \nu 2^{-n} \left[\nu 2^{-n} < f \le (\nu+1) 2^{-n}\right]$$

in $\mathcal{E}^\sigma$ and thus belongs to $\mathcal{E}^\sigma$. Lastly, since every $\phi \in \mathcal{E}$ is measurable on $\mathcal{E}_e^\sigma$, $\mathcal{E}_e^\sigma$ contains the σ-algebra $\mathcal{E}^\Sigma$ generated by $\mathcal{E}$; and since the $\mathcal{E}^\Sigma$-measurable functions form a sequentially closed collection containing $\mathcal{E}$, $\mathcal{E}_e^\sigma \subset \mathcal{E}^\Sigma$. ∎

Theorem A.3.4 (The Monotone Class Theorem) *Let $\mathcal{V}$ be a collection of real-valued functions on some set that is closed under pointwise limits of increasing or decreasing sequences – this makes it a **monotone class**. Assume further that $\mathcal{V}$ forms a real vector space and contains the constants. With any subcollection $\mathcal{M}$ of bounded functions that is closed under multiplication – a **multiplicative class** – $\mathcal{V}$ then contains every real-valued function measurable on the σ-algebra $\mathcal{M}^\Sigma$ generated by $\mathcal{M}$.*

Proof. The family $\mathcal{E}$ of all finite linear combinations of functions in $\mathcal{M} \cup \{1\}$ is an algebra of bounded functions and is contained in $\mathcal{V}$. Its uniform closure $\overline{\mathcal{E}}$ is contained in $\mathcal{V}$ as well. For if $\mathcal{E} \ni f_n \to \overline{f}$ uniformly, we may without loss of generality assume that $\|\overline{f} - f_n\|_\infty < 2^{-n}/4$. The sequence $f_n - 2^{-n} \in \mathcal{E}$ then converges increasingly to $\overline{f}$. $\overline{\mathcal{E}}$ is a vector lattice (theorem A.2.2). Let $\mathcal{E}^{\uparrow\downarrow}$ denote the smallest collection of functions that contains $\mathcal{E}$ and is closed under pointwise limits of monotone sequences; it is evidently contained in $\mathcal{V}$. We see as in $(*)$ and $(**)$ above that $\mathcal{E}^{\uparrow\downarrow}$ is a vector lattice; namely, the collections $\mathcal{E}^*$ and $\mathcal{E}^{**}$ from the proof of lemma A.3.3 are closed under limits of monotone sequences. Since $\lim f_n = \sup_N \inf_{n>N} f_n$, $\mathcal{E}^{\uparrow\downarrow}$ is sequentially closed. If f is measurable on $\mathcal{M}^\Sigma$, it is evidently measurable on $\mathcal{E}^\Sigma = \mathcal{E}_e^\sigma$ and thus belongs to $\mathcal{E}^\sigma \subset \mathcal{E}^{\uparrow\downarrow} \subset \mathcal{V}$ (lemma A.3.3). ∎

Exercise A.3.5 (The Complex Bounded Class Theorem) Let $\mathcal{V}$ be a complex vector space of complex-valued functions on some set, and assume that $\mathcal{V}$ contains the constants and is closed under taking limits of bounded pointwise convergent sequences. With any subfamily $\mathcal{M} \subset \mathcal{V}$ that is closed under multiplication and complex conjugation – a ***complex multiplicative class*** – $\mathcal{V}$ contains every bounded complex-valued function that is measurable on the σ-algebra $\mathcal{M}^\Sigma$ generated by $\mathcal{M}$. In consequence, if two σ-additive measures of totally finite variation agree on the functions of $\mathcal{M}$, then they agree on $\mathcal{M}^\Sigma$.

Exercise A.3.6 On a topological space E the class of Baire functions is the sequential closure of the class $C_b(E)$ of bounded continuous functions. If E is completely regular, then the class of Borel functions is the sequential closure of the set of differences of lower semicontinuous functions.

Exercise A.3.7 Suppose that $\mathcal{E}$ is a self-confined vector lattice closed under chopping or an algebra of bounded functions on some set E (see exercise A.2.6). Let us denote by $\mathcal{E}_{00}^\sigma$ the smallest collection of functions on f that is closed under taking pointwise limits of bounded $\mathcal{E}$-confined sequences. Show: (i) $f \in \mathcal{E}_{00}^\sigma$ if and only if $f \in \mathcal{E}^\sigma$ is bounded and $\mathcal{E}$-confined; (ii) $\mathcal{E}_{00}^\sigma$ is both a vector lattice closed under chopping and an algebra.

Measures and Integrals

A σ-additive measure on the σ-algebra $\mathcal{F}$ is a function $\mu : \mathcal{F} \to \mathbb{R}$ [21] that satisfies $\mu\big(\bigcup_n A_n\big) = \sum_n \mu\big(A_n\big)$ for every disjoint sequence (A_n) in $\mathcal{F}$. Σ-algebras have no *raison d'être* but for the σ-additive measures that live on them. However, rare is the instance that a measure appears on a σ-algebra. Rather, measures come naturally as linear functionals on some small space $\mathcal{E}$ of functions (Radon measures, Haar measure) or as set functions on a ring $\mathcal{A}$ of sets (Lebesgue measure, probabilities). They still have to undergo a lengthy extension procedure before their domain contains the σ-algebra generated by $\mathcal{E}$ or $\mathcal{A}$ and before they can integrate functions measurable on that.

Set Functions A ***ring of sets*** on a set F is a collection $\mathcal{A}$ of subsets of F that is closed under taking relative complements and finite unions, and then under taking finite intersections. A ring is an ***algebra*** if it contains the whole ambient set F, and a ***δ-ring*** if it is closed under taking countable intersections (if both, it is a ***σ-algebra*** or ***σ-field***). A *measure* on the ring $\mathcal{A}$ is a σ-additive function $\mu : \mathcal{A} \to \mathbb{R}$ of finite variation. The additivity means that $\mu(A + A') = \mu(A) + \mu(A')$ for [13] $A, A', A + A' \in \mathcal{A}$. The σ-additivity means that $\mu\big(\bigcup_n A_n\big) = \sum_n \mu\big(A_n\big)$ for every disjoint sequence (A_n) of sets in $\mathcal{A}$ whose union A happens to belong to $\mathcal{A}$. In the presence of finite additivity this is equivalent with σ-continuity: $\mu(A_n) \to 0$ for every decreasing sequence (A_n) in $\mathcal{A}$ that has void intersection. The additive set function $\mu : \mathcal{A} \to \mathbb{R}$ has ***finite variation*** on $A \subset F$ if

$$\mid \mu \mid (A) \stackrel{\text{def}}{=} \sup\{\mu(A') - \mu(A'') \ : \ A', A'' \in \mathcal{A}\,,\ A' + A'' \le A\}$$

is finite. To say that μ has finite variation means that $\mid\mu\mid(A) < \infty$ for all $A \in \mathcal{A}$. The function $\mid\mu\mid : \mathcal{A} \to \overline{\mathbb{R}}_+$ then is a positive σ-additive measure on $\mathcal{A}$, called the ***variation*** of μ. μ has ***totally finite variation*** if $\mid\mu\mid(F) < \infty$. A σ-additive set function on a σ-algebra automatically has totally finite variation. Lebesgue measure on the finite unions of intervals $(a, b]$ is an example of a measure that appears naturally as a set function on a ring of sets.

Radon Measures are examples of measures that appear naturally as linear functionals on a space of functions. Let E be a locally compact Hausdorff space and $C_{00}(E)$ the set of continuous functions with compact support. A Radon measure is simply a linear functional $\mu : C_{00}(E) \to \mathbb{R}$ that is bounded on order-bounded (confined) sets.

Elementary Integrals The previous two instances of measures look so disparate that they are often treated quite differently. Yet a little teleological thinking reveals that they fit into a common pattern. Namely, while measuring sets is a pleasurable pursuit, integrating functions surely is what measure

[21] For numerical measures, i.e., measures that are allowed to take their values in the extended reals $\overline{\mathbb{R}}$, see exercise A.3.27.

theory ultimately is all about. So, given a measure μ on a ring $\mathcal{A}$ we immediately extend it by linearity to the linear combinations of the sets in $\mathcal{A}$, thus obtaining a linear functional on functions. Call their collection $\mathcal{E}[\mathcal{A}]$. This is the family of step functions ϕ over $\mathcal{A}$, and the linear extension is the natural one: $\mu(\phi)$ is the sum of the products height–of–step times μ-size–of–step.

In both instances we now face a linear functional $\mu : \mathcal{E} \to \mathbb{R}$. If μ was a Radon measure, then $\mathcal{E} = C_{00}(E)$; if μ came as a set function on $\mathcal{A}$, then $\mathcal{E} = \mathcal{E}[\mathcal{A}]$ and μ is replaced by its linear extension. In both cases the pair $(\mathcal{E}, \mu)$ has the following properties:

(i) *$\mathcal{E}$ is an algebra and vector lattice closed under chopping.* The functions in $\mathcal{E}$ are called ***elementary integrands***.

(ii) *μ is **σ-continuous***: $\mathcal{E} \ni \phi_n \downarrow 0$ pointwise implies $\mu(\phi_n) \to 0$.

(iii) *μ has **finite variation***: for all $\phi \geq 0$ in $\mathcal{E}$

$$\left\bracevert \mu \right\bracevert(\phi) \stackrel{\text{def}}{=} \sup\left\{|\mu(\psi)| : \psi \in \mathcal{E}\,,\ |\psi| \leq \phi\right\}$$

is finite; in fact, $\left\bracevert \mu \right\bracevert$ extends to a σ-continuous positive[22] linear functional on $\mathcal{E}$, the ***variation*** of μ. We shall call such a pair $(\mathcal{E}, \mu)$ an elementary integral.

(iv) Actually, all elementary integrals that we meet in this book have a ***σ-finite domain*** $\mathcal{E}$ (exercise A.3.2). This property facilitates a number of arguments. We shall therefore subsume the requirement of σ-finiteness on $\mathcal{E}$ in the definition of an ***elementary integral*** $(\mathcal{E}, \mu)$.

Extension of Measures and Integration The reader is no doubt familiar with the way Lebesgue succeeded in 1905[23] to extend the length function on the ring of finite unions of intervals to many more sets, and with Caratheodory's generalization to positive σ-additive set functions μ on arbitrary rings of sets. The main tools are the inner and outer measures μ_* and μ^*.

Once the measure is extended there is still quite a bit to do before it can integrate functions. In 1918 the French mathematician Daniell noticed that many of the arguments used in the extension procedure for the set function and again in the integration theory of the extension are the same. He discovered a way of melding the extension of a measure and its integration theory into one procedure. This saves labor and has the additional advantage of being applicable in more general circumstances, such as the stochastic integral. We give here a short overview. This will furnish both notation and motivation for the main body of the book. For detailed treatments see for example [9] and [12]. The reader not conversant with Daniell's extension procedure can actually find it in all detail in chapter 3, if he takes Ω to consist of a single point.

Daniell's idea is really rather simple: get to the main point right away, the main point being the integration of functions. Accordingly, when given a

---

[22] A linear functional is called ***positive*** if it maps positive functions to positive numbers.
[23] A fruitful year – see page 9.

measure μ on the ring $\mathcal{A}$ of sets, extend it right away to the step functions $\mathcal{E}[\mathcal{A}]$ as above. In other words, in whichever form the elementary data appear, keep them as, or turn them into, an elementary integral.

Daniell saw further that Lebesgue's expand–contract construction of the outer measure of sets has a perfectly simple analog in an up–down procedure that produces an upper integral for functions. Here is how it works. Given a *positive* elementary integral $(\mathcal{E}, \mu)$, let $\mathcal{E}^{\uparrow}$ denote the collection of functions h on F that are pointwise suprema of some sequence in $\mathcal{E}$:

$$\mathcal{E}^{\uparrow} = \big\{ h : \exists\, \phi_1, \phi_2, \ldots \ \text{ in } \ \mathcal{E} \ \text{ with } h = \sup_n \phi_n \ \big\} .$$

Since $\mathcal{E}$ is a lattice, the sequence (ϕ_n) can be chosen increasing, simply by replacing ϕ_n with $\phi_1 \vee \cdots \vee \phi_n$. $\mathcal{E}^{\uparrow}$ corresponds to Lebesgue's collection of open sets, which are countable suprema[13] of intervals. For $h \in \mathcal{E}^{\uparrow}$ set

$$\int^* h \, d\mu \stackrel{\text{def}}{=} \sup \left\{ \int \phi \, d\mu : \mathcal{E} \ni \phi \le h \right\} . \tag{A.3.1}$$

Similarly, let $\mathcal{E}_{\downarrow}$ denote the collection of functions k on the ambient set F that are pointwise infima of some sequence in $\mathcal{E}$, and set

$$\int_* k \, d\mu \stackrel{\text{def}}{=} \inf \left\{ \int \phi \, d\mu : \mathcal{E} \ni \phi \ge k \right\} .$$

Due to the σ-continuity of μ, $\int^* \, d\mu$ and $\int_* \, d\mu$ are σ-continuous on $\mathcal{E}^{\uparrow}$ and $\mathcal{E}_{\downarrow}$, respectively, in this sense: $\mathcal{E}^{\uparrow} \ni h_n \uparrow h$ implies $\int^* h_n \, d\mu \to \int^* h \, d\mu$ and $\mathcal{E}_{\downarrow} \ni k_n \downarrow k$ implies $\int_* k_n \, d\mu \to \int_* k \, d\mu$. Then set for arbitrary functions $f : F \to \overline{\mathbb{R}}$

$$\int^* f \, d\mu \stackrel{\text{def}}{=} \inf \left\{ \int^* h \, d\mu : h \in \mathcal{E}^{\uparrow}, h \ge f \right\} \qquad \text{and}$$

$$\int_* f \, d\mu \stackrel{\text{def}}{=} \sup \left\{ \int_* k \, d\mu : k \in \mathcal{E}_{\downarrow}, k \le f \right\} \quad \Big(= -\int^* -f \, d\mu \le \int^* f \, d\mu \Big).$$

$\int^* \, d\mu$ and $\int_* \, d\mu$ are called the ***upper integral*** and ***lower integral*** associated with μ, respectively. Their restrictions to sets are precisely the ***outer and inner measures*** μ^* and μ_* of Lebesgue–Caratheodory. The upper integral is ***countably subadditive***,[24] and the lower integral is countably superadditive. A function f on F is called μ-integrable if $\int^* f \, d\mu = \int_* f \, d\mu \in \mathbb{R}$, and the common value is the integral $\int f \, d\mu$. The idea is of course that on the integrable functions the integral is countably additive. The all-important Dominated Convergence Theorem follows from the countable additivity with little effort.

The procedure outlined is intuitively just as appealing as Lebesgue's, and much faster. Its real benefit lies in a slight variant, though, which is based on

[24] $\int^* \sum_{n=1}^{\infty} f_n \le \sum_{n=1}^{\infty} \int^* f_n$.

the easy observation that a function f is ***μ-integrable*** if and only if there is a sequence (ϕ_n) of elementary integrands with $\int^* |f - \phi_n| \, d\mu \to 0$, and then $\int f \, d\mu = \lim \int \phi_n \, d\mu$. So we might as well *define* integrability and the integral this way: the integrable functions are the closure of $\mathcal{E}$ under the seminorm $f \mapsto \| f \|_\mu^* \stackrel{\text{def}}{=} \int^* |f| \, d\mu$, and the integral is the extension by continuity. One now does not even have to introduce the lower integral, saving labor, and the proofs of the main results speed up some more.

Let us rewrite the definition of the ***Daniell mean*** $\| \ \|_\mu^*$:

$$\| f \|_\mu^* = \inf_{|f| \le h \in \mathcal{E}^\uparrow} \ \sup_{\phi \in \mathcal{E}, |\phi| \le h} \left| \int \phi \, d\mu \right| . \tag{D}$$

As it stands, this makes sense even if μ is not positive. It must merely have finite variation, in order that $\| \ \|_\mu^*$ be finite on $\mathcal{E}$. Again the integral can be defined simply as the extension by continuity of the elementary integral. The famous limit theorems are all consequences of two properties of the mean: it is countably subadditive on positive functions and additive on $\mathcal{E}_+$, as it agrees with the variation $\left\bracevert \mu \right\bracevert$ there.

As it stands, (D) even makes sense for measures μ that take values in some Banach space F, or even some space more general than that; one only needs to replace the absolute value in (D) by the norm or quasinorm of F. Under very mild assumptions on F, ordinary integration theory with its beautiful limit results can be established simply by repeating the classical arguments. In chapter 3 we go this route to do stochastic integration.

The main theorems of integration theory use only the properties of the mean $\| \ \|_\mu^*$ listed in definition 3.2.1. The proofs given in section 3.2 apply of course *a fortiori* in the classical case and produce the Monotone and Dominated Convergence Theorems, etc. Functions and sets that are $\| \ \|_\mu^*$-negligible or $\| \ \|_\mu^*$-measurable[25] are usually called ***μ-negligible*** or ***μ-measurable***, respectively. Their permanence properties are the ones established in sections 3.2 and 3.4. The integrability criterion 3.4.10 characterizes μ-integrable functions in terms of their local structure: μ-measurability, and their $\| \ \|_\mu^*$-size.

Let $\mathcal{E}^\sigma$ denote the sequential closure of $\mathcal{E}$. The sets in $\mathcal{E}^\sigma$ form a σ-algebra[26] $\mathcal{E}_e^\sigma$, and $\mathcal{E}^\sigma$ consists precisely of the functions measurable on $\mathcal{E}_e^\sigma$. In the case of a Radon measure, $\mathcal{E}^\sigma$ are the Baire functions. In the case that the starting point was a ring $\mathcal{A}$ of sets, $\mathcal{E}_e^\sigma$ is the σ-algebra generated by $\mathcal{A}$. The functions in $\mathcal{E}^\sigma$ are by Egoroff's theorem 3.4.4 μ-measurable for every measure μ on $\mathcal{E}$, but their collection is in general much smaller than the collection of μ-measurable functions, even in cardinality. Proposition 3.6.6, on the other hand, supplies μ-envelopes and for every μ-measurable function an equivalent one in $\mathcal{E}^\sigma$.

---

[25] See definitions 3.2.3 and 3.4.2

[26] The assumption that $\mathcal{E}$ be σ-finite is used here – see lemma A.3.3.

For the proof of the general Fubini theorem A.3.18 below it is worth stating that $\|\ \|_\mu^*$ is ***maximal***: any other mean that agrees with $\|\ \|_\mu^*$ on $\mathcal{E}_+$ is less than $\|\ \|_\mu^*$ (see 3.6.1); and for applications of capacity theory that it is ***continuous along arbitrary increasing sequences*** (see 3.6.5):

$$0 \le f_n \uparrow f \qquad \text{pointwise implies} \qquad \|f_n\|_\mu^* \uparrow \|f\|_\mu^* . \tag{A.3.2}$$

Exercise A.3.8 (Regularity) Let Ω be a set, $\mathcal{A}$ a σ-finite ring of subsets, and μ a positive σ-additive measure on the σ-algebra $\mathcal{A}^\sigma$ generated by $\mathcal{A}$. Then μ coincides with the Daniell extension of its restriction to $\mathcal{A}$.

(i) For any μ-integrable set A, $\mu(A) = \sup\{\mu(K) : K \in \mathcal{A}_\delta,\ K \subset A\}$.

(ii) Any subset Ω' of Ω has a ***measurable envelope***. This is a subset $\widetilde{\Omega'} \in \mathcal{A}^\sigma$ that contains Ω' and has the same outer measure. Any two measurable envelopes differ μ^*-negligibly.

Order-Continuous and Tight Elementary Integrals

Order-Continuity A positive Radon measure $(C_{00}(E), \mu)$ has a continuity property stronger than mere σ-continuity. Namely, if Φ is a decreasingly directed[2] subset of $C_0(E)$ with pointwise infimum zero, not necessarily countable, then $\inf\{\mu(\phi) : \phi \in \Phi\} = 0$. This is called order-continuity and is easily established using Dini's theorem A.2.1. Order-continuity occurs in the absence of local compactness as well: for instance, Dirac measure or, more generally, any measure that is carried by a countable number of points is order-continuous.

Definition A.3.9 *Let $\mathcal{E}$ be an algebra or vector lattice closed under chopping, of bounded functions on some set E. A positive linear functional $\mu : \mathcal{E} \to \mathbb{R}$ is **order-continuous** if*

$$\inf\{\mu(\phi) : \phi \in \Phi\} = 0$$

for any decreasingly directed family $\Phi \subset \mathcal{E}$ whose pointwise infimum is zero.

Sometimes it is useful to rephrase order-continuity this way: $\mu(\sup \Phi) = \sup \mu(\Phi)$ for any increasingly directed subset $\Phi \subset \mathcal{E}_+$ with pointwise supremum $\sup \Phi$ in $\mathcal{E}$.

Exercise A.3.10 If E is separable and metrizable, then any positive σ-continuous linear functional μ on $C_b(E)$ is automatically order-continuous.

In the presence of order-continuity a slightly improved integration theory is available: let $\mathcal{E}^\Uparrow$ denote the family of all functions that are pointwise suprema of arbitrary – not only the countable – subcollections of $\mathcal{E}$, and set as in (D)

$$\|f\|_\mu^\bullet = \inf_{|f| \le h \in \mathcal{E}^\Uparrow} \ \sup_{\phi \in \mathcal{E}, |\phi| \le h} \left| \int \phi \, d\mu \right| .$$

The functional $\| \ \|^{\bullet}_{\mu}$ is a mean[27] that agrees with $\| \ \|^{*}_{\mu}$ on $\mathcal{E}_+$, so thanks to the maximality of the latter it is smaller than $\| \ \|^{*}_{\mu}$ and consequently has more integrable functions. It is order-continuous[2] in the sense that $\|\sup H\|^{\bullet}_{\mu} = \sup \|H\|^{\bullet}_{\mu}$ for any increasingly directed subset $H \subset \mathcal{E}^{\Uparrow}_{+}$, and among all order-continuous means that agree with μ on $\mathcal{E}_+$ it is the maximal one.[27] The elements of $\mathcal{E}^{\Uparrow}$ are $\| \ \|^{\bullet}_{\mu}$-measurable; in fact,[27] assume that $H \subset \mathcal{E}^{\Uparrow}$ is increasingly directed with pointwise supremum h'. If $\| h' \|^{\bullet}_{\mu} < \infty$, then[27] h' is integrable and $H \to h'$ in $\| \ \|^{\bullet}_{\mu}$-mean:

$$\inf \left\{ \left\| h' - h \right\|^{\bullet}_{\mu} : h \in H \right\} = 0 \,. \tag{A.3.3}$$

For an example most pertinent in the sequel consider a completely regular space and let μ be an order-continuous positive linear functional on the lattice algebra $\mathcal{E} = C_b(E)$. Then $\mathcal{E}^{\Uparrow}$ contains all bounded lower semicontinuous functions, in particular all open sets (lemma A.2.19). The unique extension under $\| \ \|^{\bullet}_{\mu}$ integrates all bounded semicontinuous functions, and all Borel functions – not merely the Baire functions – are $\| \ \|^{\bullet}_{\mu}$-measurable. Of course, if E is separable and metrizable, then $\mathcal{E}^{\uparrow} = \mathcal{E}^{\Uparrow}$, $\| \ \|^{*}_{\mu} = \| \ \|^{\bullet}_{\mu}$ for any σ-continuous $\mu : C_b(E) \to \mathbb{R}$, and the two integral extensions coincide.

Tightness If E is locally compact, and in fact in most cases where a positive order-continuous measure μ appears naturally, μ is tight in this sense:

Definition A.3.11 *Let E be a completely regular space. A positive order-continuous functional $\mu : C_b(E) \to \mathbb{R}$ is* ***tight*** *and is called a* ***tight measure*** *on E if its integral extension with respect to $\| \ \|^{\bullet}_{\mu}$ satisfies*

$$\mu(E) = \sup\{\mu(K) : \ K \ \textit{compact} \ \} \,.$$

Tight measures are easily distinguished from each other:

Proposition A.3.12 *Let $\mathcal{M} \subset C_b(E;\mathbb{C})$ be a multiplicative class that is closed under complex conjugation, separates the points[5] of E, and has no common zeroes. Any two tight measures μ, ν that agree on $\mathcal{M}$ agree on $C_b(E)$.*

Proof. μ and ν are of course extended in the obvious complex-linear way to complex-valued bounded continuous functions. Clearly μ and ν agree on the set $\mathcal{A}^{\mathbb{C}}$ of complex-linear combinations of functions in $\mathcal{M}$ and then on the collection $\mathcal{A}^{\mathbb{R}}$ of real-valued functions in $\mathcal{A}^{\mathbb{C}}$. $\mathcal{A}^{\mathbb{R}}$ is a real algebra of real-valued functions in $C_b(E)$, and so is its uniform closure $\mathcal{A}[\mathcal{M}]$. In fact, $\mathcal{A}[\mathcal{M}]$ is also a vector lattice (theorem A.2.2), still separates the points, and $\mu = \nu$ on $\mathcal{A}[\mathcal{M}]$. There is no loss of generality in assuming that $\mu(1) = \nu(1) = 1$.

Let $f \in C_b(E)$ and $\epsilon > 0$ be given, and set $M = \| f \|_{\infty}$. The tightness of μ, ν provides a compact set K with $\mu(K) > 1 - \epsilon/M$ and $\nu(K) > 1 - \epsilon/M$.

[27] This is left as an exercise.

The restriction $f_{|K}$ of f to K can be approximated uniformly on K to within ϵ by a function $\phi \in \mathcal{A}[\mathcal{M}]$ (ibidem). Replacing ϕ by $-M \vee \phi \wedge M$ makes sure that ϕ is not too large. Now

$$|\mu(f) - \mu(\phi)| \le \int_K |f - \phi|\, d\mu + \int_{K^c} |f - \phi|\, d\mu \le \epsilon + 2M\mu(K^c) \le 3\epsilon\ .$$

The same inequality holds for ν, and as $\mu(\phi) = \nu(\phi)$, $|\mu(f) - \nu(f)| \le 6\epsilon$. This is true for all $\epsilon > 0$, and hence $\mu(f) = \nu(f)$. ∎

Exercise A.3.13 Let E be a completely regular space and $\mu : C_b(E) \to \mathbb{R}$ a positive order-continuous measure. Then $U_0 \stackrel{\text{def}}{=} \sup\{\phi \in C_b(E) : 0 \le \phi \le 1,\ \mu(\phi) = 0\}$ is integrable. It is the largest open μ-negligible set, and its complement U_0^c, the "smallest closed set of full measure," is called the ***support*** of μ.

Exercise A.3.14 An order-continuous tight measure μ on $C_b(E)$ is ***inner regular***; that is to say, its $\|\ \|_\mu^\bullet$-extension to the Borel sets satisfies

$$\mu(B) = \sup\{\mu(K) :\ K \subset B,\ K \ \text{ compact }\}$$

for any Borel set B, in fact for every $\|\ \|_\mu^\bullet$-integrable set B. Conversely, the Daniell $\|\ \|_\mu^*$-extension of a positive σ-additive inner regular set function on the Borels of a completely regular space E is order-continuous on $C_b(E)$, and the extension of the resulting linear functional on $C_b(E)$ agrees on $\mathcal{B}^\bullet(E)$ with μ. (If E is polish or Suslin, then any σ-continuous positive measure on $C_b(E)$ is inner regular – see proposition A.6.2.)

A.3.15 The Bochner Integral Suppose $(\mathcal{E}, \mu)$ is a σ-additive positive elementary integral on E and $\mathcal{V}$ is a Fréchet space equipped with a distinguished subadditive continuous gauge $\lceil\!\lceil\ \rceil\!\rceil_{\mathcal{V}}$. Denote by $\mathcal{E} \otimes \mathcal{V}$ the collection of all functions $f : E \to \mathcal{V}$ that are finite sums of the form

$$f(x) = \sum_i v_i\, \phi_i(x)\ , \qquad\qquad v_i \in \mathcal{V}\ ,\ \phi_i \in \mathcal{E}\ ,$$

and define
$$\int_E f(x)\, \mu(dx) \stackrel{\text{def}}{=} \sum_i v_i \int_E \phi_i\, d\mu \qquad\qquad \text{for such } f\ .$$

For any $f : E \to \mathcal{V}$ set
$$\lceil\!\lceil f \rceil\!\rceil_{\mathcal{V},\mu}^* \stackrel{\text{def}}{=} \lceil\!\lceil \lceil\!\lceil f \rceil\!\rceil_{\mathcal{V}} \rceil\!\rceil_\mu^* = \int^* \lceil\!\lceil f \rceil\!\rceil_{\mathcal{V}}\, d\mu$$

and let
$$\mathfrak{F}[\lceil\!\lceil\ \rceil\!\rceil_{\mathcal{V},\mu}^*] \stackrel{\text{def}}{=} \left\{f : E \to \mathcal{V} : \lceil\!\lceil \lambda f \rceil\!\rceil_{\mathcal{V},\mu}^* \xrightarrow[\lambda \to 0]{} 0\right\}\ .$$

The ***elementary $\mathcal{V}$-valued integral*** in the second line is a linear map from $\mathcal{E} \otimes \mathcal{V}$ to $\mathcal{V}$ majorized by the gauge $\lceil\!\lceil\ \rceil\!\rceil_{\mathcal{V},\mu}^*$ on $\mathfrak{F}[\lceil\!\lceil\ \rceil\!\rceil_{\mathcal{V},\mu}^*]$. Let us call a function $f : E \to \mathcal{V}$ ***Bochner μ-integrable*** if it belongs to the closure of $\mathcal{E} \otimes \mathcal{V}$ in $\mathfrak{F}[\lceil\!\lceil\ \rceil\!\rceil_{\mathcal{V},\mu}^*]$. Their collection $\mathcal{L}_{\mathcal{V}}^1(\mu)$ forms a Fréchet space with gauge $\lceil\!\lceil\ \rceil\!\rceil_{\mathcal{V},\mu}^*$, and the elementary integral has a unique continuous linear extension to this space. This extension is called the ***Bochner integral***. Neither $\mathcal{L}_{\mathcal{V}}^1(\mu)$ nor the integral extension depend on the choice of the gauge $\lceil\!\lceil\ \rceil\!\rceil_{\mathcal{V}}$. The

Dominated Convergence Theorem holds: if $\mathcal{L}^1_{\mathcal{V}}(\mu) \ni f_n \to f$ pointwise and $\llbracket f_n \rrbracket_{\mathcal{V}} \le g \in \mathfrak{F}[\llbracket\ \rrbracket^*_\mu] \quad \forall n$, then $f_n \to f$ in $\llbracket\ \rrbracket^*_{\mathcal{V},\mu}$-mean, etc.

Exercise A.3.16 Suppose that $(E, \mathcal{E}, \mu)$ is the Lebesgue integral $(\mathbb{R}_+, \mathcal{E}(\mathbb{R}_+), \lambda)$. Then the Fundamental Theorem of Calculus holds: if $f : \mathbb{R}_+ \to \mathcal{V}$ is *continuous*, then the function $F : t \mapsto \int_0^t f(s)\ \lambda(ds)$ is differentiable on $[0, \infty)$ and has the derivative $f(t)$ at t; conversely, if $F : [0, \infty) \to \mathcal{V}$ has a continuous derivative F' on $[0, \infty)$, then $F(t) = F(0) + \int_0^t F'\ d\lambda$.

Projective Systems of Measures

Let $\mathbb{T}$ be an increasingly directed index set. For every $\tau \in \mathbb{T}$ let $(E_\tau, \mathcal{E}_\tau, \mathbb{P}_\tau)$ be a triple consisting of a set E_τ, an algebra and/or vector lattice $\mathcal{E}_\tau$ of bounded elementary integrands on E_τ that contains the constants, and a σ-continuous probability $\mathbb{P}_\tau$ on $\mathcal{E}_\tau$. Suppose further that there are given surjections $\pi^\tau_\sigma : E_\tau \to E_\sigma$ such that

$$\phi \circ \pi^\tau_\sigma \in \mathcal{E}_\tau$$

and

$$\int \phi \circ \pi^\tau_\sigma \, d\mathbb{P}_\tau = \int \phi \, d\mathbb{P}_\sigma$$

for $\sigma \le \tau$ and $\phi \in \mathcal{E}_\sigma$. The data $\big((E_\tau, \mathcal{E}_\tau, \mathbb{P}_\tau, \pi^\tau_\sigma) : \sigma \le \tau \in \mathbb{T}\big)$ are called a ***consistent family*** or ***projective system of probabilities***.

Let us call a ***thread*** on a subset $S \subset \mathbb{T}$ any element $(x_\sigma)_{\sigma \in S}$ of $\prod_{\sigma \in S} E_\sigma$ with $\pi^\tau_\sigma(x_\tau) = x_\sigma$ for $\sigma < \tau$ in S and denote by $E_{\mathbb{T}} = \varprojlim E_\tau$ the set of all threads on $\mathbb{T}$.[28] For every $\tau \in \mathbb{T}$ define the map

$$\pi_\tau : E_{\mathbb{T}} \to E_\tau \text{ by } \pi_\tau\big((x_\sigma)_{\sigma \in \mathbb{T}}\big) = x_\tau \ .$$

Clearly

$$\pi^\tau_\sigma \circ \pi_\tau = \pi_\sigma \ , \qquad \sigma < \tau \in \mathbb{T} \ .$$

A function $f : E_{\mathbb{T}} \to \mathbb{R}$ of the form $f = \phi \circ \pi_\tau$, $\phi \in \mathcal{E}_\tau$, is called a ***cylinder function*** based on π_τ. We denote their collection by

$$\mathcal{E}_{\mathbb{T}} = \bigcup\nolimits_{\tau \in \mathbb{T}} \mathcal{E}_\tau \circ \pi_\tau \ .$$

Let $f = \phi \circ \pi_\sigma$ and $g = \psi \circ \pi_\tau$ be cylinder functions based on σ, τ, respectively. Then, assuming without loss of generality that $\sigma \le \tau$, $f + g = (\phi \circ \pi^\tau_\sigma + \psi) \circ \pi_\tau$ belongs to $\mathcal{E}_{\mathbb{T}}$. Similarly one sees that $\mathcal{E}_{\mathbb{T}}$ is closed under multiplication, finite infima, etc: $\mathcal{E}_{\mathbb{T}}$ is an algebra and/or vector lattice of bounded functions on $E_{\mathbb{T}}$. If the function $f \in \mathcal{E}_{\mathbb{T}}$ is written as $f = \phi \circ \pi_\sigma = \psi \circ \pi_\tau$ with $\phi \in C_b(E_\sigma), \psi \in C_b(E_\tau)$, then there is an $\upsilon > \sigma, \tau$ in $\mathbb{T}$, and with $\rho \stackrel{\text{def}}{=} \phi \circ \pi^\upsilon_\sigma = \psi \circ \pi^\upsilon_\tau$, $\mathbb{P}_\sigma(\phi) = \mathbb{P}_\upsilon(\rho) = \mathbb{P}_\tau(\psi)$ due to the consistency. We may thus define unequivocally for $f \in \mathcal{E}_{\mathbb{T}}$, say $f = \phi \circ \pi_\sigma$,

$$\mathbb{P}(f) \stackrel{\text{def}}{=} \mathbb{P}_\sigma(\phi) \ .$$

[28] It may well be empty or at least rather small.

Clearly $\mathbb{P}:\mathcal{E}_{\mathbb{T}}\to\mathbb{R}$ is a positive linear map with $\sup\{\mathbb{P}(f):|f|\le 1\}=1$. It is denoted by $\varprojlim\mathbb{P}_\tau$ and is called the ***projective limit*** of the $\mathbb{P}_\tau$. We also call $(E_T,\mathcal{E}_T,\mathbb{P})$ the ***projective limit*** of the elementary integrals $(E_\tau,\mathcal{E}_\tau,\mathbb{P}_\tau,\pi^\tau_\sigma)$ and denote it by $=\varprojlim(E_\tau,\mathcal{E}_\tau,\mathbb{P}_\tau,\pi^\tau_\sigma)$. $\mathbb{P}$ will not in general be σ-additive. The following theorem identifies sufficient conditions under which it is. To facilitate its statement let us call the projective system ***full*** if every thread on any subset of indices can be extended to a thread on all of $\mathbb{T}$. For instance, when $\mathbb{T}$ has a countable cofinal subset then the system is full.

Theorem A.3.17 (Kolmogorov) *Assume that*

(i) the projective system $\big((E_\tau,\mathcal{E}_\tau,\mathbb{P}_\tau,\pi^\tau_\sigma):\sigma\le\tau\in\mathbb{T}\big)$ *is full;*
(ii) every $\mathbb{P}_\tau$ *is tight under the topology generated by* $\mathcal{E}_\tau$.

Then the projective limit $\mathbb{P}=\varprojlim\mathbb{P}_\tau$ *is σ-additive.*

Proof. Suppose the sequence of functions $f_n=\phi_n\circ\pi_{\tau_n}\in\mathcal{E}_{\mathbb{T}}$ decreases pointwise to zero. We have to show that $\mathbb{P}(f_n)\to 0$. By way of contradiction assume there is an $\epsilon>0$ with $\mathbb{P}(f_n)>2\epsilon\ \forall n$. There is no loss of generality in assuming that the τ_n increase with n and that $f_1\le 1$. Let K_n be a compact subset of E_{τ_n} with $\mathbb{P}_{\tau_n}(K_n)>1-\epsilon 2^{-n}$, and set $K^n=\bigcap_{N\ge n}\pi^{\tau_N}_{\tau_n}(K_N)$. Then $\mathbb{P}_{\tau_n}(K^n)\ge 1-\epsilon$, and thus $\int_{K^n}\phi_n\,d\mathbb{P}_{\tau_n}>\epsilon$ for all n. The compact sets $\overline{K}^n\stackrel{\text{def}}{=}K^n\cap[\phi_n\ge\epsilon]$ are non-void and have $\pi^{\tau_n}_{\tau_m}(\overline{K}^n)\supset\overline{K}^m$ for $m\le n$, so there is a thread $(x_{\tau_1},x_{\tau_2},\ldots)$ with $\phi_n(x_{\tau_n})\ge\epsilon$. This thread can be extended to a thread θ on all of $\mathbb{T}$, and clearly $f_n(\theta)\ge\epsilon\ \forall n$. This contradiction establishes the claim. ∎

Products of Elementary Integrals

Let $(E,\mathcal{E}_E,\mu)$ and $(F,\mathcal{E}_F,\nu)$ be positive elementary integrals. Extending μ and ν as usual, we may assume that $\mathcal{E}_E$ and $\mathcal{E}_F$ are the step functions over the σ-algebras $\mathcal{A}_E,\mathcal{A}_F$, respectively. The ***product σ-algebra*** $\mathcal{A}_E\otimes\mathcal{A}_F$ is the σ-algebra on the cartesian product $G\stackrel{\text{def}}{=}E\times F$ generated by the product paving of rectangles

$$\mathcal{A}_E\times\mathcal{A}_F\stackrel{\text{def}}{=}\big\{A\times B:\ A\in\mathcal{A}_E,B\in\mathcal{A}_F\big\}\ .$$

Let $\mathcal{E}_G$ be the collection of functions on G of the form

$$\phi(x,y)=\sum_{k=1}^{K}\phi_k(x)\psi_k(y)\ ,\quad K\in\mathbb{N},\phi_k\in\mathcal{E}_E,\psi_k\in\mathcal{E}_F\ . \tag{A.3.4}$$

Clearly[13] $\mathcal{A}_E\otimes\mathcal{A}_F$ is the σ-algebra generated by $\mathcal{E}_G$. Define the product measure $\gamma=\mu\times\nu$ on a function as in equation (A.3.4) by

$$\begin{aligned}\int_G\phi(x,y)\,\gamma(dx,dy)&\stackrel{\text{def}}{=}\sum_k\int_E\phi_k(x)\,\mu(dx)\cdot\int_F\psi_k(y)\,\nu(dy)\\&=\int_F\Big(\int_E\phi(x,y)\,\mu(dx)\Big)\,\nu(dy)\ .\end{aligned}\tag{A.3.5}$$

The first line shows that this definition is symmetric in x, y, the second that it is independent of the particular representation (A.3.4) and that γ is σ-continuous, that is to say, $\phi_n(x,y) \downarrow 0$ implies $\int \phi_n \, d\gamma \to 0$. This is evident since the inner integral in equation (A.3.5) belongs to $\mathcal{E}_F$ and decreases pointwise to zero. We can now extend the integral to all $\mathcal{A}_E \otimes \mathcal{A}_F$-measurable functions with finite upper γ-integral, etc.

Fubini's Theorem says that the integral $\int f \, d\gamma$ can be evaluated iteratively as $\int \big(\int f(x,y) \, \mu(dx) \big) \nu(dy)$ for γ-integrable f. In several instances we need a generalization, one that refers to a slightly more general setup.

Suppose that we are given for every $y \in F$ not the fixed measure $\phi(x,y) \mapsto \int_E f(x,y) \, d\mu(x)$ on $\mathcal{E}_G$ but a measure μ_y that varies with $y \in F$, but so that $y \mapsto \int \phi(x,y) \, \mu_y(dx)$ is ν-integrable for all $\phi \in \mathcal{E}_G$. We can then define a measure $\gamma = \int \mu_y \, \nu(dy)$ on $\mathcal{E}_G$ via iterated integration:

$$\int \phi \, d\gamma \stackrel{\text{def}}{=} \int \Big(\int \phi(x,y) \, \mu_y(dx) \Big) \nu(dy) \,, \qquad \phi \in \mathcal{E}_G \,.$$

If $\mathcal{E}_G \ni \phi_n \downarrow 0$, then $\mathcal{E}_F \ni \int \phi_n(x,y) \, \mu_y(dx) \downarrow 0$ and consequently $\int \phi_n \, d\gamma \to 0$: γ is σ-continuous. Fubini's theorem can be generalized to say that the γ-integral can be evaluated as an iterated integral:

Theorem A.3.18 (Fubini) *If f is γ-integrable, then $\int f(x,y) \, \mu_y(dx)$ exists for ν-almost all $y \in Y$ and is a ν-integrable function of y, and*

$$\int f \, d\gamma = \int \Big(\int f(x,y) \, \mu_y(dx) \Big) \, \nu(dy) \,.$$

Proof. The assignment $f \mapsto \int^* \left(\int^* |f(x,y)| \, \mu_y(dx) \right) \, \nu(dy)$ is a mean that coincides with the usual Daniell mean $\| \ \|_\gamma^*$ on $\mathcal{E}$, and the maximality of Daniell's mean gives

$$\int^* \Big(\int^* |f(x,y)| \, \mu_y(dx) \Big) \, \nu(dy) \le \| f \|_\gamma^* \tag{$*$}$$

for all $f : G \to \overline{\mathbb{R}}$. Given the γ-integrable function f, find a sequence (ϕ_n) of functions in $\mathcal{E}_G$ with $\sum \| \phi_n \|_\gamma^* < \infty$ and such that $f = \sum \phi_n$ both in $\| \ \|_\gamma^*$-mean and γ-almost surely. Applying $(*)$ to the set of points $(x,y) \in G$ where $\sum \phi_n(x,y) \ne f(x,y)$, we see that the set N_1 of points $y \in F$ where not $\sum \phi_n(\cdot,y) = f(\cdot,y)$ μ_y-almost surely is ν-negligible. Since

$$\Big\| \Big\| \sum |\phi_n(\cdot,y)| \Big\|_{\mu_y}^* \Big\|_\nu^* \le \Big\| \sum \| \phi_n(\cdot,y) \|_{\mu_y}^* \Big\|_\nu^* \le \sum \| \phi_n \|_\gamma^* < \infty \,,$$

the sum $g(y) \stackrel{\text{def}}{=} \| \sum |\phi_n(\cdot,y)| \|_{\mu_y}^* = \sum \int |\phi_n(x,y)| \, \mu_y(dx)$ is ν-measurable and finite in ν-mean, so it is ν-integrable. It is, in particular, finite ν-almost surely (proposition 3.2.7). Set $N_2 = [g = \infty]$ and fix a $y \notin N_1 \cup N_2$. Then $\overline{f}(\cdot,y) \stackrel{\text{def}}{=} \sum |\phi_n(\cdot,y)|$ is μ_y-almost surely finite (ibidem). Hence $\sum \phi_n(\cdot,y)$

converges μ_y-almost surely absolutely. In fact, since $y \notin N_1$, the sum is $f(\cdot, y)$. The partial sums are dominated by $\overline{f}(\cdot, y) \in \mathfrak{L}^1(\mu_y)$. Thus $f(\cdot, y)$ is μ_y-integrable with integral

$$I(y) \stackrel{\text{def}}{=} \int f(x, y)\, \mu_y(dx) = \lim_n \int \sum_{\nu \le n} \phi_\nu(x, y) \mu_y(dx) .$$

I is ν-almost surely defined and ν-measurable, with $|I| \le g$ having $\|I\|_\nu^* < \infty$; it is thus ν-integrable with integral

$$\int I(y)\, \nu(dy) = \lim \int \int \sum_{\nu \le n} \phi_\nu(x, y) \mu_y(dx)\, \nu(dy) = \int f\, d\gamma .$$

The interchange of limit and integral here is justified by the observation that $|\int \sum_{\nu \le n} \phi_\nu(x, y) \mu_y(dx)| \le \sum_{\nu \le n} \int |\phi_\nu(x, y)| \mu_y(dx) \le g(y)$ for all n. ∎

Infinite Products of Elementary Integrals

Suppose for every t in an index set $\mathcal{T}$ the triple $(E_t, \mathcal{E}_t, \mathbb{P}_t)$ is a positive σ-additive elementary integral of total mass 1. For any finite subset $\tau \subset \mathcal{T}$ let $(E_\tau, \mathcal{E}_\tau, \mathbb{P}_\tau)$ be the product of the elementary integrals $(E_t, \mathcal{E}_t, \mathbb{P}_t)$, $t \in \tau$. For $\sigma \subset \tau$ there is the obvious projection $\pi_\sigma^\tau : E_\tau \to E_\sigma$ that "forgets the components not in σ," and the projective limit (see page 401)

$$(E, \mathcal{E}, \mathbb{P}) \stackrel{\text{def}}{=} \prod_{t \in \mathcal{T}} (E_t, \mathcal{E}_t, \mathbb{P}_t) \stackrel{\text{def}}{=} \varprojlim_{\tau \subset \mathcal{T}} (E_\tau, \mathcal{E}_\tau, \mathbb{P}_\tau, \pi_\sigma^\tau)$$

of this system is the ***product of the elementary integrals*** $(E_t, \mathcal{E}_t, \mathbb{P}_t)$, $t \in \mathcal{T}$. It has for its underlying set the cartesian product $E = \prod_{t \in \mathcal{T}} E_t$. The cylinder functions of $\mathcal{E}$ are finite sums of functions of the form

$$(e_t)_{t \in \mathcal{T}} \mapsto \phi_1(e_{t_1}) \cdot \phi_2(e_{t_2}) \cdots \phi_j(e_{t_j}) , \qquad \phi_i \in \mathcal{E}_{t_i} ,$$

"that depend only on finitely many components." The projective limit $\mathbb{P} = \varprojlim \mathbb{P}_\tau$ clearly has mass 1.

Exercise A.3.19 Suppose $\mathcal{T}$ is the disjoint union of two non-void subsets $\mathcal{T}_1, \mathcal{T}_2$. Set $(E_i, \mathcal{E}_i, \mathbb{P}_i) \stackrel{\text{def}}{=} \prod_{t \in \mathcal{T}_i} (E_t, \mathcal{E}_t, \mathbb{P}_t)$, $i = 1, 2$. Then in a canonical way

$$\prod_{t \in \mathcal{T}} (E_t, \mathcal{E}_t, \mathbb{P}_t) = (E_1, \mathcal{E}_1, \mathbb{P}_1) \times (E_2, \mathcal{E}_2, \mathbb{P}_2) , \qquad \text{so that,}$$

$$\text{for } \phi \in \mathcal{E}, \quad \int \phi(e_1, e_2) \mathbb{P}(de_1, de_2) = \int \Big(\int \phi(e_1, e_2) \mathbb{P}_2(de_2) \Big) \mathbb{P}_1(de_1) . \tag{A.3.6}$$

The present projective system is clearly full, in fact so much so that no tightness is needed to deduce the σ-additivity of $\mathbb{P} = \varprojlim \mathbb{P}_\tau$ from that of the factors $\mathbb{P}_t$:

Lemma A.3.20 *If the $\mathbb{P}_t$ are σ-additive, then so is $\mathbb{P}$.*

Proof. Let (ϕ_n) be a pointwise decreasing sequence in $\mathcal{E}_+$ and assume that for all n

$$\int \phi_n(e)\,\mathbb{P}(de) \geq a > 0\ .$$

There is a countable collection $\mathcal{T}_0 = \{t_1, t_2, \ldots\} \subset \mathcal{T}$ so that every ϕ_n depends only on coordinates in $\mathcal{T}_0$. Set $\mathcal{T}_1 \stackrel{\text{def}}{=} \{t_1\}$, $\mathcal{T}_2 \stackrel{\text{def}}{=} \{t_2, t_3, \ldots\}$. By (A.3.6),

$$\int \Big(\int \phi_n(e_1, e_2)\mathbb{P}_2(de_2)\Big)\mathbb{P}_{t_1}(de_1) > a$$

for all n. This can only be if the integrands of $\mathbb{P}_{t_1}$, which form a pointwise decreasing sequence of functions in $\mathcal{E}_{t_1}$, exceed a at some common point $e'_1 \in E_{t_1}$: for all n

$$\int \phi_n(e'_1, e_2)\,\mathbb{P}_2(de_2) \geq a\ .$$

Similarly we deduce that there is a point $e'_2 \in E_{t_2}$ so that for all n

$$\int \phi_n(e'_1, e'_2, e_3)\,\mathbb{P}_3(de_3) \geq a\ ,$$

where $e_3 \in E_3 \stackrel{\text{def}}{=} E_{t_3} \times E_{t_4} \times \cdots$. There is a point $e' = (e_t) \in E$ with $e'_{t_i} = e'_i$ for $i = 1, 2, \ldots$, and clearly $\phi_n(e') \geq a$ for all n. ∎

So our product measure is σ-additive, and we can effect the usual extension upon it (see page 395 ff.).

Exercise A.3.21 State and prove Fubini's theorem for a $\mathbb{P}$-integrable function $f : E \to \mathbb{R}$.

Images, Law, and Distribution

Let $(X, \mathcal{E}_X)$ and $(Y, \mathcal{E}_Y)$ be two spaces, each equipped with an algebra of bounded elementary integrands, and let $\mu : \mathcal{E}_X \to \mathbb{R}$ be a positive σ-continuous measure on $(X, \mathcal{E}_X)$. A map $\Phi : X \to Y$ is called ***μ-measurable***[29] if $\psi \circ \Phi$ is μ-integrable for every $\psi \in \mathcal{E}_Y$. In this case the ***image of μ under Φ*** is the measure $\nu = \Phi[\mu]$ on $\mathcal{E}_Y$ defined by

$$\nu(\psi) = \int \psi \circ \Phi\, d\mu\ , \qquad \psi \in \mathcal{E}_Y\ .$$

Some authors write $\mu \circ \Phi^{-1}$ for $\Phi[\mu]$. ν is also called the ***distribution*** or ***law*** of Φ under μ. For every $x \in X$ let λ_x be the Dirac measure at $\Phi(x)$. Then clearly

$$\int_Y \psi(y)\,\nu(dy) = \int_X \int_Y \psi(y)\lambda_x(dy)\,\mu(dx)$$

[29] The "right" definition is actually this: Φ is μ-measurable if it is ***largely uniformly continuous*** in the sense of definition 3.4.2 on page 110, where of course X, Y are given the uniformities generated by $\mathcal{E}_X, \mathcal{E}_Y$, respectively.

for $\psi \in \mathcal{E}_Y$, and Fubini's theorem A.3.18 says that this equality extends to all ν-integrable functions. This fact can be read to say: if h is ν-integrable, then $h \circ \Phi$ is μ-integrable and

$$\int_Y h \, d\nu = \int_X h \circ \Phi \, d\mu \, . \tag{A.3.7}$$

We leave it to the reader to convince herself that this definition and conclusion stay *mutatis mutandis* when both μ and ν are σ-finite.

Suppose $\mathcal{X}$ and $\mathcal{Y}$ are (the step functions over) σ-algebras. If μ is a probability $\mathbb{P}$, then the law of Φ is evidently a probability as well and is given by

$$\Phi[\mathbb{P}](B) \overset{\text{def}}{=} \mathbb{P}([\Phi \in B]) \, , \quad \forall\, B \in \mathcal{Y} \, . \tag{A.3.8}$$

Suppose Φ is real-valued. Then the ***cumulative distribution function***[30] of (the law of) Φ is the function $t \mapsto F_\Phi(t) = \mathbb{P}[\Phi \le t] = \Phi[\mathbb{P}]\big((-\infty, t]\big)$. Theorem A.3.18 applied to $(y, \lambda) \mapsto \phi'(\lambda)[\Phi(y) > \lambda]$ yields

$$\begin{aligned} \int \phi \, d\Phi[\mathbb{P}] &= \int \phi \circ \Phi \, d\mathbb{P} \\ &= \int_{-\infty}^{+\infty} \phi'(\lambda)\mathbb{P}[\Phi > \lambda] \, d\lambda = \int_{-\infty}^{+\infty} \phi'(\lambda)\big(1 - F_\Phi(\lambda)\big) \, d\lambda \end{aligned} \tag{A.3.9}$$

for any differentiable function ϕ that vanishes at $-\infty$. One defines the ***cumulative distribution function*** $F = F_\mu$ for any measure μ on the line or half-line by $F(t) = \mu((-\infty, t])$, and then denotes μ by dF and the variation $¦\mu¦$ variously by $|dF|$ or by $d¦F¦$.

The Vector Lattice of All Measures

Let $\mathcal{E}$ be a σ-finite algebra and vector lattice closed under chopping, of bounded functions on some set F. We denote by $\mathfrak{M}^*[\mathcal{E}]$ the set of all measures – i.e., σ-continuous elementary integrals of finite variation – on $\mathcal{E}$. This is a vector space under the usual addition and scalar multiplication of functions. Defining an order by saying that $\mu \le \nu$ is to mean that $\nu - \mu$ is a positive measure[22] makes $\mathfrak{M}^*[\mathcal{E}]$ into a vector lattice. That is to say, for every two measures $\mu, \nu \in \mathfrak{M}^*[\mathcal{E}]$ there is a least measure $\mu \vee \nu$ greater than both μ and ν and a greatest measure $\mu \wedge \nu$ less than both. In these terms the variation $¦\mu¦$ is nothing but $\mu \vee (-\mu)$. In fact, $\mathfrak{M}^*[\mathcal{E}]$ is order-complete: suppose $\mathcal{M} \subset \mathfrak{M}^*[\mathcal{E}]$ is ***order-bounded*** from above, i.e., there is a $\nu \in \mathfrak{M}^*[\mathcal{E}]$ greater than every element of $\mathcal{M}$; then there is a least upper order bound $\bigvee \mathcal{M}$ [5].

Let $\mathcal{E}_0^\sigma = \{\phi \in \mathcal{E}^\sigma : \ |\phi| \le \psi \ \text{ for some } \ \psi \in \mathcal{E}\}$, and for every $\mu \in \mathfrak{M}^*[\mathcal{E}]$ let μ^σ denote the restriction of the extension $\int d\mu$ to $\mathcal{E}_0^\sigma$. The map $\mu \mapsto \mu^\sigma$ is an order-preserving linear isomorphism of $\mathfrak{M}^*[\mathcal{E}]$ onto $\mathfrak{M}^*\big(\mathcal{E}_0^\sigma\big)$.

---

[30] A ***distribution function*** of a measure μ on the line is any function $F : (-\infty, \infty) \to \mathbb{R}$ that has $\mu\big((a, b]\big) = F(b) - F(a)$ for $a < b$ in $(-\infty, \infty)$. Any two differ by a constant. The cumulative distribution function is thus that distribution function which has $F(-\infty) = 0$.

Every $\mu \in \mathfrak{M}^*[\mathcal{E}]$ has an extension whose σ-algebra of μ-measurable sets includes $\mathcal{E}_e^\sigma$ but is generally cardinalities bigger. The ***universal completion*** of $\mathcal{E}$ is the collection of all sets that are μ-measurable for every single $\mu \in \mathfrak{M}^*[\mathcal{E}]$. It is denoted by $\mathcal{E}_e^*$. It is clearly a σ-algebra containing $\mathcal{E}_e^\sigma$. A function f measurable on $\mathcal{E}_e^*$ is called ***universally measurable***. This is of course the same as saying that f is μ-measurable for every $\mu \in \mathfrak{M}^*[\mathcal{E}]$.

Theorem A.3.22 (Radon–Nikodym) *Let $\mu, \nu \in \mathfrak{M}^*[\mathcal{E}]$, with $\mathcal{E}$ σ-finite. The following are equivalent:*[26]

(i) $\rvert\mu\lvert = \bigvee_{k\in\mathbb{N}} \rvert\mu\lvert \wedge (k\rvert\nu\lvert)$.

(ii) For every decreasing sequence $\phi_n \in \mathcal{E}_+$, $\nu(\phi_n) \to 0 \implies \mu(\phi_n) \to 0$.

(iii) For every decreasing sequence $\phi_n \in \mathcal{E}_{0+}^\sigma$, $\nu^\sigma(\phi_n) \to 0 \implies \mu^\sigma(\phi_n) \to 0$.

(iv) For $\phi \in \mathcal{E}_0^\sigma$, $\nu^\sigma(\phi) = 0$ *implies* $\mu^\sigma(\phi) = 0$.

(v) A ν-negligible set is μ-negligible.

(vi) There exists a function $g \in \mathcal{E}^\sigma$ *such that* $\mu(\phi) = \nu^\sigma(g\phi)$ *for all* $\phi \in \mathcal{E}$.

*In this case μ is called **absolutely continuous** with respect to ν and we write $\mu \ll \nu$; furthermore, then $\int f\, d\mu = \int fg\, d\nu$ whenever either side makes sense. The function g is the **Radon–Nikodym derivative** or **density** of μ with respect to ν, and it is customary to write $\mu = g\nu$, that is to say, for $\phi \in \mathcal{E}$ we have $(g\nu)(\phi) = \nu^\sigma(g\phi)$.*

If $\mu \ll \rho$ for all $\mu \in \mathcal{M} \subset \mathfrak{M}^*[\mathcal{E}]$, then $\bigvee \mathcal{M} \ll \rho$.

Exercise A.3.23 Let $\mu, \nu : C_b(E) \to \mathbb{R}$ be σ-additive with $\mu \ll \nu$. If ν is order-continuous and tight, then so is μ.

Conditional Expectation

Let $\Phi : (\Omega, \mathcal{F}) \to (Y, \mathcal{Y})$ be a measurable map of measurable spaces and μ a positive finite measure on $\mathcal{F}$ with image $\nu \overset{\text{def}}{=} \Phi[\mu]$ on $\mathcal{Y}$.

Theorem A.3.24 *(i) For every μ-integrable function $f : \Omega \to \mathbb{R}$ there exists a ν-integrable $\mathcal{Y}$-measurable function $\mathbb{E}[f|\Phi] = \mathbb{E}^\mu[f|\Phi] : Y \to \mathbb{R}$, called the **conditional expectation of f given Φ**, such that*

$$\int_\Omega f \cdot h \circ \Phi \, d\mu = \int_Y \mathbb{E}[f|\Phi] \cdot h \, d\nu$$

for all bounded $\mathcal{Y}$-measurable $h : Y \to \mathbb{R}$. Any two conditional expectations differ at most in a ν-negligible set of $\mathcal{Y}$ and depend only on the class of f.

(ii) The map $f \mapsto \mathbb{E}^\mu[f|\Phi]$ is linear and positive, maps 1 to 1, and is contractive[31] *from $L^p(\mu)$ to $L^p(\nu)$ when $1 \le p \le \infty$.*

---

[31] A linear map $\Phi : E \to S$ between seminormed spaces is ***contractive*** if there exists a $\gamma \le 1$ such that $\|\Phi(x)\|_S \le \gamma \cdot \|x\|_E$ for all $x \in E$; the least γ satisfying this inequality is the ***modulus of contractivity*** of Φ. If the contractivity modulus is strictly less than 1, then Φ is ***strictly contractive***.

(iii) Assume $\Gamma : \mathbb{R} \to \mathbb{R}$ *is convex*[32] *and* $f : \Omega \to \mathbb{R}$ *is* $\mathcal{F}$*-measurable and such that* $\Gamma \circ f$ *is* ν*-integrable. Then if* $\mu(1) = 1$*, we have* ν*-almost surely*

$$\Gamma\big(\mathbb{E}^\mu[f|\Phi]\big) \le \mathbb{E}^\mu[\Gamma(f)|\Phi] . \tag{A.3.10}$$

Proof. (i) Consider the measure $f\mu : B \mapsto \int_B f\, d\mu$, $B \in \mathcal{F}$, and its image $\nu' = \Phi[f\mu]$. This is a measure on the σ-algebra $\mathcal{Y}$, absolutely continuous with respect to ν. The Radon–Nikodym theorem provides a derivative $d\nu'/d\nu$, which we may call $\mathbb{E}^\mu[f|\Phi]$. If f is changed μ-negligibly, then the measure ν' and thus the (class of the) derivative do not change.

(ii) The linearity and positivity are evident. The contractivity follows from (iii) and the observation that $x \mapsto |x|^p$ is convex when $1 \le p < \infty$.

(iii) There is a countable collection of linear functions $\ell_n(x) = \alpha_n + \beta_n x$ such that $\Gamma(x) = \sup_n \ell_n(x)$ at every point $x \in \mathbb{R}$. Linearity and positivity give

$$\ell_n\big(\mathbb{E}^\mu[f|\Phi]\big) = \mathbb{E}^\mu\big[\ell_n(f)|\Phi\big] \le \mathbb{E}^\mu[\Gamma(f)|\Phi] \qquad \text{a.s.} \quad \forall\, n \in \mathbb{N} .$$

Upon taking the supremum over n, ***Jensen's inequality*** (A.3.10) follows. ∎

Frequently the situation is this: Given is not a map Φ but a sub-σ-algebra $\mathcal{Y}$ of $\mathcal{F}$. In that case we understand Φ to be the identity $(\Omega, \mathcal{F}) \to (\Omega, \mathcal{Y})$. Then $\mathbb{E}^\mu[f|\Phi]$ is usually denoted by $\mathbb{E}[f|\mathcal{Y}]$ or $\mathbb{E}^\mu[f|\mathcal{Y}]$ and is called the ***conditional expectation of f given $\mathcal{Y}$***. It is thus defined by the identity

$$\int f \cdot H\, d\mu = \int \mathbb{E}^\mu[f|\mathcal{Y}] \cdot H\, d\mu , \qquad H \in \mathcal{Y}_b ,$$

and (i)–(iii) continue to hold, *mutatis mutandis*.

Exercise A.3.25 Let μ be a subprobability ($\mu(1) \le 1$) and $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ a concave function. Then for all μ-integrable functions z

$$\int \phi(|z|)\, d\mu \le \phi\Big(\int |z|\, d\mu\Big) .$$

Exercise A.3.26 On the probability triple $(\Omega, \mathcal{G}, \mathbb{P})$ let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{G}$, X an $\mathcal{F}/\mathcal{X}$-measurable map from Ω to some measurable space $(\Xi, \mathcal{X})$, and Φ a bounded $\mathcal{X} \otimes \mathcal{G}$-measurable function. For every $x \in \Xi$ set $\overline{\Phi}(x, \omega) \stackrel{\text{def}}{=} \mathbb{E}[\Phi(x, \boldsymbol{\cdot})|\mathcal{F}](\omega)$. Then $\mathbb{E}[\Phi(X(\boldsymbol{\cdot}), \boldsymbol{\cdot})|\mathcal{F}](\omega) = \overline{\Phi}(X(\omega), \omega)$ $\mathbb{P}$-almost surely.

Numerical and σ-Finite Measures

Many authors define a measure to be a triple $(\Omega, \mathcal{F}, \mu)$, where $\mathcal{F}$ is a σ-algebra on Ω and $\mu : \mathcal{F} \to \overline{\mathbb{R}}_+$ is ***numerical***, i.e., is allowed to take values in the extended reals $\overline{\mathbb{R}}$, with suitable conventions about the meaning of $r + \infty$, etc.

[32] Γ is ***convex*** if $\Gamma\big(\lambda x + (1-\lambda)x'\big) \le \lambda\Gamma(x) + (1-\lambda)\Gamma(x')$ for $x, x' \in \operatorname{dom}\Gamma$ and $0 \le \lambda \le 1$; it is ***strictly convex*** if $\Gamma\big(\lambda x + (1-\lambda)x'\big) < \lambda\Gamma(x) + (1-\lambda)\Gamma(x')$ for $x \ne x' \in \operatorname{dom}\Gamma$ and $0 < \lambda < 1$.

(see A.1.2). μ is σ-additive if it satisfies $\mu\big(\bigcup F_n\big) = \sum \mu(F_n)$ for mutually disjoint sets $F_n \in \mathcal{F}$. Unless the δ-ring $\mathcal{D}_\mu \stackrel{\text{def}}{=} \{F \in \mathcal{F} : \mu(F) < \infty\}$ generates the σ-algebra $\mathcal{F}$, examples of quite unnatural behavior can be manufactured [111]. If this requirement is made, however, then any reasonable integration theory of the measure space $(\Omega, \mathcal{F}, \mu)$ is essentially the same as the integration theory of $(\Omega, \mathcal{D}_\mu, \mu)$ explained above.

μ is called ***σ-finite*** if $\mathcal{D}_\mu$ is a σ-finite class of sets (exercise A.3.2), i.e., if there is a countable family of sets $F_n \in \mathcal{F}$ with $\mu(F_n) < \infty$ and $\bigcup_n F_n = \Omega$; in that case the requirement is met and $(\Omega, \mathcal{F}, \mu)$ is also called a ***σ-finite measure space***.

Exercise A.3.27 Consider again a measurable map $\Phi : (\Omega, \mathcal{F}) \to (Y, \mathcal{Y})$ of measurable spaces and a measure μ on $\mathcal{F}$ with image ν on $\mathcal{Y}$, and assume that both μ and ν are σ-finite on their domains.
(i) With μ_0 denoting the restriction of μ to $\mathcal{D}_\mu$, $\mu = \mu_0^*$ on $\mathcal{F}$.
(ii) Theorem A.3.24 stays, including Jensen's inequality (A.3.10).
(iii) If Γ is strictly convex, then equality holds in inequality (A.3.10) if and only if f is almost surely equal to a function of the form $f' \circ \Phi$, f' $\mathcal{Y}$-measurable.
(iv) For $h \in \mathcal{Y}_b$, $\mathbb{E}[fh \circ \Phi|\Phi] = h \cdot \mathbb{E}[f|\Phi]$ provided both sides make sense.
(v) Let $\Psi : (Y, \mathcal{Y}) \to (Z, \mathcal{Z})$ be measurable, and assume $\Psi[\nu]$ is σ-finite. Then

$$\mathbb{E}^\mu[f|\Psi \circ \Phi] = \mathbb{E}^\nu[\mathbb{E}^\mu[f|\Phi]|\Psi] \ , \text{ and } \mathbb{E}[f|\mathcal{Z}] = \mathbb{E}\big[\mathbb{E}[f|\mathcal{Y}]|\mathcal{Z}\big]$$

when $\Omega = Y = Z$ and $\mathcal{Z} \subset \mathcal{Y} \subset \mathcal{F}$.
(vi) If $\mathbb{E}[f \cdot b] = \mathbb{E}\big[f \cdot \mathbb{E}[b|\mathcal{Y}]\big]$ for all $b \in L^\infty(\mathcal{Y})$, then f is measurable on $\mathcal{Y}$.

Exercise A.3.28 The argument in the proof of Jensen's inequality theorem A.3.24 can be used in a slightly different context. Let E be a Banach space, ν a signed measure with σ-finite variation $\lvert\nu\rvert$, and $f \in L^1_E(\nu)$ (see item A.3.15). Then

$$\Big\| \int f \, d\nu \Big\|_E \le \int \| f \|_E \, d\lvert\nu\rvert \ .$$

Exercise A.3.29 Yet another variant of the same argument can be used to establish the following inequality, which is used repeatedly in chapter 5. Let $(F, \mathcal{F}, \mu)$ and $(G, \mathcal{G}, \nu)$ be σ-finite measure spaces. Let f be a function measurable on the product σ-algebra $\mathcal{F} \otimes \mathcal{G}$ on $F \times G$. Then

$$\Big\| \| f \|_{L^p(\mu)} \Big\|_{L^q(\nu)} \le \Big\| \| f \|_{L^q(\nu)} \Big\|_{L^p(\mu)}$$

for $0 < p \le q \le \infty$.

Characteristic Functions

It is often difficult to get at the law of a random variable $\Phi : (F, \mathcal{F}) \to (G, \mathcal{G})$ through its definition (A.3.8). There is a recurring situation when the powerful tool of characteristic functions can be applied. Namely, let us suppose that $\mathcal{G}$ is generated by a vector space Γ of real-valued functions. Now, inasmuch as $\gamma = -i \lim_{n\to\infty} n\big(e^{i\gamma/n} - e^{i0}\big)$, $\mathcal{G}$ is also generated by the functions

$$y \mapsto e^{i\gamma(y)} \ , \qquad \gamma \in \Gamma \ .$$

These functions evidently form a complex multiplicative class $e^{i\Gamma}$, and in view of exercise A.3.5 any σ-additive measure μ of totally finite variation on $\mathcal{G}$ is determined by its values

$$\widehat{\mu}(\gamma) = \int_G e^{i\gamma(y)}\ \mu(dy) \tag{A.3.11}$$

on them. $\widehat{\mu}$ is called the ***characteristic function*** of μ. We also write $\widehat{\mu}^\Gamma$ when it is necessary to indicate that this notion depends on the generating vector space Γ, and then talk about the characteristic function of μ for Γ. If μ is the law of $\Phi : (F, \mathcal{F}) \to (G, \mathcal{G})$ under $\mathbb{P}$, then (A.3.7) allows us to rewrite equation (A.3.11) as

$$\widehat{\Phi[\mathbb{P}]}(\gamma) = \int_G e^{i\gamma(y)}\ \Phi[\mathbb{P}](dy) = \int_F e^{i\gamma\circ\Phi}\ d\mathbb{P}\ , \qquad \gamma \in \Gamma\ .$$

$\widehat{\Phi[\mathbb{P}]} = \widehat{\Phi[\mathbb{P}]}^\Gamma$ is also called the ***characteristic function*** of Φ.

Example A.3.30 Let $G = \mathbb{R}^n$, equipped of course with its Borel σ-algebra. The vector space Γ of linear functions $x \mapsto \langle\xi|x\rangle$, one for every $\xi \in \mathbb{R}^n$, generates the topology of $\mathbb{R}^n$ and therefore also generates $\mathcal{B}^\bullet(\mathbb{R}^n)$. Thus any measure μ of finite variation on $\mathbb{R}^n$ is determined by its characteristic function for Γ

$$\mathfrak{F}[\mu(dx)](\xi) = \widehat{\mu}(\xi) = \int_{\mathbb{R}^n} e^{i\langle\xi|x\rangle}\ \mu(dx)\ , \qquad \xi \in \mathbb{R}^n\ .$$

$\widehat{\mu}$ is a bounded uniformly continuous complex-valued function on the dual $\mathbb{R}^n$ of $\mathbb{R}^n$.

Suppose that μ has a density g with respect to Lebesgue measure λ; that is to say, $\mu(dx) = g(x)\lambda(dx)$. μ has totally finite variation if and only if g is Lebesgue integrable, and in fact $¦\mu¦ = |g|\lambda$. It is customary to write $\widehat{g}$ or $\mathfrak{F}[g(x)]$ for $\widehat{\mu}$ and to call this function the ***Fourier transform***[33] of g (and of μ). The Riemann–Lebesgue lemma says that $\widehat{g} \in C_0(\mathbb{R}^n)$. As g runs through $\mathcal{L}^1(\lambda)$, the $\widehat{g}$ form a subalgebra of $C_0(\mathbb{R}^n)$ that is practically impossible to characterize. It does however contain the ***Schwartz space*** $\mathcal{S}$ of infinitely differentiable functions that together with their partials of any order decay at ∞ faster than $|\xi|^{-k}$ for any $k \in \mathbb{N}$. By theorem A.2.2 this algebra is dense in $C_0(\mathbb{R}^n)$. For $g, h \in \mathcal{S}$ (and whenever both sides make sense) and $1 \le \nu \le n$

$$\mathfrak{F}[ix^\nu g(x)](\xi) = \frac{\partial}{\partial\xi^\nu}\mathfrak{F}[g(x)](\xi) \quad\text{and}\quad \mathfrak{F}\Big[\frac{\partial g(x)}{\partial x^\nu}\Big](\xi) = -i\xi^\nu \cdot \widehat{g}(\xi)$$

$$\widehat{g\star h} = \widehat{g}\cdot\widehat{h}\ ,\ \widehat{g\cdot h} = \widehat{g}\star\widehat{h} \quad\text{and}^{34}\quad \widehat{\overset{*}{\mu}} = \overset{*}{\widehat{\mu}}\ . \tag{A.3.12}$$

[33] Actually it is the widespread custom among analysts to take for Γ the space of linear functionals $y \mapsto 2\pi\langle\xi|x\rangle$, $\xi \in \mathbb{R}^n$, and to call the resulting characteristic function the Fourier transform. This simplifies the Fourier inversion formula (A.3.13) to $g(x) = \int e^{-2\pi i\langle\xi|x\rangle}\widehat{g}(\xi)\,d\xi$.

[34] $\overset{*}{\phi}(x) \stackrel{\text{def}}{=} \phi(-x)$ and $\overset{*}{\mu}(\phi) \stackrel{\text{def}}{=} \mu(\overset{*}{\phi})$ define the ***reflections through the origin*** $\overset{*}{\phi}$ and $\overset{*}{\mu}$. Note the perhaps somewhat unexpected equality $\overset{*}{g\cdot\lambda} = (-1)^n \cdot \overset{*}{g} \cdot \lambda$.

Roughly: the Fourier transform turns partial differentiation into multiplication with $-i$ times the corresponding coordinate function, and vice versa; it turns convolution into the pointwise product, and vice versa. It commutes with ***reflection*** $\mu \mapsto \overset{*}{\mu}$ ***through the origin***. g can be recovered from its Fourier transform $\widehat{g}$ by the ***Fourier inversion formula***

$$g(x) = \mathfrak{F}^{-1}[\widehat{g}](x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\langle \xi | x \rangle} \widehat{g}(\xi)\, d\xi \,. \tag{A.3.13}$$

Example A.3.31 Next let $(G, \mathcal{G})$ be the path space $\mathscr{C}^n$, equipped with its Borel σ-algebra. $\mathcal{G} = \mathcal{B}^\bullet(\mathscr{C}^n)$ is generated by the functions $w \mapsto \langle \alpha | w \rangle_t$ (see page 15). These do not form a vector space, however, so we emulate example A.3.30. Namely, every continuous linear functional on $\mathscr{C}^n$ is of the form

$$w \mapsto \langle w | \gamma \rangle \stackrel{\text{def}}{=} \int_0^\infty \sum_{\nu=1}^n w_t^\nu \, d\gamma_t^\nu \ ,$$

where $\gamma = (\gamma^\nu)_{\nu=1}^n$ is an n-tupel of functions of finite variation and of compact support on the half-line. The continuous linear functionals do form a vector space $\Gamma = \mathscr{C}^{n*}$ that generates $\mathcal{B}^\bullet(\mathscr{C}^n)$ (ibidem). Any law $\mathbb{L}$ on $\mathscr{C}^n$ is therefore determined by its characteristic function

$$\widehat{\mathbb{L}}(\gamma) = \int_{\mathscr{C}^n} e^{i\langle w | \gamma \rangle} \, \mathbb{L}(dw) \,.$$

An aside: the topology generated[14] by Γ is the weak topology $\sigma(\mathscr{C}^n, \mathscr{C}^{n*})$ on $\mathscr{C}^n$ (item A.2.32) and is distinctly weaker than the topology of uniform convergence on compacta.

Example A.3.32 Let $\mathcal{H}$ be a countable index set, and equip the "sequence space" $\mathbb{R}^{\mathcal{H}}$ with the topology of pointwise convergence. This makes $\mathbb{R}^{\mathcal{H}}$ into a Fréchet space. The stochastic analysis of random measures leads to the space $\mathscr{D}_{\mathbb{R}^{\mathcal{H}}}$ of all càdlàg paths $[0, \infty) \to \mathbb{R}^{\mathcal{H}}$ (see page 175). This is a polish space under the Skorohod topology; it is also a vector space, but topology and linear structure do not cooperate to make it into a topological vector space. Yet it is most desirable to have the tool of characteristic functions at one's disposal, since laws on $\mathscr{D}_{\mathbb{R}^{\mathcal{H}}}$ do arise (ibidem). Here is how this can be accomplished. Let Γ denote the vector space of all functions of compact support on $[0, \infty)$ that are continuously differentiable, say. View each $\gamma \in \Gamma$ as the cumulative distribution function of a measure $d\gamma_t = \dot{\gamma}_t dt$ of compact support. Let $\Gamma_0^{\mathcal{H}}$ denote the vector space of all $\mathcal{H}$-tuples $\boldsymbol{\gamma} = \{\gamma^h : h \in \mathcal{H}\}$ of elements of Γ all but finitely many of which are zero. Each $\boldsymbol{\gamma} \in \Gamma_0^{\mathcal{H}}$ is naturally a linear functional on $\mathscr{D}_{\mathbb{R}^{\mathcal{H}}}$, via

$$\mathscr{D}_{\mathbb{R}^{\mathcal{H}}} \ni \boldsymbol{z_.} \mapsto \langle \boldsymbol{z_.} | \boldsymbol{\gamma} \rangle \stackrel{\text{def}}{=} \sum_{h \in \mathcal{H}} \int_0^\infty z_t^h \, d\gamma_t^h \ ,$$

a finite sum. In fact, the $\langle\cdot|\boldsymbol{\gamma}\rangle$ are continuous in the Skorohod topology and separate the points of $\mathscr{D}_{\mathbb{R}^{\mathcal{H}}}$; they form a linear space $\boldsymbol{\Gamma}$ of continuous linear functionals on $\mathscr{D}_{\mathbb{R}^{\mathcal{H}}}$ that separates the points. Therefore, for one good thing, the weak topology $\sigma(\mathscr{D}_{\mathbb{R}^{\mathcal{H}}}, \boldsymbol{\Gamma})$ is a Lusin topology on $\mathscr{D}_{\mathbb{R}^{\mathcal{H}}}$, for which every σ-additive probability is tight, and whose Borels agree with those of the Skorohod topology; and for another, we can define the characteristic function of any probability $\mathbb{P}$ on $\mathscr{D}_{\mathbb{R}^{\mathcal{H}}}$ by

$$\widehat{\mathbb{P}}(\boldsymbol{\gamma}) \stackrel{\text{def}}{=} \mathbb{E}\Big[e^{i\langle\cdot|\boldsymbol{\gamma}\rangle}\Big] .$$

To amplify on examples A.3.31 and A.3.32 and to prepare the way for an easy proof of the Central Limit Theorem A.4.4 we provide here a simple result:

Lemma A.3.33 *Let Γ be a real vector space of real-valued functions on some set E. The topologies generated*[14] *by Γ and by the collection $e^{i\Gamma}$ of functions $x \mapsto e^{i\gamma(x)}$, $\gamma \in \Gamma$, have the same convergent sequences.*

Proof. It is evident that the topology generated by $e^{i\Gamma}$ is coarser than the one generated by Γ. For the converse, let (x_n) be a sequence that converges to $x \in E$ in the former topology, i.e., so that $e^{i\gamma(x_n)} \to e^{i\gamma(x)}$ for all $\gamma \in \Gamma$. Set $\delta_n = \gamma(x_n) - \gamma(x)$. Then $e^{it\delta_n} \to 1$ for all t. Now

$$\begin{aligned}\frac{1}{K}\int_{-K}^{K} 1 - e^{it\delta_n}\, dt &= 2\left(1 - \frac{1}{K}\int_0^K \cos(t\delta_n)\, dt\right)\\ &= 2\left(1 - \frac{\sin(\delta_n K)}{\delta_n K}\right) \geq 2\left(1 - \frac{1}{|\delta_n K|}\right) .\end{aligned}$$

For sufficiently large indices $n \geq n(K)$ the left-hand side can be made smaller than 1, which implies $1/|\delta_n K| \geq 1/2$ and $|\delta_n| \leq 2/K$: $\delta_n \to 0$ as desired. ∎

The conclusion may fail if Γ is merely a vector space over the rationals $\mathbb{Q}$: consider the $\mathbb{Q}$-vector space Γ of rational linear functions $x \mapsto qx$ on $\mathbb{R}$. The sequence $(2\pi n!)$ converges to zero in the topology generated by $e^{i\Gamma}$, but not in the topology generated by Γ, which is the usual one. On subsets of E that are precompact in the Γ-topology, the Γ-topology and the $e^{i\Gamma}$-topology coincide, of course, whatever Γ. However,

Exercise A.3.34 A sequence $(\boldsymbol{x}_n)$ in $\mathbb{R}^d$ converges if and only if $(e^{i\langle\boldsymbol{\xi}|\boldsymbol{x}_n\rangle})$ converges for almost all $\boldsymbol{\xi} \in \mathbb{R}^d$.

Exercise A.3.35 If $\widehat{\mathbb{L}_1}(\gamma) = \widehat{\mathbb{L}_2}(\gamma)$ for all γ in the real vector space Γ, then $\mathbb{L}_1$ and $\mathbb{L}_2$ agree on the σ-algebra generated by Γ.

Independence On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ consider n $\mathbb{P}$-measurable maps $\Phi_\nu : \Omega \to E_\nu$, where E_ν is equipped with the algebra $\mathcal{E}_\nu$ of elementary integrands. If the law of the product map $(\Phi_1, \ldots, \Phi_n) : \Omega \to E_1 \times \cdots \times E_n$ – which is clearly $\mathbb{P}$-measurable if $E_1 \times \cdots \times E_n$ is equipped with $\mathcal{E}_1 \otimes \cdots \otimes \mathcal{E}_n$ –

happens to coincide with the product of the laws $\Phi_1[\mathbb{P}], \ldots, \Phi_n[\mathbb{P}]$, then one says that the family $\{\Phi_1, \ldots, \Phi_n\}$ is ***independent*** under $\mathbb{P}$. This definition generalizes in an obvious way to countable collections $\{\Phi_1, \Phi_2, \ldots\}$ (page 404).

Suppose $\mathcal{F}_1, \mathcal{F}_2, \ldots$ are sub-σ-algebras of $\mathcal{F}$. With each goes the (trivially measurable) identity map $\Phi_n : (\Omega, \mathcal{F}) \to (\Omega, \mathcal{F}_n)$. The σ-algebras $\mathcal{F}_n$ are called independent if the Φ_n are.

Exercise A.3.36 Suppose that the sequential closure of $\mathcal{E}_\nu$ is generated by the vector space Γ_ν of real-valued functions on E_ν. Write Φ for the product map $\prod_{\nu=1}^n \Phi_\nu$ from Ω to $\prod_\nu E_\nu$. Then $\Gamma \stackrel{\text{def}}{=} \bigotimes_{\nu=1}^n \Gamma_\nu$ generates the sequential closure of $\bigotimes_{\nu=1}^n \mathcal{E}_\nu$, and $\{\Phi_1, \ldots, \Phi_n\}$ is independent if and only if

$$\widehat{\Phi[\mathbb{P}]}^{\Gamma}(\gamma_1 \otimes \cdots \otimes \gamma_n) = \prod_{1 \le \nu \le n} \widehat{\Phi_\nu[\mathbb{P}]}^{\Gamma_\nu}(\gamma_\nu) .$$

Convolution

Fix a commutative locally compact group G whose topology has a countable basis. The group operation is denoted by $+$ or by juxtaposition, and it is understood that group operation and topology are compatible in the sense that "the subtraction map" $- : G \times G \to G$, $(g, g') \mapsto g - g'$, is continuous. In the instances that occur in the main body $(G, +)$ is either $\mathbb{R}^n$ with its usual addition or $\{-1, 1\}^n$ under pointwise multiplication. On such a group there is an essentially unique translation-invariant[35] Radon measure η called ***Haar measure***. In the case $G = \mathbb{R}^n$, Haar measure is taken to be Lebesgue measure, so that the mass of the unit box is unity; in the second example it is the normalized counting measure, which gives every point equal mass 2^{-n} and makes it a probability.

Let μ_1 and μ_2 be two Radon measures on G that have bounded variation: $\|\mu_i\| \stackrel{\text{def}}{=} \sup\{\mu_i(\phi) : \phi \in C_{00}(G) \ , \ |\phi| \le 1\} < \infty$. Their ***convolution*** $\mu_1 \star \mu_2$ is defined by

$$\mu_1 \star \mu_2(\phi) = \int_{G \times G} \phi(g_1 + g_2) \ \mu_1(dg_1) \mu_2(dg_2) . \tag{A.3.14}$$

In other words, apply the product $\mu_1 \times \mu_2$ to the particular class of functions $(g_1, g_2) \mapsto \phi(g_1 + g_2)$, $\phi \in C_{00}(G)$. It is easily seen that $\mu_1 \star \mu_2$ is a Radon measure on $C_{00}(G)$ of total variation $\|\mu_1 \star \mu_2\| \le \|\mu_1\| \cdot \|\mu_2\|$, and that convolution is associative and commutative. The usual sequential closure argument shows that equation (A.3.14) persists if ϕ is a bounded Baire function on G.

Suppose μ_1 has a Radon–Nikodym derivative h_1 with respect to Haar measure: $\mu_1 = h_1 \eta$ with $h_1 \in \mathcal{L}^1[\eta]$. If ϕ in equation (A.3.14) is negligible for Haar measure, then $\mu_1 \star \mu_2$ vanishes on ϕ by Fubini's theorem A.3.18.

[35] This means that $\int \phi(x + g) \ \eta(dx) = \int \phi(x) \ \eta(dx)$ for all $g \in G$ and $\phi \in C_{00}(G)$ and persists for η-integrable functions ϕ. If η is translation-invariant, then so is $c\eta$ for $c \in \mathbb{R}$, but this is the only ambiguity in the definition of Haar measure.

Therefore $\mu_1\star\mu_2$ is absolutely continuous with respect to Haar measure. Its density is then denoted by $h_1\star\mu_2$ and can be calculated easily:

$$\big(h_1\star\mu_1\big)(g) = \int_G h_1(g-g_2)\;\mu_2(dg_2)\;. \tag{A.3.15}$$

Indeed, repeated applications of Fubini's theorem give

$$\begin{aligned}
\int_G \phi(g)\;\mu_1\star\mu_2(dg) \qquad &= \int_{G\times G} \phi\big(g_1+g_2\big)\;h_1\big(g_1\big)\;\eta(dg_1)\,\mu_2(dg_2)\\
\text{by translation-invariance:}\qquad &= \int_{G\times G} \phi\big((g_1-g_2)+g_2)\big)\;h_1\big(g_1-g_2\big)\;\eta(dg_1)\,\mu_2(dg_2)\\
\text{with } g=g_1:\qquad &= \int_{G\times G} \phi\big(g\big)h_1\big(g-g_2\big)\;\eta(dg)\,\mu_2(dg_2)\\
&= \int_G \phi(g)\Big(\int_G h_1(g-g_2)\;\mu_2(dg_2)\Big)\;\eta(dg)\;,
\end{aligned}$$

which exhibits $\int h_1(g-g_2)\;\mu_2(dg_2)$ as the density of $\mu_1\star\mu_2$ and yields equation (A.3.15).

Exercise A.3.37 (i) If $h_1\in C_0(G)$, then $h_1\star\mu_2\in C_0(G)$ as well. (ii) If μ_2, too, has a density $h_2\in\mathcal{L}^1[\eta]$ with respect to Haar measure η, then the density of $\mu_1\star\mu_2$ is commonly denoted by $h_1\star h_2$ and is given by

$$(h_1\star h_2)(g) = \int h_1(g_1)\,h_2(g-g_1)\;\eta(dg_1)\;.$$

Let us compute the characteristic function of $\mu_1\star\mu_2$ in the case $G=\mathbb{R}^n$:

$$\begin{aligned}
\widehat{\mu_1\star\mu_2}(\boldsymbol{\zeta}) &= \int_{\mathbb{R}^n} e^{i\langle\boldsymbol{\zeta}|\boldsymbol{z}_1+\boldsymbol{z}_2\rangle}\;\mu_1(d\boldsymbol{z}_1)\mu_2(d\boldsymbol{z}_2)\\
&= \int e^{i\langle\boldsymbol{\zeta}|\boldsymbol{z}_1\rangle}\;\mu_1(d\boldsymbol{z}_1)\cdot\int e^{i\langle\boldsymbol{\zeta}|\boldsymbol{z}_2\rangle}\;\mu_2(d\boldsymbol{z}_2)\\
&= \widehat{\mu_1}(\boldsymbol{\zeta})\cdot\widehat{\mu_2}(\boldsymbol{\zeta})\;.
\end{aligned} \tag{A.3.16}$$

Exercise A.3.38 Convolution commutes with reflection through the origin[34]: if $\mu=\mu_1\star\mu_2$, then $\overset{*}{\mu}=\overset{*}{\mu}_1\star\overset{*}{\mu}_2$.

Liftings, Disintegration of Measures

For the following fix a σ-algebra $\mathcal{F}$ on a set F and a positive σ-finite measure μ on $\mathcal{F}$ (exercise A.3.27). We assume that $(\mathcal{F},\mu)$ is complete, i.e., that $\mathcal{F}$ equals the ***μ-completion*** $\mathcal{F}^\mu$, the σ-algebra generated by $\mathcal{F}$ and all subsets of μ-negligible sets in $\mathcal{F}$. We distinguish carefully between a function $f\in\mathcal{L}^\infty$[36] and its class modulo negligible functions $\dot f\in L^\infty$, writing $f\dot\le g$ to mean that $\dot f\le\dot g$, i.e., that $\dot f,\dot g$ contain representatives f',g' with $f'(x)\le g'(x)$ at all points $x\in F$, etc.

[36] f is $\mathcal{F}$-measurable and bounded.

Definition A.3.39 *(i) A* ***density*** *on* $(F,\mathcal{F},\mu)$ *is a map* $\theta:\mathcal{F}\to\mathcal{F}$ *with the following properties:*
a) $\theta(\emptyset)=\emptyset$ *and* $\theta(F)=F$ *; b)* $A\dot{\subseteq}B \Longrightarrow \theta(A)\subseteq\theta(B)\in\dot{B} \quad \forall\, A,B\in\mathcal{F}$ *;*
c) $A_1\cap\ldots\cap A_k\doteq\emptyset \Longrightarrow \theta(A_1)\cap\ldots\cap\theta(A_k)=\emptyset \quad \forall\, k\in\mathbb{N}, A_1,\ldots,A_k\in\mathcal{F}$ *.*

(ii) A ***dense topology*** *is a topology* $\tau\subset\mathcal{F}$ *with the following properties: a) a negligible set in* τ *is void; and b) every set of* $\mathcal{F}$ *contains a* τ*-open set from which it differs negligibly.*

(iii) A ***lifting*** *is an algebra homomorphism* $T:\mathcal{L}^\infty\to\mathcal{L}^\infty$ *that takes the constant function* 1 *to itself and obeys* $f\doteq g\Longrightarrow Tf=Tg\in\dot{g}$ *.*

Viewed as a map $T:L^\infty\to\mathcal{L}^\infty$, a lifting T is a linear multiplicative inverse of the natural quotient map $\dot{}\ : f\mapsto\dot{f}$ from $\mathcal{L}^\infty$ to L^∞. A lifting T is positive; for if $0\le f\in\mathcal{L}^\infty$, then f is the square of some function g and thus $Tf=(Tg)^2\ge 0$. A lifting T is also contractive; for if $\|f\|_\infty\le a$, then $-a\le f\le a\Longrightarrow -a\le Tf\le a\Longrightarrow \|Tf\|_\infty\le a$.

Lemma A.3.40 *Let* $(F,\mathcal{F},\mu)$ *be a complete totally finite measure space.*

(i) If $(F,\mathcal{F},\mu)$ *admits a density* θ*, then it has a dense topology* τ_θ *that contains the family* $\{\theta(A):A\in\mathcal{F}\}$ *.*

(ii) Suppose $(F,\mathcal{F},\mu)$ *admits a dense topology* τ *. Then every function* $f\in\mathcal{L}^\infty$ *is* μ*-almost surely* τ*-continuous, and there exists a lifting* T_τ *such that* $T_\tau f(x)=f(x)$ *at all* τ*-continuity points* x *of* f *.*

(iii) If $(F,\mathcal{F},\mu)$ *admits a lifting, then it admits a density.*

Proof. (i) Given a density θ, let τ_θ be the topology generated by the sets $\theta(A)\backslash N$, $A\in\mathcal{F}$, $\mu(N)=0$. It has the basis τ_0 of sets of the form

$$\bigcap\nolimits_{i=1}^{I}\theta(A_i)\setminus N_i\ , \qquad\qquad I\in\mathbb{N}\,,\ A_i\in\mathcal{F}\,,\ \mu(N_i)=0\,.$$

If such a set is negligible, it must be void by A.3.39 (ic). Also, any set $A\in\mathcal{F}$ is μ-almost surely equal to its τ_θ-open subset $\theta(A)\cap A$. The only thing perhaps not quite obvious is that $\tau_\theta\subset\mathcal{F}$. To see this, let $U\in\tau_\theta$. There is a subfamily $\mathcal{U}\subset\tau_0\subset\mathcal{F}$ with union U. The family $\mathcal{U}^{\cup f}\subset\mathcal{F}$ of finite unions of sets in $\mathcal{U}$ also has union U. Set $u=\sup\{\mu(B):B\in\mathcal{U}^{\cup f}\}$, let $\{B_n\}$ be a countable subset of $\mathcal{U}^{\cup f}$ with $u=\sup_n\mu(B_n)$, and set $B=\bigcup_n B_n$ and $C=\theta(F\backslash B)$. Thanks to A.3.39(ic), $B\cap C=\emptyset$ for all $B\in\mathcal{U}$, and thus $B\subset U\subset C^c\doteq B$. Since $\mathcal{F}$ is μ-complete, $U\in\mathcal{F}$.

(ii) A set $A\in\mathcal{F}$ is evidently continuous on its τ-interior and on the τ-interior of its complement A^c; since these two sets add up almost everywhere to F, A is almost everywhere τ-continuous. A linear combination of sets in $\mathcal{F}$ is then clearly also almost everywhere continuous, and then so is the uniform limit of such. That is to say, every function in $\mathcal{L}^\infty$ is μ-almost everywhere τ-continuous.

By theorem A.2.2 there exists a map j from F into a compact space $\widehat{F}$ such that $\widehat{f}\mapsto\widehat{f}\circ j$ is an isometric algebra isomorphism of $C(\widehat{F})$ with $\mathcal{L}^\infty$.

Fix a point $x \in F$ and consider the set $\mathcal{I}_x$ of functions $f \in \mathcal{L}^\infty$ that differ negligibly from a function $f' \in \mathcal{L}^\infty$ that is zero and τ-continuous at x. This is clearly an ideal of $\mathcal{L}^\infty$. Let $\widehat{\mathcal{I}}_x$ denote the corresponding ideal of $C(\widehat{F})$. Its zero-set $\widehat{Z}_x$ is not void. Indeed, if it were, then there would be, for every $\widehat{y} \in \widehat{F}$, a function $\widehat{f}_y \in \widehat{\mathcal{I}}_x$ with $\widehat{f}_y(\widehat{y}) \neq 0$. Compactness would produce a finite subfamily $\{\widehat{f}_{y^i}\}$ with $\widehat{f} \stackrel{\text{def}}{=} \sum \widehat{f}_{y^i}^2 \in \widehat{\mathcal{I}}_x$ bounded away from zero. The corresponding function $f = \widehat{f} \circ j \in \mathcal{I}_x$ would also be bounded away from zero, say $f > \epsilon > 0$. For a function $f' \doteq f$ continuous at x, $[f' < \epsilon]$ would be a negligible τ-neighborhood of x, necessarily void. This contradiction shows that $\widehat{Z}_x \neq \emptyset$.

Now pick for every $x \in F$ a point $\widehat{x} \in \widehat{Z}_x$ and set

$$T_\tau f(x) \stackrel{\text{def}}{=} \widehat{f}(\widehat{x}) , \qquad f \in \mathcal{L}^\infty .$$

This is the desired lifting. Clearly T_τ is linear and multiplicative. If $f \in \mathcal{L}^\infty$ is τ-continuous at x, then $g \stackrel{\text{def}}{=} f - f(x) \in \mathcal{I}_x$, $\widehat{g} \in \widehat{\mathcal{I}}_x$, and $\widehat{g}(\widehat{x}) = 0$, which signifies that $Tg(x) = 0$ and thus $T_\tau f(x) = f(x)$: the function $T_\tau f$ differs negligibly from f, namely, at most in the discontinuity points of f. If f, g differ negligibly, then $f - g$ differs negligibly from the function zero, which is τ-continuous at all points. Therefore $f - g \in \mathcal{I}_x \quad \forall x$ and thus $T(f - g) = 0$ and $Tf = Tg$.

(iii) Finally, if T is a lifting, then its restriction to the sets of $\mathcal{F}$ (see convention A.1.5) is plainly a density. ∎

Theorem A.3.41 *Let $(F, \mathcal{F}, \mu)$ be a σ-finite measure space (exercise A.3.27) and denote by $\mathcal{F}^\mu$ the μ-completion of $\mathcal{F}$.*

(i) There exists a lifting T for $(F, \mathcal{F}^\mu, \mu)$.

(ii) Let $\mathcal{C}$ be a countable collection of bounded $\mathcal{F}$-measurable functions. There exists a set $G \in \mathcal{F}$ with $\mu(G^c) = 0$ such that $G \cdot Tf = G \cdot f$ for all f that lie in the algebra generated by $\mathcal{C}$ or in its uniform closure, in fact for all bounded f that are continuous in the topology generated by $\mathcal{C}$.

Proof. (i) We assume to start with that $\mu \geq 0$ is finite. Consider the set $\mathfrak{L}$ of all pairs $(\mathcal{A}, T^{\mathcal{A}})$, where $\mathcal{A}$ is a sub-σ-algebra of $\mathcal{F}^\mu$ that contains all μ-negligible subsets of $\mathcal{F}^\mu$ and $T^{\mathcal{A}}$ is a lifting on $(F, \mathcal{A}, \mu)$. $\mathfrak{L}$ is not void: simply take for $\mathcal{A}$ the collection of negligible sets and their complements and set $T^{\mathcal{A}} f = \int f \, d\mu$. We order $\mathfrak{L}$ by saying $(\mathcal{A}, T^{\mathcal{A}}) \ll (\mathcal{B}, T^{\mathcal{B}})$ if $\mathcal{A} \subset \mathcal{B}$ and the restriction of $T^{\mathcal{B}}$ to $\mathcal{L}^\infty(\mathcal{A})$ is $T^{\mathcal{A}}$. The proof of the theorem consists in showing that this order is inductive and that a maximal element has σ-algebra $\mathcal{F}^\mu$.

Let then $\mathfrak{C} = \{(\mathcal{A}_\sigma, T^{\mathcal{A}_\sigma}) : \sigma \in \Sigma\}$ be a chain for the order $\ll$. If the index set Σ has no countable cofinal subset, then it is easy to find an upper bound for $\mathfrak{C}$: $\mathcal{A} \stackrel{\text{def}}{=} \bigcup_{\sigma \in \Sigma} \mathcal{A}_\sigma$ is a σ-algebra and $T^{\mathcal{A}}$, defined to coincide with $T^{\mathcal{A}_\sigma}$ on $\mathcal{A}_\sigma$ for $\sigma \in \Sigma$, is a lifting on $\mathcal{L}^\infty(\mathcal{A})$. Assume then that Σ does have a countable ***cofinal subset*** – that is to say, there exists a countable subset $\Sigma_0 \subset \Sigma$ such that every $\sigma \in \Sigma$ is exceeded by some index in Σ_0. We may then assume as well that $\Sigma = \mathbb{N}$. Letting $\mathcal{B}$ denote the σ-algebra generated by $\bigcup_n \mathcal{A}_n$ we define a density θ on $\mathcal{B}$ as follows:

$$\theta(B) \stackrel{\text{def}}{=} \Big[\lim_{n\to\infty} T^{\mathcal{A}_n}\mathbb{E}^\mu\big[B|\mathcal{A}_n\big] = 1\Big]\,, \qquad B \in \mathcal{B}\,.$$

The uniformly integrable martingale $\mathbb{E}^\mu\big[B|\mathcal{A}_n\big]$ converges μ-almost everywhere to B (page 75), so that $\theta(B) = B$ μ-almost everywhere. Properties a) and b) of a density are evident; as for c), observe that $B_1\cap\ldots\cap B_k \doteq \emptyset$ implies $B_1+\cdots+B_k \stackrel{.}{\le} k-1$, so that due to the linearity of the $\mathbb{E}[\cdot|\mathcal{A}_n]$ and $T^{\mathcal{A}_n}$ not all of the $\theta(B_i)$ can equal 1 at any one point: $\theta(B_1)\cap\ldots\cap\theta(B_k) = \emptyset$. Now let τ denote the dense topology τ_θ provided by lemma A.3.40 and $T^{\mathcal{B}}$ the lifting T_τ (ibidem). If $A \in \mathcal{A}_n$, then $\theta(A) = T^{\mathcal{A}_n}A$ is τ-open, and so is $T^{\mathcal{A}_n}A^c{=}1 - T^{\mathcal{A}_n}A$. This means that $T^{\mathcal{A}_n}A$ is τ-continuous at all points, and therefore $T^{\mathcal{B}}A = T^{\mathcal{A}_n}A$: $T^{\mathcal{B}}$ extends $T^{\mathcal{A}_n}$ and $(\mathcal{B}, T^{\mathcal{B}})$ is an upper bound for our chain. Zorn's lemma now provides a maximal element $(\mathcal{M}, T^{\mathcal{M}})$ of $\mathfrak{L}$.

It is left to be shown that $\mathcal{M} = \mathcal{F}$. By way of contradiction assume that there exists a set $G \in \mathcal{F}$ that does not belong to $\mathcal{M}$. Let $\tau^{\mathcal{M}}$ be the dense topology that comes with $T^{\mathcal{M}}$ considered as a density. Let $\underline{\mathring{G}} \stackrel{\text{def}}{=} \bigcup\{U \in \tau^{\mathcal{M}} : U \stackrel{.}{\subset} G\}$ denote the *essential interior* of G, and replace G by the equivalent set $(G \cup \underline{\mathring{G}}) \setminus \underline{\mathring{G^c}}$. Let $\mathcal{N}$ be the σ-algebra generated by $\mathcal{M}$ and G,

$$\mathcal{N} = \big\{(M\cap G)\cup(M'\cap G^c) : \ M, M' \in \mathcal{M}\big\}\,,$$

and

$$\tau^{\mathcal{N}} = \big\{\,(U\cap G)\,\cup\,(U'\cap G^c)\ :\ U, U' \in \tau^{\mathcal{M}}\big\}$$

the topology generated by $\tau^{\mathcal{M}}, G, G^c$. A little set algebra shows that $\tau^{\mathcal{N}}$ is a dense topology for $\mathcal{N}$ and that the lifting $T^{\mathcal{N}}$ provided by lemma A.3.40 (ii) extends $T^{\mathcal{M}}$. $(\mathcal{N}, T^{\mathcal{N}})$ strictly exceeds $(\mathcal{M}, T^{\mathcal{M}})$ in the order $\ll$, which is the desired contradiction.

If μ is merely σ-finite, then there is a countable collection $\{F_1, F_2, \ldots\}$ of mutually disjoint sets of finite measure in $\mathcal{F}$ whose union is F. There are liftings T_n for μ on the restriction of $\mathcal{F}$ to F_n. We glue them together: $T : f \mapsto \sum T_n(F_n{\cdot}f)$ is a lifting for $(F, \mathcal{F}, \mu)$.

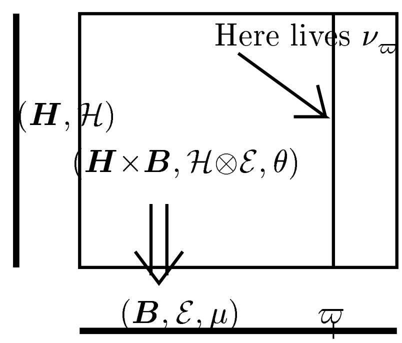
(ii) $\bigcup_{f\in\mathcal{C}}[Tf \ne f]$ is contained in a μ-negligible subset $B \in \mathcal{F}$ (see A.3.8). Set $G = B^c$. The $f \in \mathcal{L}^\infty$ with $GTf = Gf \in \mathcal{F}$ form a uniformly closed algebra that contains the algebra $\mathcal{A}$ generated by $\mathcal{C}$ and its uniform closure $\overline{\mathcal{A}}$, which is a vector lattice (theorem A.2.2) generating the same topology $\tau_{\mathcal{C}}$ as $\mathcal{C}$. Let h be bounded and continuous in that topology. There exists an increasingly directed family $\overline{\mathcal{A}}^h \subset \overline{\mathcal{A}}$ whose pointwise supremum is h (lemma A.2.19). Let $G' = G \cap TG$. Then $G'h = \sup G'\overline{\mathcal{A}}^h = \sup G'T\overline{\mathcal{A}}^h$ is lower semicontinuous in the dense topology τ_T of T. Applying this to $-h$ shows that $G'h$ is upper semicontinuous as well, so it is τ_T-continuous and therefore μ-measurable. Now $GTh \ge \sup GT\overline{\mathcal{A}}^h = \sup G\overline{\mathcal{A}}^h = Gh$. Applying this to $-h$ shows that $GTh \le Gh$ as well, so that $GTh = Gh \in \mathcal{F}$. ∎

Corollary A.3.42 (Disintegration of Measures) *Let $\boldsymbol{H}$ be a locally compact space with a countable basis for its topology and equip it with the algebra $\mathcal{H} = C_{00}(\boldsymbol{H})$ of continuous functions of compact support; let $\boldsymbol{B}$ be a set equipped with $\mathcal{E}$, a σ-finite algebra or vector lattice closed under chopping of bounded functions, and let θ be a positive σ-additive measure on $\mathcal{H} \otimes \mathcal{E}$.*

There exist a positive measure μ on $\mathcal{E}$ and a slew $\varpi \mapsto \nu_\varpi$ of positive Radon measures, one for every $\varpi \in \boldsymbol{B}$, having the following two properties: (i) for every $\phi \in \mathcal{H} \otimes \mathcal{E}$ the function $\varpi \mapsto \int \phi(\eta, \varpi)\, \nu_\varpi(d\eta)$ is measurable on the σ-algebra $\mathcal{P}$ generated by $\mathcal{E}$; (ii) for every θ-integrable function $f : \boldsymbol{H} \times \boldsymbol{B} \to \mathbb{R}$, $f(\cdot, \varpi)$ is ν_ϖ-integrable for μ-almost all $\varpi \in \boldsymbol{B}$, the function $\varpi \mapsto \int f(\eta, \varpi)\, \nu_\varpi(d\eta)$ is μ-integrable, and

$$\int_{\boldsymbol{H}\times\boldsymbol{B}} f(\eta, \varpi)\, \theta(d\eta, d\varpi) = \int_{\boldsymbol{B}} \int_{\boldsymbol{H}} f(\eta, \varpi) \nu_\varpi(d\eta)\, \mu(d\varpi)\, .$$

If $\theta(1_{\boldsymbol{H}} \otimes X) < \infty$ for all $X \in \mathcal{E}$, then the ν_ϖ can be chosen to be probabilities.



Proof. There is an increasing sequence of functions $X_i \in \mathcal{E}$ with pointwise supremum 1. The sets $P_i \stackrel{\text{def}}{=} [X_i > 1/i]$ belong to the sequential closure $\mathcal{E}^\sigma$ of $\mathcal{E}$ and increase to $\boldsymbol{B}$ (lemma A.3.3). Let $\mathcal{E}_0^\sigma$ denote the collection of those bounded functions in $\mathcal{E}^\sigma$ that vanish off one of the P_i. There is an obvious extension of θ to $\mathcal{H} \otimes \mathcal{E}_0^\sigma$. We shall denote it again by θ and prove the corollary with $\big(\mathcal{E}, \theta\big)$ replaced by $\big(\mathcal{E}_0^\sigma, \theta\big)$. The original claim is then immediate from the observation that every function $\phi \in \mathcal{E}$ is the dominated limit of the sequence $\phi{\cdot}P_i \in \mathcal{E}_0^\sigma$. There is also an increasing sequence of compacta $K_i \subset \boldsymbol{H}$ whose interiors cover $\boldsymbol{H}$.

For any $h \in \mathcal{H}$ consider the map $\mu^h : \mathcal{E}_0^\sigma \to \mathbb{R}$ defined by

$$\mu^h(X) = \int h \cdot X \, d\theta\, , \qquad X \in \mathcal{E}_0^\sigma$$

and set

$$\mu = \sum_i a_i \mu^{K_i}\, ,$$

where the $a_i > 0$ are chosen so that $\sum a_i \mu^{K_i}(P_i) < \infty$. Then μ^h is a σ-finite measure whose variation $\mu^{|h|}$ is majorized by a multiple of μ. Indeed, some K_i contains the support of h, and then $\mu^{|h|} \le a_i^{-1} \|h\|_\infty {\cdot} \mu$. There exists a bounded Radon–Nikodym derivative $\dot{g}^h = d\mu^h \big/ d\mu$. Fix now a lifting $T : L^\infty(\mu) \to \mathcal{L}^\infty(\mu)$, producing the set $\mathcal{C}$ of theorem A.3.41 (ii) by picking, for every h in a countable subcollection of $\mathcal{H}$ that generates the topology, a representative $g^h \in \dot{g}^h$ that is measurable on $\mathcal{P}$. There then exists a set

$G \in \mathcal{P}$ of μ-negligible complement such that $G \cdot g^h = G \cdot T\dot{g}^h$ for all $h \in \mathcal{H}$. We define now the maps $\nu_\varpi : \mathcal{H} \to \mathbb{R}$, one for every $\varpi \in \boldsymbol{B}$, by

$$\nu_\varpi(h) = G(\varpi) \cdot T_\varpi(\dot{g}^h) \ .$$

As positive linear functionals on $\mathcal{H}$ the ν_ϖ are Radon measures, and for every $h \in \mathcal{H}$, $\varpi \mapsto \nu_\varpi(h)$ is $\mathcal{P}$-measurable. Let $\phi = \sum_k h_k Y_k \in \mathcal{H} \otimes \mathcal{E}_0^\sigma$. Then

$$\begin{aligned}\int \phi(\eta, \varpi)\ \theta(d\eta, d\varpi) &= \sum_k \int Y_k(\varpi)\ \mu^{h_k}(d\varpi) = \sum_k \int Y_k \cdot \dot{g}^{h_k}\ d\mu \\ &= \sum_k \int Y_k \int h_k(\eta)\ \nu_\varpi(d\eta)\ \mu(d\varpi) \\ &= \int \int \phi(\eta, \varpi)\ \nu_\varpi(d\eta)\ \mu(d\varpi) \ .\end{aligned}$$

The functions ϕ for which the left-hand side and the ultimate right-hand side agree and for which $\varpi \mapsto \int \phi(\eta, \varpi)\ \nu_\varpi(d\eta)$ is measurable on $\mathcal{P}$ form a collection closed under $\mathcal{H} \otimes \mathcal{E}$-dominated sequential limits and thus contains $\big(\mathcal{H} \otimes \mathcal{E}\big)_{00}^\sigma$. This proves (i). Theorem A.3.18 on page 403 yields (ii). The last claim is left to the reader. ▮

Exercise A.3.43 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $\mathcal{C}$ a countable collection of μ-measurable functions, and τ the topology generated by $\mathcal{C}$. Every τ-continuous function is μ-measurable.

Exercise A.3.44 Let E be a separable metric space and $\mu : C_b(E) \to \mathbb{R}$ a σ-continuous positive measure. There exists a ***strong lifting***, that is to say, a lifting $T : \mathcal{L}^\infty(\mu) \to \mathcal{L}^\infty(\mu)$ such that $T\phi(x) = \phi(x)$ for all $\phi \in C_b(E)$ and all x in the support of μ.

Gaussian and Poisson Random Variables

The ***centered Gaussian distribution*** with ***variance*** t is denoted by γ_t:

$$\gamma_t(dx) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}\, dx \ .$$

A real-valued random variable X whose law is γ_t is also said to be $N(0,t)$, pronounced "normal zero–t." The ***standard deviation*** of such X or of γ_t by definition is $\sqrt{t}$; if it equals 1, then X is said to be a ***normalized Gaussian***. Here are a few elementary integrals involving Gaussian distributions. They are used on various occasions in the main text. $|\boldsymbol{a}|$ stands for the Euclidean norm $|\boldsymbol{a}|_2$.

Exercise A.3.45 A $N(0,t)$-random variable X has expectation $\mathbb{E}[X] = 0$, variance $\mathbb{E}[X^2] = t$, and its characteristic function is

$$\mathbb{E}\Big[e^{i\xi X}\Big] = \int_{\mathbb{R}} e^{i\xi x} \gamma_t(dx) = e^{-t\xi^2/2} \ . \tag{A.3.17}$$

Exercise A.3.46 The ***Gamma function*** Γ, defined for complex z with a strictly positive real part by

$$\Gamma(z) = \int_0^\infty u^{z-1} e^{-u} \, du \,,$$

is convex on $(0,\infty)$ and satisfies $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(1) = 1$, $\Gamma(z+1) = z \cdot \Gamma(z)$, and $\Gamma(n+1) = n!$ for $n \in \mathbb{N}$.

Exercise A.3.47 The Gauß kernel has moments

$$\int_{-\infty}^{+\infty} |x|^p \, \gamma_t(dx) = \frac{(2t)^{p/2}}{\sqrt{\pi}} \cdot \Gamma\Big(\frac{p+1}{2}\Big) \,, \qquad (p > -1).$$

Now let $X_1, \ldots, X_n$ be independent Gaussians distributed $N(0,t)$. Then the distribution of the vector $\boldsymbol{X} = (X_1, \ldots, X_n) \in \mathbb{R}^n$ is $\gamma_{tI}(\boldsymbol{x}) d\boldsymbol{x}$, where

$$\gamma_{tI}(\boldsymbol{x}) \stackrel{\text{def}}{=} \frac{1}{(\sqrt{2\pi t})^n} \, e^{-|\boldsymbol{x}|^2/2t}$$

is the n-dimensional ***Gauß kernel*** or ***heat kernel***. I indicates the identity matrix.

Exercise A.3.48 The characteristic function of the vector $\boldsymbol{X}$ (or of γ_{tI}) is $e^{-t|\boldsymbol{\xi}|^2/2}$. Consequently, the law of $\boldsymbol{\xi}$ is invariant under rotations of $\mathbb{R}^n$. Next let $0 < p < \infty$ and assume that $t = 1$. Then

$$\frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^n} |\boldsymbol{x}|^p \, e^{-|\boldsymbol{x}|^2/2} \, dx_1 \, dx_2 \ldots dx_n = \frac{2^{p/2} \cdot \Gamma\left(\frac{n+p}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \,,$$

and for any vector $\boldsymbol{a} \in \mathbb{R}^n$

$$\frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^n} \Big| \sum_{\nu=1}^n x_\nu \cdot a_\nu \Big|^p \, e^{-|\boldsymbol{x}|^2/2} \, dx_1 \ldots dx_n = \frac{|\boldsymbol{a}|^p \cdot 2^{p/2} \Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}} \,.$$

Consider next a symmetric ***positive semidefinite*** $d \times d$-matrix B; that is to say, $x_\eta x_\theta B^{\eta\theta} \ge 0$ for every $\boldsymbol{x} \in \mathbb{R}^d$.

Exercise A.3.49 There exists a matrix U that depends continuously on B such that $B^{\eta\theta} = \sum_{\iota=1}^n U_\iota^\eta U_\iota^\theta$.

Definition A.3.50 *The* ***Gaussian with covariance matrix B*** *or* ***centered normal distribution with covariance matrix B*** *is the image of the heat kernel γ_I under the linear map $U : \mathbb{R}^d \to \mathbb{R}^d$.*

Exercise A.3.51 The name is justified by these facts: the ***covariance matrix*** $\int_{\mathbb{R}^d} x^\eta x^\theta \, \gamma_B(d\boldsymbol{x})$ equals $B^{\eta\theta}$; for any $t > 0$ the characteristic function of γ_{tB} is given by

$$\widehat{\gamma_{tB}}(\boldsymbol{\xi}) = e^{-t\xi_\eta \xi_\theta B^{\eta\theta}/2} \,.$$

Changing topics: a random variable N that takes only positive integer values is ***Poisson*** with mean $\lambda > 0$ if

$$\mathbb{P}[N = n] = e^{-\lambda} \frac{\lambda^n}{n!} \,, \qquad n = 0, 1, 2 \ldots \,.$$

Exercise A.3.52 Its characteristic function $\widehat{N}$ is given by

$$\widehat{N}(\alpha) \stackrel{\text{def}}{=} \mathbb{E}\Big[e^{i\alpha N}\Big] = e^{\lambda(e^{i\alpha}-1)} \,.$$

The sum of independent Poisson random variables N_i with means λ_i is Poisson with mean $\sum \lambda_i$.

A.4 Weak Convergence of Measures

In this section we fix a completely regular space E and consider σ-continuous measures μ of finite total variation $\lvert\mu\rvert(1) = \|\mu\|$ on the lattice algebra $C_b(E)$. Their collection is $\mathfrak{M}^*(E)$. Each has an extension that integrates all bounded Baire functions and more. The order-continuous elements of $\mathfrak{M}^*(E)$ form the collection $\mathfrak{M}^{\bullet}(E)$. The positive[22] σ-continuous measures of total mass 1 are the ***probabilities*** on E, and their collection is denoted by $\mathfrak{M}^*_{1,+}(E)$ or $\mathfrak{P}^*(E)$. We shall be concerned mostly with the order-continuous probabilities on E; their collection is denoted by $\mathfrak{M}^{\bullet}_{1,+}(E)$ or $\mathfrak{P}^{\bullet}(E)$. Recall from exercise A.3.10 that $\mathfrak{P}^*(E) = \mathfrak{P}^{\bullet}(E)$ when E is separable and metrizable. The advantage that the order-continuity of a measure conveys is that every Borel set, in particular every compact set, is integrable with respect to it under the integral extension discussed on pages 398–400.

Equipped with the uniform norm $C_b(E)$ is a Banach space, and $\mathfrak{M}^*(E)$ is a subset of the dual $C_b^*(E)$ of $C_b(E)$. The pertinent topology on $\mathfrak{M}^*(E)$ is the trace of the weak*-topology on $C_b^*(E)$; unfortunately, probabilists call the corresponding notion of convergence ***weak convergence***,[37] and so *nolens volens* will we: a sequence[38] (μ_n) in $\mathfrak{M}^*(E)$ converges weakly to $\mu \in \mathfrak{P}^*(E)$, written $\mu_n \Rightarrow \mu$, if

$$\mu_n(\phi) \xrightarrow[n\to\infty]{} \mu(\phi) \qquad\qquad \forall\, \phi \in C_b(E)\ .$$

In the typical application made in the main body E is a path space $\mathscr{C}$ or $\mathscr{D}$, and μ_n, μ are the laws of processes $X^{(n)}, X$ considered as E-valued random variables on probability spaces $\big(\Omega^{(n)}, \mathcal{F}^{(n)}, \mathbb{P}^{(n)}\big)$, which may change with n. In this case one also writes $X^{(n)} \Rightarrow X$ and says that $X^{(n)}$ ***converges to*** X ***in law*** or ***in distribution***.

It is generally hard to verify the convergence of μ_n to μ on every single function of $C_b(E)$. Our first objective is to reduce the verification to fewer functions.

Proposition A.4.1 *Let $\mathcal{M} \subset C_b(E;\mathbb{C})$ be a multiplicative class that is closed under complex conjugation and generates the topology,*[14] *and let μ_n, μ belong to $\mathfrak{P}^{\bullet}(E)$.*[39] *If $\mu_n(\phi) \to \mu(\phi)$ for all $\phi \in \mathcal{M}$, then $\mu_n \Rightarrow \mu$*[38]*; moreover,*

(i) For h bounded and lower semicontinuous, $\int h\,d\mu \le \liminf_{n\to\infty} \int h\,d\mu_n$; and for k bounded and upper semicontinuous, $\int k\,d\mu \ge \limsup_{n\to\infty} \int k\,d\mu_n$.

(ii) If f is a bounded function that is integrable for every one of the μ_n and is μ-almost everywhere continuous, then still $\int f\,d\mu = \lim_{n\to\infty} \int f\,d\mu_n$.

---

[37] Sometimes called "strict convergence," "convergence étroite" in French. In the parlance of functional analysts, weak convergence of measures is convergence for the trace of the weak*-topology (!) $\sigma\big(C_b^*(E), C_b(E)\big)$ on $\mathfrak{P}^*(E)$; they reserve the words "weak convergence" for the weak topology $\sigma\big(C_b(E), C_b^*(E)\big)$ on $C_b(E)$. See item A.2.32 on page 381.
[38] Everything said applies to nets and filters as well.
[39] The proof shows that it suffices to check that the μ_n are σ-continuous and that μ is order-continuous on the real part of the algebra generated by $\mathcal{M}$.

Proof. Since $\mu_n(1) = \mu(1) = 1$, we may assume that $1 \in \mathcal{M}$. It is easy to see that the lattice algebra $\mathcal{A}[\mathcal{M}]$ constructed in the proof of proposition A.3.12 on page 399 still generates the topology and that $\mu_n(\phi) \to \mu(\phi)$ for all of its functions ϕ. In other words, we may assume that $\mathcal{M}$ is a lattice algebra.

(i) We know from lemma A.2.19 on page 376 that there is an increasingly directed set $\mathcal{M}^h \subset \mathcal{M}$ whose pointwise supremum is h. If $a < \int h\ d\mu$,[40] there is due to the order-continuity of μ a function $\phi \in \mathcal{M}$ with $\phi \le h$ and $a < \mu(\phi)$. Then $a < \liminf \mu_n(\phi) \le \liminf \int h\ d\mu_n$. Consequently $\int h\ d\mu \le \liminf \int h\ d\mu_n$. Applying this to $-k$ gives the second claim of (i).

(ii) Set $k(x) \stackrel{\text{def}}{=} \limsup_{y \to x} f(y)$ and $h(x) \stackrel{\text{def}}{=} \liminf_{y \to x} f(y)$. Then k is upper semicontinuous and h lower semicontinuous, both bounded. Due to (i),

$$\limsup \int f\, d\mu_n \le \limsup \int k\, d\mu_n$$

as $h = f = k$ μ-a.e.:

$$\le \int k\, d\mu = \int f\, d\mu = \int h\, d\mu$$

$$\le \liminf \int h\, d\mu_n \le \liminf \int f\, d\mu_n \ :$$

equality must hold throughout. *A fortiori*, $\mu_n \Rightarrow \mu$. ∎

For an application consider the case that E is separable and metrizable. Then every σ-continuous measure on $C_b(E)$ is automatically order-continuous (see exercise A.3.10). If $\mu_n \to \mu$ on *uniformly* continuous bounded functions, then $\mu_n \Rightarrow \mu$ and the conclusions (i) and (ii) persist. Proposition A.4.1 not only reduces the need to check $\mu_n(\phi) \to \mu(\phi)$ to fewer functions ϕ, it can also be used to deduce $\mu_n(f) \to \mu(f)$ for more than μ-almost surely continuous functions f once $\mu_n \Rightarrow \mu$ is established:

Corollary A.4.2 *Let E be a completely regular space and (μ_n) a sequence*[38] *of order-continuous probabilities on E that converges weakly to $\mu \in \mathfrak{P}^{\bullet}(E)$. Let F be a subset of E, not necessarily measurable, that has full measure for every μ_n and for μ (i.e., $\int^{\bullet} F\ d\mu = \int^{\bullet} F\ d\mu_n = 1 \quad \forall\, n$). Then $\int f\ d\mu_n \to \int f\ d\mu$ for every bounded function f that is integrable for every μ_n and for μ and whose restriction to F is μ-almost everywhere continuous.*

Proof. Let $\mathcal{E}$ denote the collection of restrictions $\phi_{|F}$ to F of functions ϕ in $C_b(E)$. This is a lattice algebra of bounded functions on F and generates the induced topology. Let us define a positive linear functional $\mu_{|F}$ on $\mathcal{E}$ by

$$\mu_{|F}\big(\phi_{|F}\big) \stackrel{\text{def}}{=} \mu(\phi)\ , \qquad \phi \in C_b(E)\ .$$

$\mu_{|F}$ is well-defined; for if $\phi, \phi' \in C_b(E)$ have the same restriction to F, then $F \subset [\phi = \phi']$, so that the Baire set $[\phi \ne \phi']$ is μ-negligible and consequently

[40] This is of course the μ-integral under the extension discussed on pages 398–400.

$\mu(\phi) = \mu(\phi')$. $\mu_{|F}$ is also order-continuous on $\mathcal{E}$. For if $\Phi \subset \mathcal{E}$ is decreasingly directed with pointwise infimum zero on F, without loss of generality consisting of the restrictions to F of a decreasingly directed family $\Psi \subset C_b(E)$, then $[\inf \Psi = 0]$ is a Borel set of E containing F and thus has μ-negligible complement: $\inf \mu_{|F}(\Phi) = \inf \mu(\Psi) = \mu(\inf \Psi) = 0$. The extension of $\mu_{|F}$ discussed on pages 398–400 integrates all bounded functions of $\mathcal{E}^{\Uparrow}$, among which are the bounded continuous functions of F (lemma A.2.19), and the integral is order-continuous on $C_b(F)$ (equation (A.3.3)). We might as well identify $\mu_{|F}$ with the probability in $\mathfrak{P}^{\bullet}(F)$ so obtained. The order-continuous mean $f \to \| f_{|F} \|^{\bullet}_{\mu_{|F}}$ is the same whether built with $\mathcal{E}$ or $C_b(F)$ as the elementary integrands,[27] agrees with $\| \ \|^{\bullet}_{\mu}$ on $C_{b+}(E)$, and is thus smaller than the latter. From this observation it is easy to see that if $f : E \to \mathbb{R}$ is μ-integrable, then its restriction to F is $\mu_{|F}$-integrable and

$$\int f_{|F} \, d\mu_{|F} = \int f \, d\mu \,. \tag{A.4.1}$$

The same remarks apply to the $\mu_{n|F}$. We are in the situation of proposition A.4.1: $\mu_n \Rightarrow \mu$ clearly implies $\mu_{n|F}\big(\psi_{|F}\big) \to \mu_{|F}\big(\psi_{|F}\big)$ for all $\psi_{|F}$ in the multiplicative class $\mathcal{E}$ that generates the topology, and therefore $\mu_{n|F}(\phi) \to \mu_{|F}(\phi)$ for all bounded functions ϕ on F that are $\mu_{|F}$-almost everywhere continuous.

This translates easily into the claim. Namely, the set of points in F where $f_{|F}$ is discontinuous is by assumption μ-negligible, so by (A.4.1) it is $\mu_{|F}$-negligible: $f_{|F}$ is $\mu_{|F}$-almost everywhere continuous. Therefore

$$\int f \, d\mu_n = \int f_{|F} \, d\mu_{n|F} \to \int f_{|F} \, d\mu_{|F} = \int f \, d\mu \,. \qquad \blacksquare$$

Proposition A.4.1 also yields the Continuity Theorem on $\mathbb{R}^d$ without further ado. Namely, since the complex multiplicative class $\{x \mapsto e^{i\langle x|\alpha\rangle} : \alpha \in \mathbb{R}^d\}$ generates the topology of $\mathbb{R}^d$ (lemma A.3.33), the following is immediate:

Corollary A.4.3 (The Continuity Theorem) *Let μ_n be a sequence of probabilities on $\mathbb{R}^d$ and assume that their characteristic functions $\widehat{\mu}_n$ converge pointwise to the characteristic function $\widehat{\mu}$ of a probability μ. Then $\mu_n \Rightarrow \mu$, and the conclusions (i) and (ii) of proposition A.4.1 continue to hold.*

Theorem A.4.4 (Central Limit Theorem with Lindeberg Criteria) *For $n \in \mathbb{N}$ let $X_n^1, \ldots, X_n^{r_n}$ be independent random variables, defined on probability spaces $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$. Assume that $\mathbb{E}_n[X_n^k = 0]$ and $(\sigma_n^k)^2 \stackrel{\text{def}}{=} \mathbb{E}_n[|X_n^k|^2] < \infty$ for all $n \in \mathbb{N}$ and $k \in \{1, \ldots, r_n\}$, set $S_n = \sum_{k=1}^{r_n} X_n^k$ and $s_n^2 = \mathrm{var}(S_n) = \sum_{k=1}^{r_n} |\sigma_n^k|^2$, and assume the* ***Lindeberg condition***

$$\frac{1}{s_n^2} \sum_{k=1}^{r_n} \int_{|X_n^k| > \epsilon s_n} |X_n^k|^2 \, d\mathbb{P}_n \xrightarrow[n\to\infty]{} 0 \quad \textit{for all } \epsilon > 0 \,. \tag{A.4.2}$$

Then S_n/s_n converges in law to a normalized Gaussian random variable.

Proof. Corollary A.4.3 reduces the problem to showing that the characteristic functions $\widehat{S_n}(\xi)$ of S_n converge to $e^{-\xi^2/2}$ (see A.3.45). This is a standard estimate [10]: replacing X_n^k by X_n^k/s_n we may assume that $s_n = 1$. The inequality[27]

$$\Big|e^{i\xi x} - \big(1 + i\xi x - \xi^2 x^2/2\big)\Big| \le |\xi x|^2 \wedge |\xi x|^3$$

results in the inequality

$$\begin{aligned}\Big|\widehat{X_n^k}(\xi) - \big(1 - \xi^2|\sigma_n^k|^2/2\big)\Big| &\le \mathbb{E}_n\Big[\big|\xi X_n^k\big|^2 \wedge \big|\xi X_n^k\big|^3\Big] \\ &\le \xi^2 \int_{|X_n^k|\ge\epsilon} \big|X_n^k\big|^2 \, d\mathbb{P}_n + \int_{|X_n^k|<\epsilon} \big|\xi X_n^k\big|^3 \, d\mathbb{P}_n \\ &\le \xi^2 \int_{|X_n^k|\ge\epsilon} \big|X_n^k\big|^2 \, d\mathbb{P}_n + \epsilon|\xi|^3|\sigma_n^k|^2 \qquad \forall\, \epsilon > 0\end{aligned}$$

for the characteristic function of X_n^k. Since the $|\sigma_n^k|^2$ sum to 1 and $\epsilon > 0$ is arbitrary, Lindeberg's condition produces

$$\sum_{k=1}^{r_n} \Big|\widehat{X_n^k}(\xi) - \big(1 - \xi^2|\sigma_n^k|^2/2\big)\Big| \xrightarrow[n\to\infty]{} 0 \tag{A.4.3}$$

for any fixed ξ. Now for $\epsilon > 0$, $|\sigma_n^k|^2 \le \epsilon^2 + \int_{|X_n^k|\ge\epsilon} \big|X_n^k\big|^2 \, d\mathbb{P}_n$, so Lindeberg's condition also gives $\max_{k=1}^{r_n} \sigma_n^k \xrightarrow[n\to\infty]{} 0$. Henceforth we fix a ξ and consider only indices n large enough to ensure that $\big|1 - \xi^2|\sigma_n^k|^2/2\big| \le 1$ for $1 \le k \le r_n$. Now if $z_1, \ldots, z_m$ and $w_1, \ldots, w_m$ are complex numbers of absolute value less than or equal to one, then

$$\Big|\prod_{k=1}^m z_k - \prod_{k=1}^m w_k\Big| \le \sum_{k=1}^m \big|z_k - w_k\big| \,, \tag{A.4.4}$$

so (A.4.3) results in

$$\widehat{S_n}(\xi) = \prod_{k=1}^{r_n} \widehat{X_n^k}(\xi) = \prod_{k=1}^{r_n} \big(1 - \xi^2|\sigma_n^k|^2/2\big) + R_n \,, \tag{A.4.5}$$

where $R_n \xrightarrow[n\to\infty]{} 0$: it suffices to show that the product on the right converges to $e^{-\xi^2/2} = \prod_{k=1}^{r_n} e^{-\xi^2|\sigma_n^k|^2/2}$. Now (A.4.4) also implies that

$$\Big|\prod_{k=1}^{r_n} e^{-\xi^2|\sigma_n^k|^2/2} - \prod_{k=1}^{r_n} \big(1 - \xi^2|\sigma_n^k|^2/2\big)\Big| \le \sum_{k=1}^{r_n} \Big|e^{-\xi^2|\sigma_n^k|^2/2} - \big(1 - \xi^2|\sigma_n^k|^2/2\big)\Big| \,.$$

Since $|e^{-x} - (1-x)| \le x^2$ for $x \in \mathbb{R}_+$,[27] the left-hand side above is majorized by

$$\xi^4 \sum_{k=1}^{r_n} |\sigma_n^k|^4 \le \xi^4 \max_{k=1}^{r_n} |\sigma_n^k|^2 \times \sum_{k=1}^{r_n} |\sigma_n^k|^2 \xrightarrow[n\to\infty]{} 0 \,.$$

This in conjunction with (A.4.5) yields the claim. ∎

Uniform Tightness

Unless the underlying completely regular space E is $\mathbb{R}^d$, as in corollary A.4.3, or the topology of E is rather weak, it is hard to find multiplicative classes of bounded functions that define the topology, and proposition A.4.1 loses its utility. There is another criterion for the weak convergence $\mu_n \Rightarrow \mu$, though, one that can be verified in many interesting cases, to wit that the family $\{\mu_n\}$ be uniformly tight and converge on a multiplicative class that separates the points.

Definition A.4.5 *The set $\mathfrak{M}$ of measures in $\mathfrak{M}^{\cdot}(E)$ is* ***uniformly tight*** *if $M \stackrel{\text{def}}{=} \sup\{\,\vert\mu\vert(1) : \mu \in \mathfrak{M}\}$ is finite and if for every $\alpha > 0$ there is a compact subset $K_\alpha \subset E$ such that $\sup\{\,\vert\mu\vert(K_\alpha^c)$*[40] *$: \mu \in \mathfrak{M}\} < \alpha$.*

A set $\mathfrak{P} \subset \mathfrak{P}^{\cdot}(E)$ clearly is uniformly tight if and only if for every $\alpha < 1$ there is a compact set K_α such that $\mu(K_\alpha) \ge 1 - \alpha$ for all $\mu \in \mathfrak{P}$.

Proposition A.4.6 (Prokhoroff) *A uniformly tight collection $\mathfrak{M} \subset \mathfrak{M}^{\cdot}(E)$ is relatively compact in the topology of weak convergence of measures*[37]*; the closure of $\mathfrak{M}$ belongs to $\mathfrak{M}^{\cdot}(E)$ and is uniformly tight as well.*

Proof. The theorem of Alaoglu, a simple consequence of Tychonoff's theorem, shows that the closure of $\mathfrak{M}$ in the topology $\sigma\big(C_b^*(E), C_b(E)\big)$ consists of linear functionals on $C_b(E)$ of total variation less than M. What may not be entirely obvious is that a limit point μ' of $\mathfrak{M}$ is order-continuous. This is rather easy to see, though. Namely, let $\Phi \subset C_b(E)$ be decreasingly directed with pointwise infimum zero. Pick a $\phi_0 \in \Phi$. Given an $\alpha > 0$, find a compact set K_α as in definition A.4.5. Thanks to Dini's theorem A.2.1 there is a $\phi_\alpha \le \phi_0$ in Φ smaller than α on all of K_α. For any $\phi \in \Phi$ with $\phi \le \phi_\alpha$

$$|\mu(\phi)| \le \alpha \vert\mu\vert(K_\alpha) + \int_{K_\alpha^c} \phi_\alpha \, d\vert\mu\vert \le \alpha\big(M + \|\phi_0\|_\infty\big) \qquad \forall\, \mu \in \mathfrak{M}\,.$$

This inequality will also hold for the limit point μ'. That is to say, $\mu'(\Phi) \to 0$: μ' is order-continuous.

If ϕ is any continuous function less than[13] K_α^c, then $|\mu(\phi)| \le \alpha$ for all $\mu \in \mathfrak{M}$ and so $|\mu'(\phi)| \le \alpha$. Taking the supremum over such ϕ gives $\vert\mu'\vert(K_\alpha^c) \le \alpha$: the closure of $\mathfrak{M}$ is "just as uniformly tight as $\mathfrak{M}$ itself." ∎

Corollary A.4.7 *Let (μ_n) be a uniformly tight sequence*[38] *in $\mathfrak{P}^{\cdot}(E)$ and assume that $\mu(\phi) = \lim \mu_n(\phi)$ exists for all ϕ in a complex multiplicative class $\mathcal{M}$ of bounded continuous functions that separates the points. Then (μ_n) converges weakly*[37] *to an order-continuous tight measure that agrees with μ on $\mathcal{M}$. Denoting this limit again by μ we also have the conclusions (i) and (ii) of proposition A.4.1.*

Proof. All limit points of $\{\mu_n\}$ agree on $\mathcal{M}$ and are therefore identical (see proposition A.3.12). ∎

Exercise A.4.8 There exists a partial converse of proposition A.4.6, which is used in section 5.5 below: if E is polish, then a relatively compact subset $\mathfrak{P}$ of $\mathfrak{P}^{\bullet}(E)$ is uniformly tight.

Application: Donsker's Theorem

Recall the normalized random walk

$$Z_t^{(n)} = \frac{1}{\sqrt{n}} \sum_{k \le tn} X_k^{(n)} = \frac{1}{\sqrt{n}} \sum_{k \le [tn]} X_k^{(n)} \,, \qquad t \ge 0 \,,$$

of example 2.5.26. The $X_k^{(n)}$ are independent Bernoulli random variables with $\mathbb{P}^{(n)}[X_k^{(n)} = \pm 1] = 1/2$; they may be living on probability spaces $\big(\Omega^{(n)}, \mathcal{F}^{(n)}, \mathbb{P}^{(n)}\big)$ that vary with n. The Central Limit Theorem easily shows[27] that, for every fixed instant t, $Z_t^{(n)}$ converges in law to a Gaussian random variable with expectation zero and variance t. Donsker's theorem extends this to the whole path: viewed as a random variable with values in the path space $\mathscr{D}$, $Z^{(n)}$ converges in law to a standard Wiener process W. The pertinent topology on $\mathscr{D}$ is the topology of uniform convergence on compacta; it is defined by, and complete under, the metric

$$\rho(z, z') = \sum_{u \in \mathbb{N}} 2^{-u} \wedge \sup_{0 \le s \le u} \big| z(s) - z'(s) \big| \,, \qquad z, z' \in \mathscr{D} \,.$$

What we mean by $Z^{(n)} \Rightarrow W$ is this: for all continuous bounded functions ϕ on $\mathscr{D}$,

$$\mathbb{E}^{(n)}\big[\phi(Z_{\bullet}^{(n)})\big] \xrightarrow[n \to \infty]{} \mathbb{E}\big[\phi(W_{\bullet})\big] \,. \tag{A.4.6}$$

It is necessary to spell this out, since a priori the law $\mathbb{W}$ of a Wiener process is a measure on $\mathscr{C}$, while the $Z^{(n)}$ take values in $\mathscr{D}$ – so how then can the law $\mathbb{Z}^{(n)}$ of $Z^{(n)}$, which lives on $\mathscr{D}$, converge to $\mathbb{W}$? Equation (A.4.6) says how: read Wiener measure as the probability

$$\overline{\mathbb{W}} : \phi \mapsto \int \phi_{|\mathscr{C}} \, d\mathbb{W} \,, \qquad \phi \in C_b(\mathscr{D}) \,,$$

on $\mathscr{D}$. Since the restrictions $\phi_{|\mathscr{C}}$, $\phi \in C_b(\mathscr{D})$, belong to $C_b(\mathscr{C})$, $\overline{\mathbb{W}}$ is actually order-continuous (exercise A.3.10). Now $\mathscr{C}$ is a Borel set in $\mathscr{D}$ (exercise A.2.23) that carries $\overline{\mathbb{W}}$,[27] so we shall henceforth identify $\mathbb{W}$ with $\overline{\mathbb{W}}$ and simply write $\mathbb{W}$ for both.

The left-hand side of (A.4.6) raises a question as well: what is the meaning of $\mathbb{E}^{(n)}\big[\phi(Z^{(n)})\big]$? Observe that $Z^{(n)}$ takes values in the subspace $\mathscr{D}^{(n)} \subset \mathscr{D}$ of paths that are constant on intervals of the form $[k/n, (k+1)/n)$, $k \in \mathbb{N}$, and take values in the discrete set $\mathbb{N}/\sqrt{n}$. One sees as in exercise 1.2.4 that $\mathscr{D}^{(n)}$ is separable and complete under the metric ρ and that the evaluations

$z \mapsto z_t$, $t \geq 0$, generate the Borel σ-algebra of this space. We see as above for Wiener measure that

$$\mathbb{Z}^{(n)} : \phi \mapsto \mathbb{E}^{(n)}\big[\phi(Z_{\bullet}^{(n)})\big] , \qquad \phi \in C_b(\mathscr{D}) ,$$

defines an order-continuous probability on $\mathscr{D}$. It makes sense to state the

Theorem A.4.9 (Donsker) $\mathbb{Z}^{(n)} \Rightarrow \mathbb{W}$. *In other words, the $Z^{(n)}$ converge in law to a standard Wiener process.*

We want to show this using corollary A.4.7, so there are two things to prove: **1)** the laws $\mathbb{Z}^{(n)}$ form a uniformly tight family of probabilities on $C_b(\mathscr{D})$, and **2)** there is a multiplicative class $\mathcal{M} = \overline{\mathcal{M}} \subset C_b(\mathscr{D};\mathbb{C})$ separating the points so that

$$\mathbb{E}^{\mathbb{Z}^{(n)}}[\phi] \xrightarrow[n\to\infty]{} \mathbb{E}^{\mathbb{W}}[\phi] \qquad \forall\, \phi \in \mathcal{M} .$$

We start with point 2). Let Γ denote the vector space of all functions $\gamma : [0,\infty) \to \mathbb{R}$ of compact support that have a continuous derivative $\dot\gamma$. We view $\gamma \in \Gamma$ as the cumulative distribution function of the measure $d\gamma_t = \dot\gamma_t dt$ and also as the functional $z_{\bullet} \mapsto \langle z_{\bullet}|\gamma\rangle \stackrel{\text{def}}{=} \int_0^\infty z_t \; d\gamma_t$. We set $\mathcal{M} = e^{i\Gamma} \stackrel{\text{def}}{=} \{e^{i\langle\cdot|\gamma\rangle} : \gamma \in \Gamma\}$ as on page 410. Clearly $\mathcal{M}$ is a multiplicative class closed under complex conjugation and separating the points; for if $e^{i\int_0^\infty z_t \; d\gamma_t} = e^{i\int_0^\infty z'_t \; d\gamma_t}$ for all $\gamma \in \Gamma$, then the two right-continuous paths $z, z' \in \mathscr{D}$ must coincide.

Lemma A.4.10 $\mathbb{E}^{(n)}\Big[e^{i\int_0^\infty Z_t^{(n)} \; d\gamma_t}\Big] \xrightarrow[n\to\infty]{} e^{-\frac{1}{2}\int_0^\infty \gamma_t^2 \, dt}$.

Proof. Repeated applications of l'Hospital's rule show that $(\tan x - x)/x^3$ has a finite limit as $x \to 0$, so that $\tan x = x + O(x^3)$ at $x = 0$. Integration gives $\ln\cos x = -x^2/2 + O(x^4)$. Since γ is continuous and bounded and vanishes after some finite instant, therefore,

$$\sum_{k=1}^{\infty} \ln\cos\Big(\frac{\gamma_{k/n}}{\sqrt{n}}\Big) = -\frac{1}{2}\sum_{k=1}^{\infty}\frac{\gamma_{k/n}^2}{n} + O(1/n) \xrightarrow[n\to\infty]{} -\frac{1}{2}\int_0^\infty \gamma_t^2 \; dt ,$$

and so
$$\prod_{k=1}^{\infty} \cos\Big(\frac{\gamma_{k/n}}{\sqrt{n}}\Big) \xrightarrow[n\to\infty]{} e^{-\frac{1}{2}\int_0^\infty \gamma_t^2 \, dt} . \qquad (*)$$

Now
$$\int_0^\infty Z_t^{(n)} \; d\gamma_t = -\int_0^\infty \gamma_t \; dZ_t^{(n)} = -\sum_{k=1}^{\infty} \frac{\gamma_{k/n}}{\sqrt{n}} \cdot X_k^{(n)} ,$$

and so
$$\mathbb{E}^{(n)}\Big[e^{i\int_0^\infty Z_t^{(n)} \; d\gamma_t}\Big] = \mathbb{E}^{(n)}\Big[e^{-i\sum_{k=1}^\infty \frac{\gamma_{k/n}}{\sqrt{n}}\cdot X_k^{(n)}}\Big]$$
$$= \prod_{k=1}^{\infty} \cos\Big(\frac{\gamma_{k/n}}{\sqrt{n}}\Big) \xrightarrow[n\to\infty]{} e^{-\frac{1}{2}\int_0^\infty \gamma_t^2 \, dt} .$$

∎

Now to point 1), the tightness of the $\mathbb{Z}^{(n)}$. To start with we need a criterion for compactness in $\mathscr{D}$. There is an easy generalization of the Ascoli–Arzela theorem A.2.38:

Lemma A.4.11 *A subset $\mathcal{K} \subset \mathscr{D}$ is relatively compact if and only if the following two conditions are satisfied:*

(a) For every $u \in \mathbb{N}$ there is a constant M^u such that $|z_t| \le M^u$ for all $t \in [0,u]$ and all $z \in \mathcal{K}$.

(b) For every $u \in \mathbb{N}$ and every $\epsilon > 0$ there exists a finite collection $\mathcal{T}^{u,\epsilon} = \{0 = t_0^{u,\epsilon} < t_1^{u,\epsilon} < \ldots < t_{N(\epsilon)}^{u,\epsilon} = u\}$ of instants such that for all $z \in \mathcal{K}$

$$\sup\{|z_s - z_t| : s,t \in [t_{n-1}^{u,\epsilon}, t_n^{u,\epsilon})\} \le \epsilon \,, \qquad 1 \le n \le N(\epsilon) \,. \tag{$*$}$$

Proof. We shall need, and therefore prove, only the sufficiency of these two conditions. Assume then that they are satisfied and let $\mathfrak{F}$ be a filter on $\mathcal{K}$. Tychonoff's theorem A.2.13 in conjunction with (a) provides a refinement $\mathfrak{F}'$ that converges pointwise to some path z. Clearly z is again bounded by M^u on $[0,u]$ and satisfies $(*)$. A path $z' \in \mathcal{K}$ that differs from z in the points $t_n^{u,\epsilon}$ by less than ϵ is uniformly as close as 3ϵ to z on $[0,u]$. Indeed, for $t \in [t_{n-1}^{u,\epsilon}, t_n^{u,\epsilon})$

$$\left|z_t - z'_t\right| \le \left|z_t - z_{t_{n-1}^{u,\epsilon}}\right| + \left|z_{t_{n-1}^{u,\epsilon}} - z'_{t_{n-1}^{u,\epsilon}}\right| + \left|z'_{t_{n-1}^{u,\epsilon}} - z'_t\right| < 3\epsilon \,.$$

That is to say, the refinement $\mathfrak{F}'$ converges uniformly $[0,u]$ to z. This holds for all $u \in \mathbb{N}$, so $\mathfrak{F}' \to z \in \mathscr{D}$ uniformly on compacta. ∎

We use this to prove the tightness of the $\mathbb{Z}^{(n)}$:

Lemma A.4.12 *For every $\alpha > 0$ there exists a compact set $\mathcal{K}_\alpha \subset \mathscr{D}$ with the following property: for every $n \in \mathbb{N}$ there is a set $\Omega_\alpha^{(n)} \in \mathcal{F}^{(n)}$ such that*

$$\mathbb{P}^{(n)}\big[\Omega_\alpha^{(n)}\big] > 1 - \alpha \quad \textit{and} \quad Z^{(n)}\big(\Omega_\alpha^{(n)}\big) \subset \mathcal{K}_\alpha \,.$$

Consequently the laws $\mathbb{Z}^{(n)}$ form a uniformly tight family.

Proof. For $u \in \mathbb{N}$, let

$$M_\alpha^u \stackrel{\text{def}}{=} \sqrt{u}\, 2^{u+1}/\alpha$$

and set

$$\Omega_{\alpha,1}^{(n)} \stackrel{\text{def}}{=} \bigcap_{u \in \mathbb{N}} \big[\big|Z^{(n)}\big|_u^* \le M_\alpha^u\big] \,.$$

Now $Z^{(n)}$ is a martingale that at the instant u has square expectation u, so Doob's maximal lemma 2.5.18 and a summation give

$$\mathbb{P}^{(n)}\Big[\big|Z^{(n)}\big|_u^* > M_\alpha^u\Big] < \sqrt{u}/M_\alpha^u \quad \text{and} \quad \mathbb{P}\big[\Omega_{\alpha,1}^{(n)}\big] > 1 - \alpha/2 \,.$$

For $\omega \in \Omega_{\alpha,1}^{(n)}$, $Z_{\bullet}^{(n)}(\omega)$ is bounded by M_α^u on $[0,u]$.

The construction of a large set $\Omega^{(n)}_{\alpha,2}$ on which the paths of $Z^{(n)}$ satisfy (b) of lemma A.4.11 is slightly more complicated. Let $0 \le s \le \tau \le t$. $Z^{(n)}_\tau - Z^{(n)}_s = \sum_{[sn]<k\le[\tau n]} X^{(n)}_k$ has the same distribution as $M^{(n)}_\tau \stackrel{\text{def}}{=} \sum_{k\le[\tau n]-[sn]} X^{(n)}_k$. Now $M^{(n)}_t$ has fourth moment

$$\mathbb{E}\Big[\big|M^{(n)}_t\big|^4\Big] = 3\big([tn]-[sn]\big)^2/n^2 \le 3\big(t-s+1/n\big)^2 .$$

Since $M^{(n)}$ is a martingale, Doob's maximal theorem 2.5.19 gives

$$\mathbb{E}\Big[\sup_{s\le\tau\le t}\big|Z^{(n)}_\tau - Z^{(n)}_s\big|^4\Big] \le \left(\frac{4}{3}\right)^4 \cdot 3(t-s+1/n)^2 < 10(t-s+1/n)^2 \qquad (*)$$

for all $n \in \mathbb{N}$. Since $\sum N2^{-N/2} < \infty$, there is an index N_α such that

$$40 \sum_{N\ge N_\alpha} N2^{-N/2} < \alpha/2 .$$

If $n \le 2^{N_\alpha}$, we set $\Omega^{(n)}_{\alpha,2} = \Omega^{(n)}$. For $n > 2^{N_\alpha}$ let $\mathcal{N}^{(n)}$ denote the set of integers $N \ge N_\alpha$ with $2^N \le n$. For every one of them $(*)$ and Chebyscheff's inequality produce

$$\mathbb{P}\Big[\sup_{k2^{-N}\le\tau\le(k+1)2^{-N}}\big|Z^{(n)}_\tau - Z^{(n)}_{k2^{-N}}\big| > 2^{-N/8}\Big]$$
$$\le 2^{N/2}\cdot 10\big(2^{-N}+1/n\big)^2 \le 40\cdot 2^{-3N/2} , \qquad k = 0,1,2,\ldots .$$

Hence $$\bigcup_{N\in\mathcal{N}^{(n)}}\ \bigcup_{0\le k<N2^N}\Big[\sup_{k2^{-N}\le\tau\le(k+1)2^{-N}}\big|Z^{(n)}_\tau - Z^{(n)}_{k2^{-N}}\big| > 2^{-N/8}\Big]$$

has measure less than $\alpha/2$. We let $\Omega^{(n)}_{\alpha,2}$ denote its complement and set

$$\Omega^{(n)}_\alpha = \Omega^{(n)}_{\alpha,1} \cap \Omega^{(n)}_{\alpha,2} .$$

This is a set of $\mathbb{P}^{(n)}$-measure greater than $1-\alpha$.

For $N \in \mathbb{N}$, let $\mathcal{T}^N$ be the set of instants that are of the form k/l, $k \in \mathbb{N}, l \le 2^{N_\alpha}$, or of the form $k2^{-N}$, $k \in \mathbb{N}$. For the set $\mathcal{K}_\alpha$ we take the collection of paths z that satisfy the following description: for every $u \in \mathbb{N}$ z is bounded on $[0,u]$ by M^u_α and varies by less than $2^{-N/8}$ on any interval $[s,t)$ whose endpoints s,t are consecutive points of $\mathcal{T}^N$. Since $\mathcal{T}^N \cap [0,u]$ is finite, $\mathcal{K}_\alpha$ is compact (lemma A.4.11).

It is left to be shown that $Z^{(n)}_.(\omega) \in \mathcal{K}_\alpha$ for $\omega \in \Omega^{(n)}_\alpha$. This is easy when $n \le 2^{N_\alpha}$: the path $Z^{(n)}_.(\omega)$ is actually constant on $[s,t)$, whatever $\omega \in \Omega^{(n)}$. If $n > 2^{N_\alpha}$ and s,t are consecutive points in $\mathcal{T}^N$, then $[s,t)$ lies in an interval of the form $\big[k2^{-N},(k+1)2^{-N}\big)$, $N \in \mathcal{N}^{(n)}$, and $Z^{(n)}_.(\omega)$ varies by less than $2^{-N/8}$ on $[s,t)$ as long as $\omega \in \Omega^{(n)}_\alpha$.

Thanks to equation (A.3.7),

$$\mathbb{Z}^{(n)}(\mathcal{K}_\alpha) = \mathbb{E}\Big[\mathcal{K}_\alpha \circ Z^{(n)}\Big] \geq \mathbb{P}^{(n)}\big[\Omega^{(n)}\big] \geq 1-\alpha\ , \qquad n \in \mathbb{N}\ .$$

The family $\{\mathbb{Z}^{(n)} : n \in \mathbb{N}\}$ is thus uniformly tight. ∎

Proof of Theorem A.4.9. Lemmas A.4.10 and A.4.12 in conjunction with criterion A.4.7 allow us to conclude that $(\mathbb{Z}^{(n)})$ converges weakly to an order-continuous tight (proposition A.4.6) probability $\mathbb{Z}$ on $\mathscr{D}$ whose characteristic function $\widehat{\mathbb{Z}}^{\Gamma}$ is that of Wiener measure (corollary 3.9.5). By proposition A.3.12, $\mathbb{Z} = \mathbb{W}$. ∎

Example A.4.13 Let $\delta_1, \delta_2, \ldots$ be strictly positive numbers. On $\mathscr{D}$ define by induction the functions $\tau_0 = \tau_0^+ = 0$,

$$\tau_{k+1}(z) = \inf\Big\{t : \big|z_t - z_{\tau_k(z)}\big| \geq \delta_{k+1}\Big\}$$

and

$$\tau_{k+1}^+(z) = \inf\Big\{t : \Big|z_t - z_{\tau_k^+(z)}\Big| > \delta_{k+1}\Big\}\ .$$

Let $\Phi = \Phi(t_1, \zeta_1, t_2, \zeta_2, \ldots)$ be a bounded continuous function on $\mathbb{R}^{\mathbb{N}}$ and set

$$\phi(z) = \Phi\big(\tau_1(z), z_{\tau_1(z)}, \tau_2(z), z_{\tau_2(z)}, \ldots\big)\ , \qquad z \in \mathscr{D}\ .$$

Then

$$\mathbb{E}^{\mathbb{Z}^{(n)}}[\phi] \xrightarrow[n\to\infty]{} \mathbb{E}^{\mathbb{W}}[\phi]\ .$$

Proof. The processes $Z^{(n)}, W$ take their values in the Borel[41] subset

$$\mathscr{D}^{\simeq} \stackrel{\text{def}}{=} \bigcup_n \mathscr{D}^{(n)} \cup \mathscr{C}$$

of $\mathscr{D}$. $\mathscr{D}^{\simeq}$ therefore[27] has full measure for their laws $\mathbb{Z}^{(n)}, \mathbb{W}$. Henceforth we consider τ_k, τ_k^+ as functions on this set. At a point $z^0 \in \mathscr{D}^{(n)}$ the functions τ_k, τ_k^+, $z \mapsto z_{\tau_k(z)}$, and $z \mapsto z_{\tau_k^+(z)}$ are continuous. Indeed, pick an instant of the form p/n, where p, n are relatively prime. A path $z \in \mathscr{D}^{\simeq}$ closer than $1/(6\sqrt{n})$ to z^0 uniformly on $[0, 2p/n]$ must jump at p/n by at least $2/(3\sqrt{n})$, and no path in $\mathscr{D}^{\simeq}$ other than z^0 itself does that. In other words, every point of $\bigcup_n \mathscr{D}^{(n)}$ is an isolated point in $\mathscr{D}^{\simeq}$, so that every function is continuous at it: we have to worry about the continuity of the functions above only at points $w \in \mathscr{C}$. Several steps are required.

a) If $z \to z_{\tau_k(z)}$ is continuous on $E_k \stackrel{\text{def}}{=} \mathscr{C} \cap \big[\tau_k = \tau_k^+ < \infty\big]$, then (a1) τ_{k+1} is lower semicontinuous and (a2) τ_{k+1}^+ is upper semicontinuous on this set.

To see (a1) let $w \in E_k$, set $s = \tau_k(w)$, and pick $t < \tau_{k+1}(w)$. Then $\alpha \stackrel{\text{def}}{=} \delta_{k+1} - \sup_{s\leq\sigma\leq t}|w_\sigma - w_s| > 0$. If $z \in \mathscr{D}^{\simeq}$ is so close to w uniformly

[41] See exercise A.2.23 on page 379.

on a suitable interval containing $[0,t+1]$ that $|w_s - z_{\tau_k(z)}| < \alpha/2$, and is uniformly as close as $\alpha/2$ to w there, then

$$\begin{aligned}\left|z_\sigma - z_{\tau_k(z)}\right| &\le |z_\sigma - w_\sigma| + |w_\sigma - w_s| + \left|w_s - z_{\tau_k(z)}\right| \\ &< \alpha/2 + (\delta_{k+1} - \alpha) + \alpha/2 = \delta_{k+1}\end{aligned}$$

for all $\sigma \in [s,t]$ and consequently $\tau_{k+1}(z) > t$. Therefore we have as desired $\liminf_{z\to w} \tau_{k+1}(z) \ge \tau_{k+1}(w)$.

To see (a2) consider a point $w \in \left[\tau^+_{k+1} < u\right] \cap E_k$. Set $s = \tau^+_k(w)$. There is an instant $t \in (s,u)$ at which $\alpha \stackrel{\text{def}}{=} |w_t - w_s| - \delta_{k+1} > 0$. If $z \in \mathscr{D}^{\simeq}$ is sufficiently close to w uniformly on some interval containing $[0,u]$, then $|z_t - w_t| < \alpha/2$ and $|w_s - z_{\tau_k(z)}| < \alpha/2$ and therefore

$$\begin{aligned}\left|z_t - z_{\tau_k(z)}\right| &\ge -|z_t - w_t| + |w_t - w_s| - \left|w_s - z_{\tau_k(z)}\right| \\ &> -\alpha/2 + (\delta_{k+1} + \alpha) - \alpha/2 = \delta_{k+1}\ .\end{aligned}$$

That is to say, $\tau^+_{k+1} < u$ in a whole neighborhood of w in $\mathscr{D}^{\simeq}$, wherefore as desired $\limsup_{z\to w} \tau^+_{k+1}(z) \le \tau^+_{k+1}(w)$.

b) $z \to z_{\tau_k(z)}$ is continuous on E_k for all $k \in \mathbb{N}$. This is trivially true for $k = 0$. Asssume it for k. By a) τ_{k+1} and τ^+_{k+1}, which on E_{k+1} agree and are finite, are continuous there. Then so is $z \to z_{\tau_k(z)}$.

c) $\mathbb{W}[E_k] = 1$, for $k = 1, 2, \ldots$. This is plain for $k = 0$. Assuming it for $1, \ldots, k$, set $E^k = \bigcap_{\kappa\le k} E_\kappa$. This is then a Borel subset of $\mathscr{C} \subset \mathscr{D}$ of Wiener measure 1 on which plainly $\tau_{k+1} \le \tau^+_{k+1}$. Let $\delta = \delta_1 + \cdots + \delta_{k+1}$. The stopping time τ^+_{k+1} occurs before $T \stackrel{\text{def}}{=} \inf\left\{t : |w_t| > \delta\right\}$, which is integrable (exercise 4.2.21). The continuity of the paths results in

$$\delta^2_{k+1} = \left(w_{\tau_{k+1}} - w_{\tau_k}\right)^2 = \left(w_{\tau^+_{k+1}} - w_{\tau^+_k}\right)^2 .$$

Thus $$\begin{aligned}\delta^2_{k+1} &= \mathbb{E}^{\mathbb{W}}\Big[\left(w_{\tau_{k+1}} - w_{\tau_k}\right)^2\Big] = \mathbb{E}^{\mathbb{W}}\Big[w^2_{\tau_{k+1}} - w^2_{\tau_k}\Big] \\ &= \mathbb{E}^{\mathbb{W}}\Big[2\int_{\tau_k}^{\tau_{k+1}} w_s\, dw_s + \left(\tau_{k+1} - \tau_k\right)\Big] = \mathbb{E}^{\mathbb{W}}[\tau_{k+1} - \tau_k]\ .\end{aligned}$$

The same calculation can be made for τ^+_{k+1}, so that $\mathbb{E}^{\mathbb{W}}[\tau^+_{k+1} - \tau_{k+1}] = 0$ and consequently $\tau^+_{k+1} = \tau_{k+1}$ $\mathbb{W}$-almost surely on E^k: we have $\mathbb{W}[E_{k+1}] = 1$, as desired.

Let $E = \bigcup_n \mathscr{D}^{(n)} \cup \bigcap_k E^k$. This is a Borel subset of $\mathscr{D}$ with $\mathbb{W}[E] = \mathbb{Z}^{(n)}[E] = 1 \quad \forall\, n$. The restriction of ϕ to it is continuous. Corollary A.4.2 applies and gives the claim. ∎

Exercise A.4.14 Assume the coupling coefficient f of the markovian SDE (5.6.4), which reads

$$X_t = x + \int_0^t f(X^x_{s-})\, dZ_s\ , \tag{A.4.7}$$

is a bounded Lipschitz vector field. As Z runs through the sequence $Z^{(n)}$ the solutions $X^{(n)}$ converge in law to the solution of (A.4.7) driven by Wiener process.

A.5 Analytic Sets and Capacity

The *preimages* under continuous maps of open sets are open; the *preimages* under measurable maps of measurable sets are measurable. Nothing can be said in general about *direct* or *forward* images, with one exception: the continuous image of a compact set is compact (exercise A.2.14). (Even Lebesgue himself made a mistake here, thinking that the projection of a Borel set would be Borel.) This dearth is alleviated slightly by the following abbreviated theory of analytic sets, initiated by Lusin and his pupil Suslin. The presentation follows [20] – see also [17]. The class of *analytic sets* is designed to be invariant under direct images of certain simple maps, projections. Their theory implicitly uses the fact that continuous direct images of compact sets are compact.

Let F be a set. Any collection $\mathcal{F}$ of subsets of F is called a ***paving*** of F, and the pair $(F, \mathcal{F})$ is a ***paved set***. $\mathcal{F}_\sigma$ denotes the collection of subsets of F that can be written as countable unions of sets in $\mathcal{F}$, and $\mathcal{F}_\delta$ denotes the collection of subsets of F that can be written as countable intersections of members of $\mathcal{F}$. Accordingly $\mathcal{F}_{\sigma\delta}$ is the collection of sets that are countable intersections of sets each of which is a countable union of sets in $\mathcal{F}$, etc. If $(K, \mathcal{K})$ is another paved set, then the ***product paving*** $\mathcal{K} \times \mathcal{F}$ consists of the "rectangles" $A \times B$, $A \in \mathcal{K}, B \in \mathcal{F}$. The family $\mathcal{K}$ of subsets of K constitutes a ***compact paving*** if it has the *finite intersection property*: whenever a subfamily $\mathcal{K}' \subset \mathcal{K}$ has void intersection there exists a finite subfamily $\mathcal{K}'_0 \subset \mathcal{K}'$ that already has void intersection. We also say that K is ***compactly paved*** by $\mathcal{K}$.

Definition A.5.1 (Analytic Sets) *Let $(F, \mathcal{F})$ be a paved set. A set $A \subset F$ is called $\mathcal{F}$**-analytic** if there exist an auxiliary set K equipped with a compact paving $\mathcal{K}$ and a set $B \in (\mathcal{K} \times \mathcal{F})_{\sigma\delta}$ such that A is the projection of B on F:*

$$A = \pi_F(B) .$$

Here $\pi_F = \pi_F^{K \times F}$ is the natural projection of $K \times F$ onto its second factor F – see figure A.16. The collection of $\mathcal{F}$-analytic sets is denoted by $\mathcal{A}[\mathcal{F}]$.

Theorem A.5.2 *The sets of $\mathcal{F}$ are $\mathcal{F}$-analytic. The intersection and the union of countably many $\mathcal{F}$-analytic sets are $\mathcal{F}$-analytic.*

Proof. The first statement is obvious. For the second, let $\{A_n : n = 1, 2 \ldots\}$ be a countable collection of $\mathcal{F}$-analytic sets. There are auxiliary spaces K_n equipped with compact pavings $\mathcal{K}_n$ and $(\mathcal{K}_n \times \mathcal{F})_{\sigma\delta}$-sets $B_n \subset K_n \times F$ whose projection onto F is A_n. Each B_n is the countable intersection of sets $B_n^j \in (\mathcal{K}_n \times \mathcal{F})_\sigma$.

To see that $\bigcap A_n$ is $\mathcal{F}$-analytic, consider the product $K = \prod_{n=1}^\infty K_n$. Its paving $\mathcal{K}$ is the product paving, consisting of sets $C = \prod_{n=1}^\infty C_n$, where

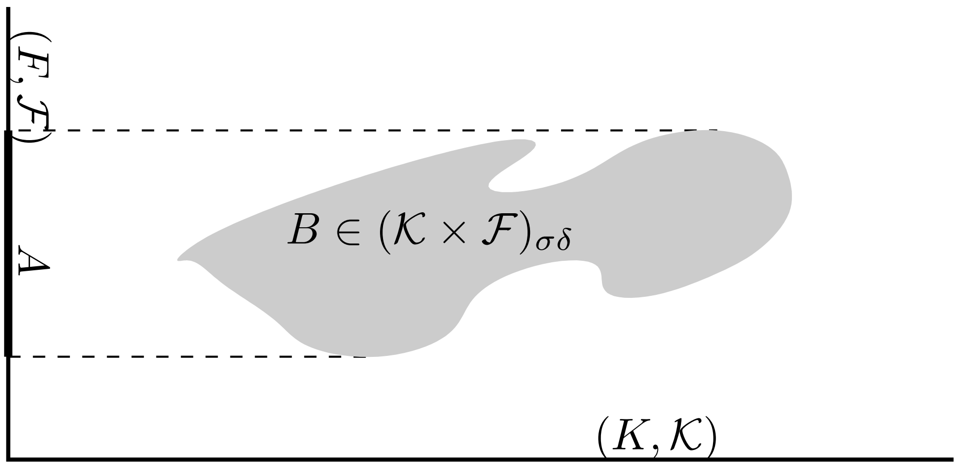


Figure A.16 An $\mathcal{F}$-analytic set A

$C_n = K_n$ for all but finitely many indices n and $C_n \in \mathcal{K}_n$ for the finitely many exceptions. $\mathcal{K}$ is compact. For if $\{C^\alpha\} = \{\prod_n C_n^\alpha : \alpha \in A\} \subset \mathcal{K}$ has void intersection, then one of the collections $\{C_n^\alpha : \alpha\}$, say C_1^α, must have void intersection, otherwise $\prod_n \bigcap_\alpha C_n^\alpha \neq \emptyset$ would be contained in $\bigcap_\alpha C^\alpha$. There are then $\alpha_1, \ldots, \alpha_k$ with $\bigcap_i C_1^{\alpha_i} = \emptyset$, and thus $\bigcap_i C^{\alpha_i} = \emptyset$. Let

$$B'_n = \prod_{m \neq n} K_m \times B_n = \prod_{m \neq n} K_m \times \bigcap_{j=1}^{\infty} B_n^j \subset F \times K .$$

Clearly $B = \bigcap B'_n$ belongs to $(\mathcal{K} \times \mathcal{F})_{\sigma\delta}$ and has projection $\bigcap A_n$ onto F. Thus $\bigcap A_n$ is $\mathcal{F}$-analytic.

For the union consider instead the disjoint union $K = \biguplus_n K_n$ of the K_n. For its paving $\mathcal{K}$ we take the direct sum of the $\mathcal{K}_n$: $C \subset K$ belongs to $\mathcal{K}$ if and only if $C \cap K_n$ is void for all but finitely many indices n and a member of $\mathcal{K}_n$ for the exceptions. $\mathcal{K}$ is clearly compact. The set $B \stackrel{\text{def}}{=} \biguplus_n B_n$ equals $\bigcap_{j=1}^{\infty} \biguplus_n B_n^j$ and has projection $\bigcup A_n$. Thus $\bigcup A_n$ is $\mathcal{F}$-analytic. ∎

Corollary A.5.3 *$\mathcal{A}[\mathcal{F}]$ contains the σ-algebra generated by $\mathcal{F}$ if and only if the complement of every set in $\mathcal{F}$ is $\mathcal{F}$-analytic. In particular, if the complement of every set in $\mathcal{F}$ is the countable union of sets in $\mathcal{F}$, then $\mathcal{A}[\mathcal{F}]$ contains the σ-algebra generated by $\mathcal{F}$.*

Proof. Under the hypotheses the collection of sets $A \subset F$ such that both A and its complement A^c are $\mathcal{F}$-analytic contains $\mathcal{F}$ and is a σ-algebra. ∎

The *direct* or *forward* image of an analytic set is analytic, under certain projections. The precise statement is this:

Proposition A.5.4 *Let $(K, \mathcal{K})$ and $(F, \mathcal{F})$ be paved sets, with $\mathcal{K}$ compact. The projection of a $\mathcal{K} \times \mathcal{F}$-analytic subset B of $K \times F$ onto F is $\mathcal{F}$-analytic.*

Proof. There exist an auxiliary compactly paved space $(K', \mathcal{K}')$ and a set $C \in \big(\mathcal{K}' \times (\mathcal{K} \times \mathcal{F})\big)_{\sigma\delta}$ whose projection on $K \times F$ is B. Set $K'' = K' \times K$ and let $\mathcal{K}''$ be its product paving, which is compact. Clearly C belongs to $(\mathcal{K}'' \times \mathcal{F})_{\sigma\delta}$, and $\pi_F^{K'' \times F}(C) = \pi_F^{K' \times F}(B)$. This last set is therefore $\mathcal{F}$-analytic. ▮

Exercise A.5.5 Let $(F, \mathcal{F})$ and $(G, \mathcal{G})$ be paved sets. Then $\mathcal{A}[\mathcal{A}[\mathcal{F}]] = \mathcal{A}[\mathcal{F}]$ and $\mathcal{A}[\mathcal{F}] \times \mathcal{A}[\mathcal{G}] \subset \mathcal{A}[\mathcal{F} \times \mathcal{G}]$. If $f : G \to F$ has $f^{-1}(\mathcal{F}) \subset \mathcal{G}$, then $f^{-1}(\mathcal{A}[\mathcal{F}]) \subset \mathcal{A}[\mathcal{G}]$.

Exercise A.5.6 Let $(K, \mathcal{K})$ be a compactly paved set. (i) The intersections of arbitrary subfamilies of $\mathcal{K}$ form a compact paving $\mathcal{K}^{\cap_a}$. (ii) The collection $\mathcal{K}^{\cup_f}$ of all unions of finite subfamilies of $\mathcal{K}$ is a compact paving. (iii) There is a compact topology on K (possibly far from being Hausdorff) such that $\mathcal{K}$ is a collection of compact sets.

Definition A.5.7 (Capacities and Capacitability) *Let $(F, \mathcal{F})$ be a paved set. (i) An $\mathcal{F}$-**capacity** is a positive numerical set function I that is defined on all subsets of F and is increasing: $A \subseteq B \implies I(A) \le I(B)$; is continuous along arbitrary increasing sequences: $F \supset A_n \uparrow A \implies I(A_n) \uparrow I(A)$; and is continuous along decreasing sequences of $\mathcal{F}$: $\mathcal{F} \ni F_n \downarrow F \implies I(F_n) \downarrow I(F)$. (ii) A subset C of F is called $(\mathcal{F}, I)$-capacitable, or **capacitable** for short, if*

$$I(C) = \sup\{I(K) : K \subset C\,,\ K \in \mathcal{F}_\delta\}\ .$$

The point of the compactness that is required of the auxiliary paving is the following consequence:

Lemma A.5.8 *Let $(K, \mathcal{K})$ and $(F, \mathcal{F})$ be paved sets, with $\mathcal{K}$ compact and $\mathcal{F}$ closed under finite unions. Denote by $\mathcal{K} \otimes \mathcal{F}$ the paving $(\mathcal{K} \times \mathcal{F})^{\cup_f}$ of finite unions of rectangles from $\mathcal{K} \times \mathcal{F}$. (i) For any decreasing sequence (C_n) in $\mathcal{K} \otimes \mathcal{F}$, $\pi_F\big(\bigcap_n C_n\big) = \bigcap_n \pi_F(C_n)$.*
(ii) If I is an $\mathcal{F}$-capacity, then

$$I \circ \pi_F : A \mapsto I\big(\pi_F(A)\big)\,, \qquad A \subset K \times F\,,$$

is a $\mathcal{K} \otimes \mathcal{F}$-capacity.

Proof. (i) Let $x \in \bigcap_n \pi_F(C_n)$. The sets $K_n^x \stackrel{\text{def}}{=} \{k \in K : (k, x) \in C_n\}$ belong to $\mathcal{K}^{\cup_f}$, are non-void, and decreasing in n. Exercise A.5.6 furnishes a point k in their intersection, and clearly (k, x) is a point in $\bigcap_n C_n$ whose projection on F is x. Thus $\bigcap_n \pi_F(C_n) \subset \pi_F\big(\bigcap_n C_n\big)$. The reverse inequality is obvious.

Here is a direct proof that avoids ultrafilters. Let x be a point in $\bigcap_n \pi_F(C_n)$ and let us show that it belongs to $\pi_F\big(\bigcap_n C_n\big)$. Now the sets $K_n^x = \{k \in K : (k, x) \in C_n\}$ are not void. K_n^x is a finite union of sets in $\mathcal{K}$, say $K_n^x = \bigcup_{i=1}^{I(n)} K_{n,i}^x$. For at least one index $i = n(1)$, $K_{1,n(1)}^x$ must intersect all of the subsequent sets K_n^x, $n > 1$, in a non-void set. Replacing K_n^x by $K_{1,n(1)}^x \cap K_n^x$ for $n = 1, 2 \ldots$ reduces the situation to $K_1^x \in \mathcal{K}^{\cap_f}$. For at least one index $i = n(2)$, $K_{2,n(2)}^x$ must intersect all of the subsequent sets K_n^x, $n > 2$, in a non-void set. Replacing K_n^x by $K_{2,n(2)}^x \cap K_n^x$ for $n = 2, 3 \ldots$

reduces the situation to $K_2^x \in \mathcal{K}^{\cap f}$. Continue on. The K_n^x so obtained belong to $\mathcal{K}^{\cap f}$, still decrease with n, and are non-void. There is thus a point $k \in \bigcap_n K_n^x$. The point $(k,x) \in \bigcap_n C_n$ evidently has $\pi_F(k,x) = x$, as desired.

(ii) First, it is evident that $I \circ \pi_F$ is increasing and continuous along arbitrary sequences; indeed, $K \times F \supset A_n \uparrow A$ implies $\pi_F(A_n) \uparrow \pi_F(A)$, whence $I \circ \pi_F(A_n) \uparrow I \circ \pi_F(A_n)$. Next, if C_n is a decreasing sequence in $\mathcal{K} \otimes \mathcal{F}$, then $\pi_F(C_n)$ is a decreasing sequence of $\mathcal{F}$, and by (i) $I \circ \pi_F(C_n)$ $= I\big(\pi_F(C_n)\big)$ decreases to $I\big(\bigcap_n \pi_F(C_n)\big) = I \circ \pi_F\big(\bigcap_n C_n\big)$: the continuity along decreasing sequences of $\mathcal{K} \otimes \mathcal{F}$ is established as well. ∎

Theorem A.5.9 (Choquet's Capacitability Theorem) *Let $\mathcal{F}$ be a paving that is closed under finite unions and finite intersections, and let I be an $\mathcal{F}$-capacity. Then every $\mathcal{F}$-analytic set A is $(\mathcal{F}, I)$-capacitable.*

Proof. To start with, let $A \in \mathcal{F}_{\sigma\delta}$. There is a sequence of sets $F_n^\sigma \in \mathcal{F}_\sigma$ whose intersection is A. Every one of the F_n^σ is the union of a countable family $\{F_n^j : j \in \mathbb{N}\} \subset \mathcal{F}$. Since $\mathcal{F}$ is closed under finite unions, we may replace F_n^j by $\bigcup_{i=1}^j F_n^i$ and thus assume that F_n^j increases with j: $F_n^j \uparrow F_n^\sigma$. Suppose $I(A) > r$. We shall construct by induction a sequence (F_n') in $\mathcal{F}$ such that
$$F_n' \subset F_n^\sigma \quad \text{and} \quad I(A \cap F_1' \cap \ldots \cap F_n') > r\ .$$
Since
$$I(A) = I(A \cap F_1^\sigma) = \sup_j I(A \cap F_1^j) > r\ ,$$
we may choose for F_1' an F_1^j with sufficiently high index j. If $F_1', \ldots, F_n'$ in $\mathcal{F}$ have been found, we note that
$$\begin{aligned} I(A \cap F_1' \ldots \cap F_n') &= I(A \cap F_1' \cap \ldots \cap F_n' \cap F_{n+1}^\sigma) \\ &= \sup_j I(A \cap F_1' \cap \ldots \cap F_n' \cap F_{n+1}^j) > r\ ; \end{aligned}$$
for F_{n+1}' we choose F_{n+1}^j with j sufficiently large. The construction of the F_n' is complete. Now $F^\delta \stackrel{\text{def}}{=} \bigcap_{n=1}^\infty F_n'$ is an $\mathcal{F}_\delta$-set and is contained in A, inasmuch as it is contained in every one of the F_n^σ. The continuity along decreasing sequences of $\mathcal{F}$ gives $I(F^\delta) \ge r$. The claim is proved for $A \in \mathcal{F}_{\sigma\delta}$. Now let A be a general $\mathcal{F}$-analytic set and $r < I(A)$. There are an auxiliary compactly paved set $(K, \mathcal{K})$ and an $(\mathcal{K} \times \mathcal{F})_{\sigma\delta}$-set $B \subset K \times F$ whose projection on F is A. We may assume that $\mathcal{K}$ is closed under taking finite intersections by the simple expedient of adjoining to $\mathcal{K}$ the intersections of its finite subcollections (exercise A.5.6). The paving $\mathcal{K} \otimes \mathcal{F}$ of $K \times F$ is then closed under both finite unions and finite intersections, and B still belongs to $(\mathcal{K} \otimes \mathcal{F})_{\sigma\delta}$. Due to lemma A.5.8 (ii), $I \circ \pi_F$ is a $\mathcal{K} \otimes \mathcal{F}$-capacity with $r < I \circ \pi_F(B)$, so the above provides a set $C \subset B$ in $(\mathcal{K} \otimes \mathcal{F})_\delta$ with $r < I\big(\pi_F(C)\big)$. Clearly $F_r \stackrel{\text{def}}{=} \pi_F(C)$ is a subset of A with $r < I(F_r)$. Now C is the intersection of a decreasing family $C_n \in \mathcal{K} \otimes \mathcal{F}$, each of which has $\pi_F(C_n) \in \mathcal{F}$, so by lemma A.5.8 (i) $F_r = \bigcap_n \pi_F(C_n) \in \mathcal{F}_\delta$. Since $r < I(A)$ was arbitrary, A is $(\mathcal{F}, I)$-capacitable. ∎

Applications to Stochastic Analysis

Theorem A.5.10 (The Measurable Section Theorem) *Let* $(\Omega, \mathcal{F})$ *be a measurable space and* $B \subset \mathbb{R}_+ \times \Omega$ *measurable on* $\mathcal{B}^\bullet(\mathbb{R}_+) \otimes \mathcal{F}$.

(i) For every $\mathcal{F}$*-capacity*[42] I *and* $\epsilon > 0$ *there is an* $\mathcal{F}$*-measurable function* $R : \Omega \to \overline{\mathbb{R}}_+$, *"an* $\mathcal{F}$*-measurable random time," whose graph is contained in* B *and such that*

$$I[R < \infty] > I[\pi_\Omega(B)] - \epsilon .$$

(ii) $\pi_\Omega(B)$ *is measurable on the universal completion* $\mathcal{F}^*$.

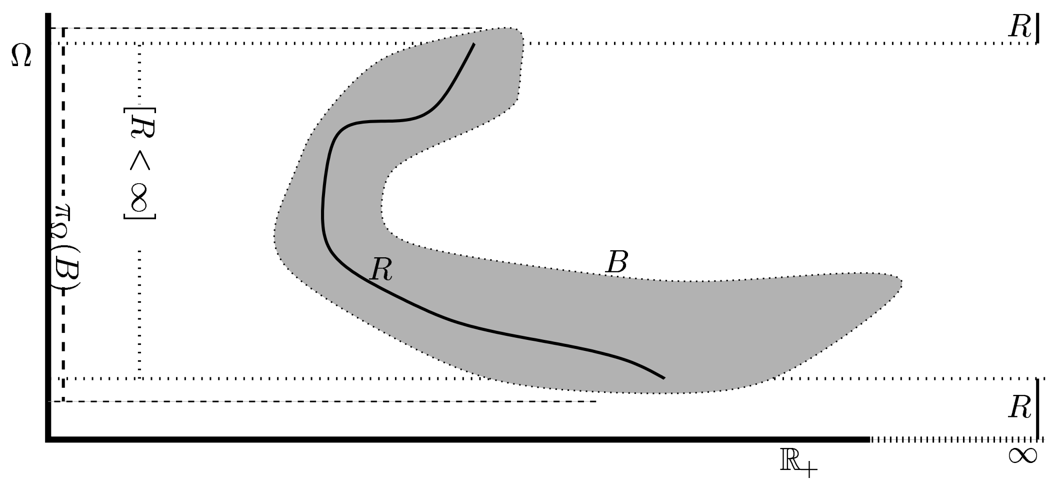


Figure A.17 The Measurable Section Theorem

Proof. (i) π_Ω denotes, of course, the natural projection of $\boldsymbol{B}$ onto Ω. We equip $\mathbb{R}_+$ with the paving $\mathcal{K}$ of compact intervals. On $\Omega \times \mathbb{R}_+$ consider the pavings

$$\mathcal{K} \times \mathcal{F} \quad \text{and} \quad \mathcal{K} \otimes \mathcal{F} .$$

The latter is closed under finite unions and intersections and generates the σ-algebra $\mathcal{B}^\bullet(\mathbb{R}_+) \otimes \mathcal{F}$. For every set $M = \bigcup_i [s_i, t_i] \times A_i$ in $\mathcal{K} \otimes \mathcal{F}$ and every $\omega \in \Omega$ the path $M_\bullet(\omega) = \bigcup_{i : A_i(\omega) \neq \emptyset} [s_i, t_i]$ is a compact subset of $\mathbb{R}_+$. Inasmuch as the complement of every set in $\mathcal{K} \times \mathcal{F}$ is the countable union of sets in $\mathcal{K} \times \mathcal{F}$, the paving of $\mathbb{R}_+ \times \Omega$ which generates the σ-algebra $\mathcal{B}^\bullet(\mathbb{R}_+) \otimes \mathcal{F}$, every set of $\mathcal{B}^\bullet(\mathbb{R}_+) \otimes \mathcal{F}$, in particular B, is $\mathcal{K} \times \mathcal{F}$-analytic (corollary A.5.3) and *a fortiori* $\mathcal{K} \otimes \mathcal{F}$-analytic. Next consider the set function

$$F \mapsto J(F) \stackrel{\text{def}}{=} I[\pi(F)] = I \circ \pi_\Omega(F) , \qquad F \subset \boldsymbol{B}.$$

According to lemma A.5.8, J is a $\mathcal{K} \otimes \mathcal{F}$-capacity. Choquet's theorem provides a set $K \in (\mathcal{K} \otimes \mathcal{F})_\delta$, the intersection of a decreasing countable family

[42] In most applications I is the outer measure $\mathbb{P}^*$ of a probability $\mathbb{P}$ on $\mathcal{F}$, which by equation (A.3.2) is a capacity.

$\{C_n\} \subset \mathcal{K} \otimes \mathcal{F}$, that is contained in B and has $J(K) > J(B) - \epsilon$. The "left edges" $R_n(\omega) \stackrel{\text{def}}{=} \inf\{t : (t, \omega) \in C_n\}$ are simple $\mathcal{F}$-measurable random variables, with $R_n(\omega) \in C_m(\omega)$ for $n \geq m$ at points ω where $R_n(\omega) < \infty$. Therefore $R \stackrel{\text{def}}{=} \sup_n R_n$ is $\mathcal{F}$-measurable, and thus $R(\omega) \in \bigcap_m C_m(\omega)$ $= K(\omega) \subset B(\omega)$ where $R(\omega) < \infty$. Clearly $[R < \infty] = \pi_\Omega[K] \in \mathcal{F}$ has $I[R < \infty] > I[\pi_\Omega(B)] - \epsilon$.

(ii) To say that the filtration $\mathcal{F}_.$ is ***universally complete*** means of course that $\mathcal{F}_t$ is universally complete for all $t \in [0, \infty]$ ($\mathcal{F}_t = \mathcal{F}_t^*$; see page 407); and this is certainly the case if $\mathcal{F}_.$ is $\mathfrak{P}$-regular, no matter what the collection $\mathfrak{P}$ of pertinent probabilities. Let then $\mathbb{P}$ be a probability on $\mathcal{F}$, and R_n $\mathcal{F}$-measurable random times whose graphs are contained in B and that have $\mathbb{P}^*[\pi_\Omega(B)] < \mathbb{P}[R_n < \infty] + 1/n$. Then $A \stackrel{\text{def}}{=} \bigcup_n [R_n < \infty] \in \mathcal{F}$ is contained in $\pi_\Omega(B)$ and has $\mathbb{P}[A] = \mathbb{P}^*[\pi_\Omega(B)]$: the inner and outer measures of $\pi_\Omega(B)$ agree, and so $\pi_\Omega(B)$ is $\mathbb{P}$-measurable. This is true for every probability $\mathbb{P}$ on $\mathcal{F}$, so $\pi_\Omega(B)$ is universally measurable. ∎

A slight refinement of the argument gives further information:

Corollary A.5.11 *Suppose that the filtration $\mathcal{F}_.$ is universally complete, and let T be a stopping time. Then the projection $\pi_\Omega[B]$ of a progressively measurable set $B \subset [\![0, T]\!]$ is measurable on $\mathcal{F}_T$.*

Proof. Fix an instant $t < \infty$. We have to show that $\pi_\Omega[B] \cap [T \leq t] \in \mathcal{F}_t$. Now this set equals the intersection of $\pi_\Omega[B^t]$ with $[T \leq t]$, so as $[T \leq t] \in \mathcal{F}_T$ it suffices to show that $\pi_\Omega[B^t] \in \mathcal{F}_t$. But this is immediate from theorem A.5.10 (ii) with $\mathcal{F} = \mathcal{F}_t$, since the stopped process B^t is measurable on $\mathcal{B}^\bullet(\mathbb{R}_+) \otimes \mathcal{F}_t$ by the very definition of progressive measurability. ∎

Corollary A.5.12 (First Hitting Times Are Stopping Times) *If the filtration $\mathcal{F}_.$ is right-continuous and universally complete, in particular if it satisfies the natural conditions, then the* ***debut***

$$D_B(\omega) \stackrel{\text{def}}{=} \inf\{t : (t, \omega) \in B\}$$

of a progressively measurable set $B \subset \boldsymbol{B}$ is a stopping time.

Proof. Let $0 \leq t < \infty$. The set $B \cap [\![0, t)\!)$ is progressively measurable and contained in $[\![0, t]\!]$, and its projection on Ω is $[D_B < t]$. Due to the universal completeness of $\mathcal{F}_.$, $[D_B < t]$ belongs to $\mathcal{F}_t$ (corollary A.5.11). Due to the right-continuity of $\mathcal{F}_.$, D_B is a stopping time (exercise 1.3.30 (i)). ∎

Corollary A.5.12 is a pretty result. Consider for example a progressively measurable process Z with values in some measurable state space $(S, \mathcal{S})$, and let $A \in \mathcal{S}$. Then $T_A \stackrel{\text{def}}{=} \inf\{t : Z_t \in A\}$ is the debut of the progressively measurable set $B \stackrel{\text{def}}{=} [Z \in A]$ and is therefore a stopping time. T_A is the "first time Z hits A," or better "the last time Z has not touched A." We can of course not claim that Z is in A at that time. If Z is right-continuous, though, and A is closed, then B is left-closed and $Z_{T_A} \in A$.

Corollary A.5.13 (The Progressive Measurability of the Maximal Process) *If the filtration $\mathcal{F}_{\bullet}$ is universally complete, then the maximal process $X^\star$ of a progressively measurable process X is progressively measurable.*

Proof. Let $0 \le t < \infty$ and $a > 0$. The set $[X^\star_t > a]$ is the projection on Ω of the $\mathcal{B}^\bullet[0,\infty) \otimes \mathcal{F}_t$-measurable set $[|X^t| > a]$ and is by theorem A.5.10 measurable on $\mathcal{F}_t = \mathcal{F}^*_t$: $X^\star$ is adapted. Next let T be the debut of $[X^\star > a]$. It is identical with the debut of $[|X| > a]$, a progressively measurable set, and so is a stopping time on the right-continuous version $\mathcal{F}_{\bullet+}$ (corollary A.5.12). So is its reduction S to $[|X|_T > a] \in \mathcal{F}_{T+}$ (proposition 1.3.9). Now clearly $[X^\star > a] = [\![S]\!] \cup (\!(T,\infty)\!)$. This union is progressively measurable for the filtration $\mathcal{F}_{\bullet+}$. This is obvious for $(\!(T,\infty)\!)$ (proposition 1.3.5 and exercise 1.3.17) and also for the set $[\![S]\!] = \bigcap [\![S, S+1/n)\!)$ (ibidem). Since this holds for all $a > 0$, $X^\star$ is progressively measurable for $\mathcal{F}_{\bullet+}$. Now apply exercise 1.3.30 (v). ∎

Theorem A.5.14 (Predictable Sections) *Let $(\Omega, \mathcal{F}_{\bullet})$ be a filtered measurable space and $B \subset \boldsymbol{B}$ a predictable set. For every $\mathcal{F}_\infty$-capacity I[42] and every $\epsilon > 0$ there exists a predictable stopping time R whose graph is contained in B and that satisfies (see figure A.17)*

$$I[R < \infty] > I[\pi_\Omega(B)] - \epsilon \ .$$

Proof. Consider the collection $\mathcal{M}$ of finite unions of stochastic intervals of the form $[\![S,T]\!]$, where S, T are predictable stopping times. The arbitrary left-continuous stochastic intervals

$$(\!(S,T]\!] = \bigcup_n \bigcap_k [\![S+1/n, T+1/k]\!] \in \mathcal{M}_{\delta\sigma} \ ,$$

with S, T arbitrary stopping times, generate the predictable σ-algebra $\mathcal{P}$ (exercise 2.1.6), and then so does $\mathcal{M}$. Let $[\![S,T]\!]$, S, T predictable, be an element of $\mathcal{M}$. Its complement is $[\![0,S)\!) \cup (\!(T,\infty)\!)$. Now $(\!(T,\infty)\!) = \bigcup [\![T+1/n, n]\!]$ is $\mathcal{M}$-analytic as a member of $\mathcal{M}_\sigma$ (corollary A.5.3), and so is $[\![0,S)\!)$. Namely, if S_n is a sequence of stopping times announcing S, then $[\![0,S)\!) = \bigcup [\![0,S_n]\!]$ belongs to $\mathcal{M}_\sigma$. Thus every predictable set, in particular B, is $\mathcal{M}$-analytic.

Consider next the set function

$$F \mapsto J(F) \overset{\text{def}}{=} I[\pi_\Omega(F)] \ , \qquad F \subset \boldsymbol{B}.$$

We see as in the proof of theorem A.5.10 that J is an $\mathcal{M}$-capacity. Choquet's theorem provides a set $K \in \mathcal{M}_\delta$, the intersection of a decreasing countable family $\{M_n\} \subset \mathcal{M}$, that is contained in B and has $J(K) > J(B) - \epsilon$. The "left edges" $R_n(\omega) \overset{\text{def}}{=} \inf\{t : (t,\omega) \in M_n\}$ are predictable stopping times with $R_n(\omega) \in M_m(\omega)$ for $n \ge m$. Therefore $R \overset{\text{def}}{=} \sup_n R_m$ is a predictable stopping time (exercise 3.5.10). Also, $R(\omega) \in \bigcap_m M_m(\omega) = K(\omega) \subset B(\omega)$ where $R(\omega) < \infty$. Evidently $[R < \infty] = \pi_\Omega[K]$, and therefore $I[R < \infty] > I[\pi_\Omega(B)] - \epsilon$. ∎

Corollary A.5.15 (The Predictable Projection) *Let X be a bounded measurable process. For every probability $\mathbb{P}$ on $\mathcal{F}_\infty$ there exists a predictable process $X^{\mathcal{P},\mathbb{P}}$ such that for all predictable stopping times T*

$$\mathbb{E}\big[X_T[T<\infty]\big] = \mathbb{E}\big[X_T^{\mathcal{P},\mathbb{P}}[T<\infty]\big] \ .$$

*$X^{\mathcal{P},\mathbb{P}}$ is called **a predictable $\mathbb{P}$-projection** of X. Any two predictable $\mathbb{P}$-projections of X cannot be distinguished with $\mathbb{P}$.*

Proof. Let us start with the uniqueness. If $X^{\mathcal{P},\mathbb{P}}$ and $\overline{X}^{\mathcal{P},\mathbb{P}}$ are predictable projections of X, then $N \stackrel{\text{def}}{=} \big[X^{\mathcal{P},\mathbb{P}} > \overline{X}^{\mathcal{P},\mathbb{P}}\big]$ is a predictable set. It is $\mathbb{P}$-evanescent. Indeed, if it were not, then there would exist a predictable stopping time T with its graph contained in N and $\mathbb{P}[T<\infty]>0$; then we would have $\mathbb{E}\big[X_T^{\mathcal{P},\mathbb{P}}\big] > \mathbb{E}\big[\overline{X}_T^{\mathcal{P},\mathbb{P}}\big]$, a plain impossibility. The same argument shows that if $X \le Y$, then $X^{\mathcal{P},\mathbb{P}} \le Y^{\mathcal{P},\mathbb{P}}$ except in a $\mathbb{P}$-evanescent set.

Now to the existence. The family $\mathcal{M}$ of bounded processes that have a predictable projection is clearly a vector space containing the constants, and a monotone class. For if X^n have predictable projections $X^{n\,\mathcal{P},\mathbb{P}}$ and, say, increase to X, then $\limsup X^{n\,\mathcal{P},\mathbb{P}}$ is evidently a predictable projection of X. $\mathcal{M}$ contains the processes of the form $X = (t,\infty)\times g$, $g \in L^\infty(\mathcal{F}^*_\infty)$, which generate the measurable σ-algebra. Indeed, a predictable projection of such a process is $M^g_{.-}\cdot(t,\infty)$. Here M^g is the right-continuous martingale $M^g_t = \mathbb{E}[g|\mathcal{F}_t]$ (proposition 2.5.13) and $M^g_{.-}$ its left-continuous version. For let T be a predictable stopping time, announced by (T_n), and recall from lemma 3.5.15 (ii) that the strict past of T is $\bigvee \mathcal{F}_{T_n}$ and contains $[T>t]$.

Thus $\qquad \mathbb{E}[g|\mathcal{F}_{T-}] = \mathbb{E}\big[g|\bigvee \mathcal{F}_{T_n}\big]$

by exercise 2.5.5: $\qquad = \lim \mathbb{E}\big[g|\mathcal{F}_{T_n}\big]$

by theorem 2.5.22: $\qquad = \lim M^g_{T_n} = M^g_{T-}$

$\mathbb{P}$-almost surely, and therefore

$$\begin{aligned}\mathbb{E}\big[X_T\cdot[T<\infty]\big] &= \mathbb{E}\big[g\cdot[T>t]\big] \\ &= \mathbb{E}\big[\mathbb{E}[g|\mathcal{F}_{T-}]\cdot[T>t]\big] = \mathbb{E}\big[M^g_{T-}\cdot[T>t]\big] \\ &= \mathbb{E}\big[\big(M^g_{.-}\cdot(t,\infty)\big)_T\cdot[T<\infty]\big] \ .\end{aligned}$$

This argument has a flaw: M^g is generally adapted only to the natural enlargement $\mathcal{F}^{\mathbb{P}}_{.+}$ and $M^g_{.-}$ only to the $\mathbb{P}$-regularization $\mathcal{F}^{\mathbb{P}}$. It can be fixed as follows. For every dyadic rational q let $\overline{M}^g_q$ be an $\mathcal{F}_q$-measurable random variable $\mathbb{P}$-nearly equal to M^g_{q-} (exercise 1.3.33) and set

$$\overline{M}^{g,n} \stackrel{\text{def}}{=} \sum_k \overline{M}^g_{k2^{-n}}\big(k2^{-n},(k+1)2^{-n}\big] \ .$$

This is a predictable process, and so is $\overline{M}^g \stackrel{\text{def}}{=} \limsup_n \overline{M}^{g,n}$. Now the paths of $\overline{M}^g$ differ from those of $M^g_{.-}$ only in the $\mathbb{P}$-nearly empty set $\bigcup_q [M^g_{q-} \neq \overline{M}^g_q]$. So $\overline{M}^g \cdot (t,\infty)$ is a predictable projection of $X = (t,\infty) \times g$. An application of the monotone class theorem A.3.4 finishes the proof. ∎

Exercise A.5.16 For any predictable right-continuous increasing process I

$$\mathbb{E}\Big[\int X \, dI\Big] = \mathbb{E}\Big[\int X^{\mathcal{P},\mathbb{P}} \, dI\Big] .$$

Supplements and Additional Exercises

Definition A.5.17 (Optional or Well-Measurable Processes) *The σ-algebra generated by the càdlàg adapted processes is called the σ-algebra of **optional** or **well-measurable** sets on $\boldsymbol{B}$ and is denoted by $\mathcal{O}$. A function measurable on $\mathcal{O}$ is an **optional** or **well-measurable process**.*

Exercise A.5.18 The optional σ-algebra $\mathcal{O}$ is generated by the right-continuous stochastic intervals $[\![S,T)\!)$, contains the previsible σ-algebra $\mathcal{P}$, and is contained in the σ-algebra of progressively measurable sets. For every optional process X there exist a predictable process X' and a countable family $\{T^n\}$ of stopping times such that $[X \neq X']$ is contained in the union $\bigcup_n [\![T^n]\!]$ of their graphs.

Corollary A.5.19 (The Optional Section Theorem) *Suppose that the filtration $\mathcal{F}_{\bullet}$ is right-continuous and universally complete, let $\mathcal{F} = \mathcal{F}_\infty$, and let $B \subset \boldsymbol{B}$ be an optional set. For every $\mathcal{F}$-capacity I [42] and every $\epsilon > 0$ there exists a stopping time R whose graph is contained in B and which satisfies (see figure A.17)*

$$I[R < \infty] > I[\pi_\Omega(B)] - \epsilon .$$

[Hint: Emulate the proof of theorem A.5.14, replacing $\mathcal{M}$ by the finite unions of arbitrary right-continuous stochastic intervals $[\![S,T)\!)$.]

Exercise A.5.20 (The Optional Projection) Let X be a measurable process. For every probability $\mathbb{P}$ on $\mathcal{F}_\infty$ there exists a process $X^{\mathcal{O},\mathbb{P}}$ that is measurable on the optional σ-algebra of the natural enlargement $\mathcal{F}^{\mathbb{P}}_{.+}$ and such that for all stopping times

$$\mathbb{E}[X_T[T < \infty]] = \mathbb{E}[X^{\mathcal{O},\mathbb{P}}_T[T < \infty]] .$$

$X^{\mathcal{O},\mathbb{P}}$ is called ***an optional*** $\mathbb{P}$***-projection*** of X. Any two optional $\mathbb{P}$-projections of X are indistinguishable with $\mathbb{P}$.

Exercise A.5.21 (The Optional Modification) Assume that the measured filtration is right-continuous and regular. Then an adapted measurable process X has an optional modification.

A.6 Suslin Spaces and Tightness of Measures

Polish and Suslin Spaces

A topological space is ***polish*** if it is Hausdorff and separable and if its topology can be defined by a metric under which it is complete. Exercise 1.2.4 on page 15 amounts to saying that the path space $\mathscr{C}$ is polish. The name seems to have arisen this way: the Poles decided that, being a small nation, they should concentrate their mathematical efforts in one area and do it well rather than spread themselves thin. They chose analysis, extending the achievements of the great Pole Banach. They excelled. The theory of analytic spaces, which are the continuous images of polish spaces, is essentially due to them. A Hausdorff topological space F is called a ***Suslin space*** if there exists a polish space P and a continuous surjection $p : P \to F$. A Suslin space is evidently separable.[43] If a continuous injective surjection $p : P \to F$ can be found, then F is a ***Lusin space***. A subset of a Hausdorff space is a ***Suslin set*** or a ***Lusin set***, of course, if it is Suslin or Lusin in the induced topology. The attraction of Suslin spaces in the context of measure theory is this: they contain an abundance of large compact sets: every σ-additive measure on their Borels is tight. Henceforth $\mathcal{G}$ or $\mathcal{K}$ denote the open or compact sets of the topological space at hand, respectively.

Exercise A.6.1 (i) If P is polish, then there exists a compact metric space $\widehat{P}$ and a homeomorphism j of P onto a dense subset of $\widehat{P}$ that is both a $\mathcal{G}_\delta$-set and a $\mathcal{K}_{\sigma\delta}$-set of $\widehat{P}$.

(ii) A closed subset and a continuous Hausdorff image of a Suslin set are Suslin. This fact is the clue to everthing that follows.

(iii) The union and intersection of countably many Suslin sets are Suslin.

(iv) In a metric Suslin space every Borel set is Suslin.

Proposition A.6.2 *Let F be a Suslin space and $\mathcal{E}$ be an algebra of continuous bounded functions that separates the points of F, e.g., $\mathcal{E} = C_b(F)$.*

(i) Every Borel subset of F is $\mathcal{K}$-analytic.

(ii) $\mathcal{E}$ contains a countable algebra $\mathcal{E}_0$ over $\mathbb{Q}$ that still separates the points of F. The topology generated by $\mathcal{E}_0$ is metrizable, Suslin, and weaker than the given one (but not necessarily strictly weaker).

(iii) Let $m : \mathcal{E} \to \mathbb{R}$ be a positive σ-continuous linear functional with $\|m\| \stackrel{\text{def}}{=} \sup\{m(\phi) : \phi \in \mathcal{E}, |\phi| \le 1\} < \infty$. Then the Daniell extension of m integrates all bounded Borel functions. There exists a unique σ-additive measure μ on $\mathcal{B}^\bullet(F)$ that represents m:

$$m(\phi) = \int \phi \, d\mu \,, \qquad \phi \in \mathcal{E} \,.$$

This measure is tight and inner regular, and order-continuous on $C_b(F)$.

Proof. Scaling reduces the situation to the case that $\|m\| = 1$. Also, the Daniell extension of m certainly integrates any function in the uniform closure of $\mathcal{E}$ and is σ-continuous thereon. We may thus assume without loss of generality that $\mathcal{E}$ is uniformly closed and thus is both an algebra and a vector lattice (theorem A.2.2). Fix a polish space P and a continuous surjection $p : P \to F$. There are several steps.

(ii) There is a countable subset $\Phi \subset \mathcal{E}$ that still separates the points of F. To see this note that $P \times P$ is again separable and metrizable, and

[43] In the literature Suslin and Lusin spaces are often metrizable by definition. We don't require this, so we don't have to check a topology for metrizability in an application.

let $\mathcal{U}_0[P{\times}P]$ be a countable uniformly dense subset of $\mathcal{U}[P{\times}P]$ (see lemma A.2.20). For every $\phi \in \mathcal{E}$ set $g_\phi(x',y') \stackrel{\text{def}}{=} |\phi(p(x')) - \phi(p(y'))|$, and let $\mathcal{U}^{\mathcal{E}}$ denote the countable collection

$$\{f \in \mathcal{U}_0[P{\times}P] :\ \exists \phi \in \mathcal{E} \text{ with } f \le g_\phi\} .$$

For every $f \in \mathcal{U}^{\mathcal{E}}$ select one particular $\phi_f \in \mathcal{E}$ with $f \le g_{\phi_f}$, thus obtaining a countable subcollection Φ of $\mathcal{E}$. If $x = p(x') \ne p(y') = y$ in F, then there are a $\phi \in \mathcal{E}$ with $0 < |\phi(x) - \phi(y)| = g_\phi(x',y')$ and an $f \in \mathcal{U}[P \times P]$ with $f \le g_\phi$ and $f(x',y') > 0$. The function $\phi_f \in \Phi$ has $g_{\phi_f}(x',y') > 0$, which signifies that $\phi_f(x) \ne \phi_f(y)$: $\Phi \subset \mathcal{E}$ still separates the points of F. The finite $\mathbb{Q}$-linear combinations of finite products of functions in Φ form a countable $\mathbb{Q}$-algebra $\mathcal{E}_0$.

Henceforth m' denotes the restriction m to the uniform closure $\mathcal{E}' \stackrel{\text{def}}{=} \overline{\mathcal{E}}_0$ in $\mathcal{E} = \overline{\mathcal{E}}$. Clearly $\mathcal{E}'$ is a vector lattice and algebra and m' is a positive linear σ-continuous functional on it.

Let $j : F \to \widehat{F}$ denote the local $\mathcal{E}_0$-compactification of F provided by theorem A.2.2 (ii). $\widehat{F}$ is metrizable and j is injective, so that we may identify F with the dense subset $j(F)$ of $\widehat{F}$. Note however that the Suslin topology of F is a priori finer than the topology induced by $\widehat{F}$. Every $\phi \in \mathcal{E}'$ has a unique extension $\widehat{\phi} \in C_0(\widehat{F})$ that agrees with ϕ on F, and $\widehat{\mathcal{E}'} = C_0(\widehat{F})$. Let us define $\widehat{m}' : C(\widehat{F}) \to \mathbb{R}$ by $\widehat{m}'(\widehat{\phi}) = m'(\phi)$. This is a Radon measure on $\widehat{F}$. Thanks to Dini's theorem A.2.1, $\widehat{m}'$ is automatically σ-continuous. We convince ourselves next that F has upper integral (= outer measure) 1. Indeed, if this were not so, then the inner measure $\widehat{m}'_*(\widehat{F} - F) = 1 - \widehat{m}'^*(F)$ of its complement would be strictly positive. There would be a function $k \in C_\downarrow(\widehat{F})$, pointwise limit of a decreasing sequence $\widehat{\phi}_n$ in $C(\widehat{F})$, with $k \le \widehat{F} - F$ and $0 < \widehat{m}'_*(k) = \lim \widehat{m}'(\widehat{\phi}_n) = \lim m'(\phi_n)$. Now (ϕ_n) decreases to zero on F and therefore the last limit is zero, a plain contradiction.

So far we do not know that F *is* $\widehat{m}'$*-measurable in* $\widehat{F}$. To prove that it is it will suffice to show that F is $\mathcal{K}$-analytic in $\widehat{F}$: since $\widehat{m}'^*$ is a $\mathcal{K}$-capacity, Choquet's theorem will then provide compact sets $K_n \subset F$ with $\sup \widehat{m}'(K_n) = \sup \widehat{m}'^*(K_n) = 1$, showing that the inner measure of F equals 1 as well, so that F is measurable. To show the analyticity of F we embed P homeomorphically as a $\mathcal{G}_\delta$ in a compact metric space $\widehat{P}$, as in exercise A.6.1. Actually, we shall view P as a $\mathcal{G}_\delta$ in $\Pi \stackrel{\text{def}}{=} \widehat{F} \times \widehat{P}$, embedded via $x \mapsto (p(x), x)$. The projection π of Π on its first factor $\widehat{F}$ coincides with p on P.

Now the topologies of both $\widehat{F}$ and $\widehat{P}$ have a countable basis of relatively compact open balls, and the rectangles made from them constitute a countable basis for the topology of Π. Therefore every open set of Π is the countable union of compact rectangles and is thus a $\mathcal{K}^\times_\sigma$-set for the paving $\mathcal{K}^\times \stackrel{\text{def}}{=} \mathcal{K}[\widehat{F}] \times \mathcal{K}[\widehat{P}]$. The complement of a set in $\mathcal{K}^\times$ is the countable union of sets in $\mathcal{K}^\times$, so every Borel set of Π is $\mathcal{K}^\times$-analytic (corollary A.5.3). In particular,

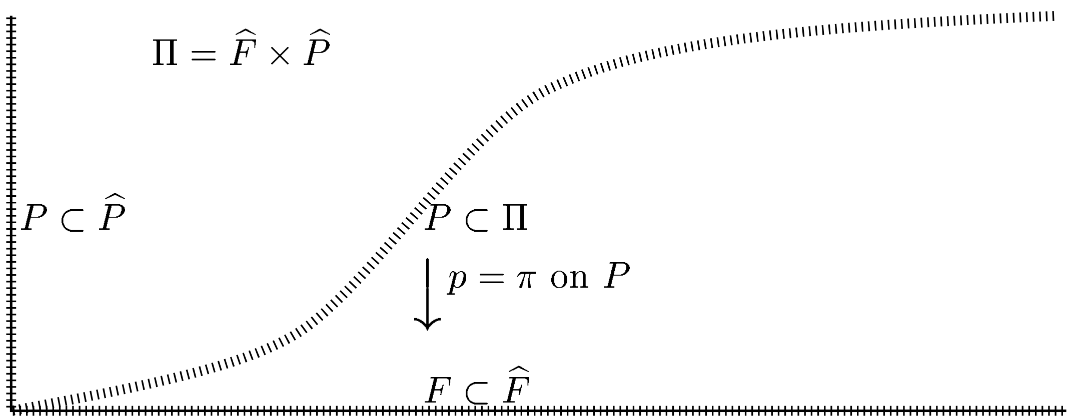


Figure A.18 Compactifying the situation

P, which is the countable intersection of open subsets of Π (exercise A.6.1), is $\mathcal{K}^{\times}$-analytic. By proposition A.5.4, $F = \pi(P)$ is $\mathcal{K}[\widehat{F}]$-analytic and *a fortiori* $\mathcal{K}[F]$-analytic.

This argument also shows that *every Borel set B of F is analytic.* Namely, $B^{\uparrow} \stackrel{\text{def}}{=} p^{-1}(B)$ is a Borel set of P and then of Π, and its image $B = \pi(B^{\uparrow})$ is $\mathcal{K}[\widehat{F}]$-analytic in F and in $\widehat{F}$. As above we see that therefore $\widehat{m}'^{*}(B) = \sup\{\int K \, d\widehat{m}' : K \subset B, K \in \mathcal{K}[\widehat{F}]\}$: $B \mapsto \mu(B) \stackrel{\text{def}}{=} \int B \, d\widehat{m}'$ is an inner regular measure representing m.

Only the uniqueness of μ remains to be shown. Assume then that μ' is another measure on the Borel sets of F that agrees with m on $\mathcal{E}$. Taking $C_b(F)$ instead of $\mathcal{E}$ in the previous arguments shows that μ' is inner regular. By theorem A.3.4, μ and μ' agree on the sets of $\mathcal{K}[\widehat{F}]$ in F, by the inner regularity on all Borel sets. ∎

A.7 The Skorohod Topology

In this section (E, ρ) is a fixed complete separable metric space. Replacing if necessary ρ by $\rho \wedge 1$, we may and shall assume that $\rho \le 1$. We consider the set $\mathscr{D}_E$ of all paths $z : [0, \infty) \to E$ that are right-continuous and have left limits $z_{t-} \stackrel{\text{def}}{=} \lim_{s \uparrow t} z_s \in E$ at all finite instants $t > 0$. Integrators and solutions of stochastic differential equations are random variables with values in $\mathscr{D}_{\mathbb{R}^n}$. In the stochastic analysis of them the maximal process plays an important role. It is finite at finite times (lemma 2.3.2), which seems to indicate that the topology of uniform convergence on bounded intervals is the appropriate topology on $\mathscr{D}$. This topology is not Suslin, though, as it is not separable: the functions $[0, t)$, $t \in \mathbb{R}_+$, are uncountable in number, but any two of them have uniform distance 1 from each other. Results invoking tightness, as for instance proposition A.6.2 and theorem A.3.17, are not applicable.

Skorohod has given a polish topology on $\mathscr{D}_E$. It rests on the idea that temporal as well as spatial measurements are subject to errors, and that paths that can be transformed into each other by small deformations of space and time should be considered close. For instance, if $s \approx t$, then $[0,s)$ and $[0,t)$ should be considered close. Skorohod's topology of section A.7 makes the above-mentioned results applicable. It is not a panacea, though. It is not compatible with the vector space structure of $\mathscr{D}_{\mathbb{R}}$, thus rendering tools from Fourier analysis such as the characteristic function unusable; the rather useful example A.4.13 genuinely uses the topology of uniform convergence – the functions ϕ appearing therein are not continuous in the Skorohod topology.

It is most convenient to study Skorohod's topology first on a bounded time-interval $[0,u]$, that is to say, on the subspace $\mathscr{D}_E^u \subset \mathscr{D}_E$ of paths z that stop at the instant u: $z = z^u$ in the notation of page 23. We shall follow rather closely the presentation in Billingsley [10]. There are two equivalent metrics for the Skorohod topology whose convenience of employ depends on the situation. Let Λ denote the collection of all strictly increasing functions from $[0,\infty)$ onto itself, the "time transformations." The first Skorohod metric $d^{(0)}$ on $\mathscr{D}_E^u$ is defined as follows: for $z,y \in \mathscr{D}_E^u$, $d^{(0)}(z,y)$ is the infimum of the numbers $\epsilon > 0$ for which there exists a $\lambda \in \Lambda$ with

$$\|\lambda\|^{(0)} \stackrel{\text{def}}{=} \sup_{0\le t<\infty} |\lambda(t) - t| < \epsilon$$

and

$$\sup_{0\le t<\infty} \rho\big(z_t, y_{\lambda(t)}\big) < \epsilon \,.$$

It is left to the reader to check that

$$\|\,\lambda\,\|^{(0)} = \big\|\,\lambda^{-1}\,\big\|^{(0)} \quad \text{and} \quad \|\,\lambda\circ\mu\,\|^{(0)} \le \|\,\lambda\,\|^{(0)} + \|\,\mu\,\|^{(0)}\,, \qquad \lambda,\mu \in \Lambda\,,$$

and that $d^{(0)}$ is a metric satisfying $d^{(0)}(z,y) \le \sup_t \rho(z_t,y_t)$. The topology of $d^{(0)}$ is called the Skorohod topology on $\mathscr{D}_E^u$. It is coarser than the uniform topology. A sequence $\big(z^{(n)}\big)$ of $\mathscr{D}_E^u$ converges to $z \in \mathscr{D}_E^u$ if and only if there exist $\lambda_n \in \Lambda$ with $\|\,\lambda_n\,\|^{(0)} \to 0$ and $z^{(n)}_{\lambda_n(t)} \to z_t$ uniformly in t.

We now need a couple of tools. The ***oscillation*** of any stopped path $z : [0,\infty) \to E$ on an interval $I \subset \mathbb{R}_+$ is $o_I[z] \stackrel{\text{def}}{=} \sup\big\{\rho(z_s,z_t) : s,t \in I\big\}$, and there are two pertinent ***moduli of continuity***:

$$\gamma^\delta[z] \stackrel{\text{def}}{=} \sup_{0\le t<\infty} o_{[t,t+\delta]}[z]$$

and $\gamma^\delta_{[)}[z] \stackrel{\text{def}}{=} \inf\Big\{\sup_i o_{[t_i,t_{i+1})}[z] : \ 0 = t_0 < t_1 < \ldots\ ;\ |t_{i+1} - t_i| > \delta\Big\}\,.$

They are related by $\gamma^\delta_{[)}[z] \le \gamma^{2\delta}[z]$. A stopped path $z = z^u : [0,\infty) \to\ E$ is in $\mathscr{D}_E^u$ if and only if $\gamma^\delta_{[)}[z] \xrightarrow[\delta\to 0]{} 0$ and is continuous if and only if $\gamma^\delta[z] \xrightarrow[\delta\to 0]{} 0$ – we then write $z \in \mathscr{C}_E^u$. Evidently

$$\begin{aligned}\rho\big(z_t, z^{(n)}_t\big) &\le \rho\Big(z_t, z_{\lambda_n^{-1}(t)}\Big) + \rho\Big(z_{\lambda_n^{-1}(t)}, z^{(n)}_t\Big) \le \rho\Big(z_t, z_{\lambda_n^{-1}(t)}\Big) + d^{(0)}\big(z, z^{(n)}\big) \\ &\le \gamma^{\|\lambda_n^{-1}\|}[z] + d^{(0)}\big(z, z^{(n)}\big)\,.\end{aligned}$$

The first line shows that $d^{(0)}\big(z, z^{(n)}\big) \to 0$ implies $z_t^{(n)} \to z_t$ in continuity points t of z, and the second that the convergence is uniform if $z = z^u$ is continuous (and then uniformly continuous). The Skorohod topology therefore coincides on $\mathscr{C}_E^u$ with the topology of uniform convergence.

It is separable. To see this let E_0 be a countable dense subset of E and let $\mathscr{D}^{(k)}$ be the collection of paths in $\mathscr{D}_E^u$ that are constant on intervals of the form $[i/k, (i+1)/k)$, $i \le ku$, and take a value from E_0 there. $\mathscr{D}^{(k)}$ is countable and $\bigcup_k \mathscr{D}^{(k)}$ is $d^{(0)}$-dense in $\mathscr{D}_E^u$.

However, $\mathscr{D}_E^u$ is not complete under this metric – $[u/2 - 1/n, u/2)$ is a $d^{(0)}$-Cauchy sequence that has no limit. There is an equivalent metric, one that defines the same topology, under which $\mathscr{D}_E^u$ is complete; thus $\mathscr{D}_E^u$ is polish in the Skorohod topology. The second Skorohod metric $d^{(1)}$ on $\mathscr{D}_E^u$ is defined as follows: for $z, y \in \mathscr{D}_E^u$, $d^{(1)}(z, y)$ is the infimum of the numbers $\epsilon > 0$ for which there exists a $\lambda \in \Lambda$ with

$$\|\lambda\|^{(1)} \stackrel{\text{def}}{=} \sup_{0 \le s < t < \infty} \ln \left| \frac{\lambda(t) - \lambda(s)}{t - s} \right| < \epsilon \tag{A.7.1}$$

and
$$\sup_{0 \le t < \infty} \rho\big(z_t, y_{\lambda(t)}\big) < \epsilon \,. \tag{A.7.2}$$

Roughly speaking, (A.7.1) restricts the time transformations to those with "slope close to unity." Again it is left to the reader to show that

$$\|\lambda\|^{(1)} = \big\|\lambda^{-1}\big\|^{(1)} \quad \text{and} \quad \|\lambda \circ \mu\|^{(1)} \le \|\lambda\|^{(1)} + \|\mu\|^{(1)} \,, \qquad \lambda, \mu \in \Lambda \,,$$

and that $d^{(1)}$ is a metric.

Theorem A.7.1 *(i) $d^{(0)}$ and $d^{(1)}$ define the same topology on $\mathscr{D}_E^u$.*

(ii) $(\mathscr{D}_E^u, d^{(1)})$ is complete. $\mathscr{D}_E$ is separable and complete under the metric

$$d(z, y) \stackrel{\text{def}}{=} \sum_{u \in \mathbb{N}} 2^{-u} \wedge d^{(1)}(z^u, y^u) \,, \qquad z, y \in \mathscr{D}_E \,.$$

The polish topology τ of d on $\mathscr{D}_E$ is called the ***Skorohod topology*** *and coincides on $\mathscr{C}_E$ with the topology of uniform convergence on bounded intervals and on $\mathscr{D}_E^u \subset \mathscr{D}_E$ with the topology of $d^{(0)}$ or $d^{(1)}$. The stopping maps $z \mapsto z^u$ are continuous projections from $(\mathscr{D}_E, d)$ onto $(\mathscr{D}_E^u, d^{(1)})$, $0 < u < \infty$.*

(iii) The Hausdorff topology σ on $\mathscr{D}_{\mathbb{R}^d}$ generated by the linear functionals

$$\boldsymbol{z} \mapsto \langle \boldsymbol{z} | \boldsymbol{\phi} \rangle \stackrel{\text{def}}{=} \int_0^\infty \sum_{a=1}^d z_s^{(a)} \phi_a(s)\, ds \,, \qquad \boldsymbol{\phi} \in C_{00}(\mathbb{R}_+, \mathbb{R}^d) \,,$$

is weaker than τ and makes $\mathscr{D}_{\mathbb{R}^d}$ into a Lusin topological vector space.

(iv) Let $\mathcal{F}_t^0$ denote the σ-algebra generated by the evaluations $z \mapsto z_s$, $0 \le s \le t$, the ***basic filtration of path space****. Then $\mathcal{F}_t^0 \subset \mathcal{B}^\bullet(\mathscr{D}_E^t, \tau)$, and $\mathcal{F}_t^0 = \mathcal{B}^\bullet(\mathscr{D}_E^t, \sigma)$ if $E = \mathbb{R}^d$.*

(v) Suppose $\mathbb{P}_t$ is, for every $t < \infty$, a σ-additive probability on $\mathcal{F}_t^0$, such that the restriction of $\mathbb{P}_t$ to $\mathcal{F}_s^0$ equals $\mathbb{P}_s$ for $s < t$. Then there exists a unique tight probability $\mathbb{P}$ on the Borels of $(\mathscr{D}_E, \tau)$ that equals $\mathbb{P}_t$ on $\mathcal{F}_t$ for all t.

Proof. (i) If $d^{(1)}(z,y) < \epsilon < 1/(4+u)$, then there is a $\lambda \in \Lambda$ with $\|\lambda\|^{(1)} < \epsilon$ satisfying (A.7.2). Since $\lambda(0) = 0$, $\ln(1-2\epsilon) < -\epsilon < \ln\big(\lambda(t)/t\big) < \epsilon < \ln(1+2\epsilon)$, which results in $|\lambda(t) - t| \le 2\epsilon t \le 2\epsilon(u+1) < 1/2$ for $0 \le t \le u+1$. Changing λ on $[u+1, \infty)$ to continue with slope 1 we get $\|\lambda\|^{(0)} \le 2\epsilon(u+1)$ and $d^{(0)}(z,y) \le 2\epsilon(u+1)$. That is to say, $d^{(1)}(z, z^{(n)}) \to 0$ implies $d^{(0)}(z, z^{(n)}) \to 0$. For the converse we establish the following claim: *if $d^{(0)}(z,y) < \delta^2 < 1/4$, then $d^{(1)}(z,y) \le 4\delta + \gamma_{\mathbb{D}}^{\delta}[z]$.* To see this choose instants $0 = t_0 < t_1 \ldots$ with $t_{i+1} - t_i > \delta$ and $o_{[t_i, t_{i+1})}[z] < \gamma_{\mathbb{D}}^{\delta}[z] + \delta$ and $\mu \in \Lambda$ with $\|\mu\|^{(0)} < \delta^2$ and $\sup_t \rho\big(z_{\mu^{-1}(t)}, y_t\big) < \delta^2$. Let λ be that element of Λ which agrees with μ at the instants t_i and is linear in between, $i = 0, 1, \ldots$. Clearly $\mu^{-1} \circ \lambda$ maps $[t_i, t_{i+1})$ to itself, and

$$\begin{aligned} \rho\big(z_t, y_{\lambda(t)}\big) &\le \rho\big(z_t, z_{\mu^{-1}\circ\lambda(t)}\big) + \rho\big(z_{\mu^{-1}\circ\lambda(t)}, y_{\lambda(t)}\big) \\ &\le \gamma_{\mathbb{D}}^{\delta}[z] + \delta + \delta^2 < 4\delta + \gamma_{\mathbb{D}}^{\delta}[z] \,. \end{aligned}$$

So if $d^{(0)}(z, z^{(n)}) \to 0$ and $0 < \epsilon < 1/2$ is given, we choose $0 < \delta < \epsilon/8$ so that $\gamma_{\mathbb{D}}^{\delta}[z] < \epsilon/2$ and then N so that $d^{(0)}(z, z^{(n)}) < \delta^2$ for $n > N$. The claim above produces $d^{(1)}(z, z^{(n)}) < \epsilon$ for such n.

(ii) Let $\big(z^{(n)}\big)$ be a $d^{(1)}$-Cauchy sequence in $\mathscr{D}_E^u$. Since it suffices to show that a subsequence converges, we may assume that $d^{(1)}\big(z^{(n)}, z^{(n+1)}\big) < 2^{-n}$. Choose $\mu_n \in \Lambda$ with

$$\|\mu_n\|^{(1)} < 2^{-n} \qquad \text{and} \qquad \sup_t \rho\big(z_t^{(n)}, z_{\mu_n(t)}^{(n+1)}\big) < 2^{-n} \,.$$

Denote by μ_n^{n+m} the composition $\mu_{n+m} \circ \mu_{n+m-1} \circ \cdots \circ \mu_n$. Clearly

$$\sup_t \Big| \mu_{n+m+1} \circ \mu_n^{n+m}(t) - \mu_n^{n+m}(t) \Big| < 2^{-n-m-1} \,,$$

and by induction $$\sup_t \Big| \mu_n^{n+m'}(t) - \mu_n^{n+m}(t) \Big| \le 2^{-n-m} \,, \quad 1 \le m < m' \,.$$

The sequence $\big(\mu_n^{n+m}\big)_{m=1}^{\infty}$ is thus uniformly Cauchy and converges uniformly to some function λ_n that has $\lambda_n(0) = 0$ and is increasing. Now

$$\ln \left| \frac{\mu_n^{n+m}(t) - \mu_n(s)}{t-s} \right| \le \sum\nolimits_{i=n+1}^{n+m} \|\mu_i\|^{(1)} \le 2^{-n} \,,$$

so that $\|\lambda_n\|^{(1)} \le 2^{-n}$. Therefore λ_n is strictly increasing and belongs to Λ. Also clearly $\lambda_n = \lambda_{n+1} \circ \mu_n$. Thus

$$\sup_t \rho\Big(z_{\lambda_n^{-1}(t)}^{(n)}, z_{\lambda_{n+1}^{-1}(t)}^{(n+1)}\Big) = \sup_t \rho\Big(z_t^{(n)}, z_{\mu^n(t)}^{(n+1)}\Big) \le 2^{-n} \,,$$

and the paths $z_{\lambda_n^{-1}}^{(n)}$ converge uniformly to some right-continuous path z

with left limits. Since $z^{(n)}$ is constant on $[\lambda_n^{-1} \ge u]$, and since $\lambda_n^{-1}(t) \to t$ uniformly on $[0, u+1]$, z is constant on $(u, \infty) \subset \liminf[\lambda_n^{-1} \ge u]$ and by right-continuity belongs to $\mathscr{D}_E^u$. Since $\|\lambda_n\|^{(1)} \xrightarrow[n\to\infty]{} 0$ and $\sup_t \rho\big(z_t, z^{(n)}_{\lambda_n^{-1}(t)}\big)$ converges to 0, $d^{(1)}\big(z, z^{(n)}\big) \xrightarrow[n\to\infty]{} 0$: $\big(\mathscr{D}_E^u, d^{(1)}\big)$ is indeed complete. The remaining statements of (ii) are left to the reader.

(iii) Let $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_d) \in C_{00}(\mathbb{R}_+, \mathbb{R}^d)$. To see that the linear functional $\boldsymbol{z} \mapsto \langle \boldsymbol{z} | \boldsymbol{\phi} \rangle$ is continuous in the Skorohod topology, let u be the supremum of supp $\boldsymbol{\phi}$. If $\boldsymbol{z}^{(n)} \to \boldsymbol{z}$, then $\sup_{n\in\mathbb{N}, t\le u} \big|\boldsymbol{z}_t^{(n)}\big|$ is finite and $\boldsymbol{z}_t^{(n)} \xrightarrow[n\to\infty]{} \boldsymbol{z}_t$ in all but the countably many points t where $\boldsymbol{z}_t$ jumps. By the DCT the integrals $\langle \boldsymbol{z}^{(n)} | \boldsymbol{\phi} \rangle$ converge to $\langle \boldsymbol{z} | \boldsymbol{\phi} \rangle$. The topology σ is thus weaker than τ. Since $\langle \boldsymbol{z}^{(n)} | \boldsymbol{\phi} \rangle = 0$ for all continuous $\boldsymbol{\phi} : [0, \infty) \to \mathbb{R}^d$ with compact support implies $\boldsymbol{z} \equiv 0$, this evidently linear topology is Lusin. Writing $\langle \boldsymbol{z} | \boldsymbol{\phi} \rangle$ as a limit of Riemann sums shows that the σ-Borels of $\mathscr{D}^t_{\mathbb{R}^d}$ are contained in $\mathcal{F}_t^0$. Conversely, letting $\boldsymbol{\phi}$ run through an ***approximate identity*** $\boldsymbol{\phi}_n$ supported on the right of $s \le t$ [44] shows that $\boldsymbol{z}_s = \lim \langle \boldsymbol{z} | \boldsymbol{\phi}_n \rangle$ is measurable on $\mathcal{B}^\bullet(\sigma)$, so that there is coincidence.

(iv) In general, when E is just some polish space, one can define a Lusin topology σ weaker than τ as well: it is the topology generated by the τ-continuous functions $z \mapsto \int_0^\infty \psi(z_s) \cdot \phi(s)\, ds$, where $\phi \in C_{00}(\mathbb{R}_+)$ and $\psi : E \to \mathbb{R}$ is uniformly continuous. It follows as above that $\mathcal{F}_t^0 = \mathcal{B}^\bullet(\mathscr{D}_E^t, \sigma)$.

(v) The $\mathbb{P}_t$ viewed as probabilities on the Borels of $(\mathscr{D}_E^t, \sigma)$ form a tight (proposition A.6.2) projective system, evidently full, under the stopping maps $\pi_u^t(z) \stackrel{\text{def}}{=} z^t$ [45]. There is thus on $\bigcup_t C_b\big(\mathscr{D}_E^t, \sigma\big) \circ \pi^t$ a σ-additive projective limit $\mathbb{P}$ (theorem A.3.17). This algebra of bounded Skorohod-continuous functions separates the points of $\mathscr{D}_E$, so $\mathbb{P}$ has a unique extension to a tight probability on the Borels of the Skorohod topology τ (proposition A.6.2). ∎

Proposition A.7.2 *A subset $\mathscr{K} \subset \mathscr{D}_E$ is relatively compact if and only if both (i) $\{z_t : z \in \mathscr{K}\}$ is relatively compact in E for every $t \in [0, \infty)$ and (ii) for every instant $u \in [0, \infty)$, $\lim_{\delta\to 0} \gamma^\delta_{[)}[z^u] = 0$ uniformly in $z \in \mathscr{K}$. In this case the sets $\{z_t : 0 \le t \le u, z \in \mathscr{K}\}$, $u < \infty$, are relatively compact in E.*
For a proof see for example Ethier and Kurtz [34, page 123].

Proposition A.7.3 *A family $\mathcal{M}$ of probabilities on $\mathscr{D}_E$ is uniformly tight provided that for every instant $u < \infty$ we have, uniformly in $\mu \in \mathcal{M}$,*

$$\int_{\mathscr{D}_E} \gamma^\delta_{[)}[z^u] \wedge 1\; \mu(dz) \xrightarrow[\delta\to 0]{} 0$$

or $$\sup\Big\{ \int_{\mathscr{D}_E} \rho(z_T^u, z_S^u) \wedge 1\; \mu(dz) : S, T \in \mathfrak{T},\ 0 \le S \le T \le S + \delta \Big\} \xrightarrow[\delta\to 0]{} 0\,.$$

Here $\mathfrak{T}$ denotes collection of stopping times for the right-continuous version of the basic filtration on $\mathscr{D}_E$.

44 supp $\boldsymbol{\phi}_n \in [s, s+1/n]$, $\phi_n^\eta \ge 0$, and $\int \phi_n^\eta(r)\, dr = 1$.
45 The set of threads is identified naturally with $\mathscr{D}_E$.

A.8 The L^p-Spaces

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and recall from page 33 the following measurements of the size of a measurable function f:

$$\lceil\!\lceil f \rceil\!\rceil_p = \lceil\!\lceil f \rceil\!\rceil_{L^p(\mathbb{P})} \stackrel{\text{def}}{=} \begin{cases} \|f\|_p = \left(\int |f|^p \, d\mathbb{P}\right)^{1/p} & \text{for } 1 \le p < \infty, \\ \|f\|_p^p = \int |f|^p \, d\mathbb{P} & \text{for } 0 < p \le 1, \\ \inf\left\{\lambda : \mathbb{P}\left[|f| > \lambda\right] \le \lambda\right\} & \text{for } p = 0; \end{cases}$$

and

$$\|f\|_{[\alpha]} = \|f\|_{[\alpha;\mathbb{P}]} = \inf\{\lambda > 0 : \mathbb{P}[|f| > \lambda] \le \alpha\} \qquad \text{for } p = 0 \text{ and } \alpha > 0.$$

The space $L^p(\mathbb{P})$ is the collection of all measurable functions f that satisfy $\lceil\!\lceil rf \rceil\!\rceil_{L^p(\mathbb{P})} \xrightarrow[r \to 0]{} 0$. Customarily, the slew of L^p-spaces is extended at $p = \infty$ to include the space $L^\infty = L^\infty(\mathbb{P}) = L^\infty(\mathcal{F}, \mathbb{P})$ of bounded measurable functions equipped with the seminorm

$$\|f\|_\infty = \|f\|_{L^\infty(\mathbb{P})} \stackrel{\text{def}}{=} \inf\{c : \mathbb{P}[|f| > c] = 0\} ,$$

which we also write $\lceil\!\lceil \ \rceil\!\rceil_\infty$, if we want to stress its subadditivity. L^∞ plays a minor role in this book, since it is not suited to be the range of a vector measure such as the stochastic integral.

Exercise A.8.1 (i) $\lceil\!\lceil f \rceil\!\rceil_0 \le a \iff \mathbb{P}[|f| > a] \le a$.

(ii) For $1 \le p \le \infty$, $\lceil\!\lceil \ \rceil\!\rceil_p$ is a seminorm. For $0 \le p < 1$, it is subadditive but not homogeneous.

(iii) Let $0 \le p \le \infty$. A measurable function f is said to be ***finite in p-mean*** if

$$\lim_{r \to 0} \lceil\!\lceil rf \rceil\!\rceil_p = 0 .$$

For $0 < p \le \infty$ this means simply that $\lceil\!\lceil f \rceil\!\rceil_p < \infty$. A numerical measurable function f belongs to L^p if and only if it is finite in p-mean.

(iv) Let $0 \le p \le \infty$. The spaces L^p are vector lattices, i.e., are closed under taking finite linear combinations and finite pointwise maxima and minima. They are not in general algebras, except for L^0, which is one. They are complete under the metric $dist_p(f, g) = \lceil\!\lceil f - g \rceil\!\rceil_p$, and every mean-convergent sequence has an almost surely convergent subsequence.

(v) Let $0 \le p < \infty$. The simple measurable functions are p-mean dense. (A measurable function is simple if it takes only finitely many values, all of them finite.)

Exercise A.8.2 For $0 < p < 1$, the homogeneous functionals $\| \ \|_p$ are not subadditive, but there is a substitute for subadditivity:

$$\|f_1 + \ldots + f_n\|_p \le n^{0 \vee (1-p)/p} \cdot \left(\|f_1\|_p + \ldots + \|f_n\|_p\right) \qquad 0 < p \le \infty .$$

Exercise A.8.3 For any set K and $p \in [0, \infty]$, $\lceil\!\lceil K \rceil\!\rceil_{L^p(\mathbb{P})} = (\mathbb{P}[K])^{1 \wedge 1/p}$.

Theorem A.8.4 (Hölder's Inequality) *For any three exponents $0 < p, q, r \le \infty$ with $1/r = 1/p + 1/q$ and any two measurable functions f, g,*

$$\| fg \|_r \le \| f \|_p \cdot \| g \|_q \,.$$

If $p, p' \in [1, \infty]$ are related by $1/p + 1/p' = 1$, then p, p' are called ***conjugate exponents****. Then $p' = p/(p-1)$ and $p = p'/(p'-1)$. For conjugate exponents p, p'*

$$\| f \|_p = \sup\{ \int fg : g \in L^{p'} \,, \| g \|_{p'} \le 1 \} \,.$$

All of this remains true if the underlying measure is merely σ-finite.[46]

Proof: Exercise. ∎

Exercise A.8.5 Let μ be a positive σ-additive measure and f a μ-measurable function. The set $I_f \stackrel{\text{def}}{=} \{1/p \in \mathbb{R}_+ : \| f \|_p < \infty\}$ is either empty or is an interval. The function $1/p \mapsto \| f \|_p$ is logarithmically convex and therefore continuous on the interior $\mathring{I}_f$. Consequently, for $0 < p_0 < p_1 < \infty$,

$$\sup_{p_0 < p < p_1} \| f \|_p = \| f \|_{p_0} \vee \| f \|_{p_1} \,.$$

Uniform Integrability Let $0 < p < \infty$ and let μ be a positive measure. A collection $\mathcal{C}$ of L^p-integrable functions is ***uniformly p-integrable*** if for every $\epsilon > 0$ there is a function $g_\epsilon \in L^p(\mu)$ such that for every $f \in \mathcal{C}$

the distance $\quad dist(f, [-g_\epsilon, g_\epsilon]) \stackrel{\text{def}}{=} \inf\{\|f - g\|_{L^p(\mu)} : -g_\epsilon \le g \le g_\epsilon\}$

of f from the

order interval $\quad [-g_\epsilon, g_\epsilon] \stackrel{\text{def}}{=} \{g \in L^p(\mu) : -g_\epsilon \le g \le g_\epsilon\}$

is less than ϵ. In this case the infimum above is minimized at the member $-g_\epsilon \vee f \wedge g_\epsilon$ of $[-g_\epsilon, g_\epsilon]$. Uniform integrability generalizes the domination condition in the DCT in this sense: if there is a $g \in L^p$ with $|f| \le g$ for all $f \in \mathcal{C}$, then $\mathcal{C}$ is uniformly p-integrable. Using this notion one can establish the most general and sharp convergence theorem of Lebesgue's theory – it makes a good exercise:

Theorem A.8.6 (Dominated Convergence Theorem on L^p) *Let $\mathbb{P}$ be a probability and let $0 < p < \infty$. A sequence (f_n) in $L^p(\mathbb{P})$ converges in p-mean if and only if it is uniformly p-integrable and converges in measure.*

Exercise A.8.7 (Fatou's Lemma) (i) Let (f_n) be a sequence of positive measurable functions. Then

$$\Big\|\liminf_{n\to\infty} f_n\Big\|_{L^p} \le \liminf_{n\to\infty} \|f_n\|_{L^p} \,, \qquad 0 < p \le \infty;$$

$$\Big[\!\!\Big[\liminf_{n\to\infty} f_n\Big]\!\!\Big]_{L^p} \le \liminf_{n\to\infty} [\![f_n]\!]_{L^p} \,, \qquad 0 \le p < \infty;$$

$$\Big\|\liminf_{n\to\infty} f_n\Big\|_{[\alpha]} \le \liminf_{n\to\infty} \|f_n\|_{[\alpha]} \,, \qquad 0 < \alpha < \infty.$$

---

[46] I.e., $L^1(\mu)$ is σ-finite (exercise A.3.2).

(ii) Let (f_n) be a sequence in L^0 that converges in measure to f. Then

$$\begin{aligned} \|f\|_{L^p} &\le \liminf_{n\to\infty} \|f_n\|_{L^p} , && 0 < p \le \infty; \\ \lceil\!\lceil f \rceil\!\rceil_{L^p} &\le \liminf_{n\to\infty} \lceil\!\lceil f_n \rceil\!\rceil_{L^p} , && 0 \le p < \infty; \\ \|f\|_{[\alpha]} &\le \liminf_{n\to\infty} \|f_n\|_{[\alpha]} , && 0 < \alpha < \infty. \end{aligned}$$

A.8.8 Convergence in Measure — the Space L^0 The space L^0 of almost surely finite functions plays a large role in this book. Exercise A.8.10 amounts to saying that L^0 is a topological vector space for the topology of convergence in measure, whose neighborhood system at 0 has a countable basis, and which is defined by the gauge $\lceil\!\lceil\ \rceil\!\rceil_0$. Here are a few exercises that are used in the main text.

Exercise A.8.9 The functional $\lceil\!\lceil\ \rceil\!\rceil_0$ is by no means the canonical way with which to gauge the size of a.s. finite measurable functions. Here is another one that serves just as well and is sometimes easier to handle: $\lceil\!\lceil f \rceil\!\rceil_{\odot} \stackrel{\text{def}}{=} \mathbb{E}[|f| \wedge 1]$, with associated distance

$$dist_{\odot}(f,g) \stackrel{\text{def}}{=} \lceil\!\lceil f-g \rceil\!\rceil_{\odot} = \mathbb{E}[|f-g| \wedge 1] .$$

Show: (i) $\lceil\!\lceil\ \rceil\!\rceil_{\odot}$ is subadditive, and $dist_{\odot}$ is a pseudometric. (ii) A sequence (f_n) in L^0 converges to $f \in L^0$ in measure if and only if $\lceil\!\lceil f - f_n \rceil\!\rceil_{\odot} \xrightarrow[n\to\infty]{} 0$. In other words, $\lceil\!\lceil\ \rceil\!\rceil_{\odot}$ is a gauge on $L^0(\mathbb{P})$.

Exercise A.8.10 L^0 is an algebra. The maps $(f,g) \mapsto f+g$ and $(f,g) \mapsto f \cdot g$ from $L^0 \times L^0$ to L^0 and $(r,f) \mapsto r \cdot f$ from $\mathbb{R} \times L^0$ to L^0 are continuous. The neighborhood system at 0 has a basis of "balls" of the form

$$B_r(0) \stackrel{\text{def}}{=} \{f : \lceil\!\lceil f \rceil\!\rceil_0 < r\} \text{ and of the form } B_r(0) \stackrel{\text{def}}{=} \{f : \lceil\!\lceil f \rceil\!\rceil_{\odot} < r\} , \quad r > 0 .$$

Exercise A.8.11 Here is another gauge on $L^0(\mathbb{P})$, which is used in proposition 3.6.20 on page 129 to study the behavior of the stochastic integral under a change of measure. Let $\mathbb{P}'$ be a probability equivalent to $\mathbb{P}$ on $\mathcal{F}$, i.e., a probability having the same negligible sets as $\mathbb{P}$. The Radon–Nikodym theorem A.3.22 on page 407 provides a strictly positive $\mathcal{F}$-measurable function g' such that $\mathbb{P}' = g'\mathbb{P}$. Show that the spaces $L^0(\mathbb{P})$ and $L^0(\mathbb{P}')$ coincide; moreover, the topologies of convergence in $\mathbb{P}$-measure and in $\mathbb{P}'$-measure coincide as well.

The mean $\lceil\!\lceil\ \rceil\!\rceil_{L^0(\mathbb{P}')}$ thus describes the topology of convergence in $\mathbb{P}$-measure just as satisfactorily as does $\lceil\!\lceil\ \rceil\!\rceil_{L^0(\mathbb{P})}$: $\lceil\!\lceil\ \rceil\!\rceil_{L^0(\mathbb{P}')}$ is a gauge on $L^0(\mathbb{P})$.

Exercise A.8.12 Let $(\Omega, \mathcal{F})$ be a measurable space and $\mathbb{P} \ll \mathbb{P}'$ two probabilities on $\mathcal{F}$. There exists an increasing right-continuous function $\Phi : (0,1] \to (0,1]$ with $\lim_{r\to 0} \Phi(r) = 0$ such that $\lceil\!\lceil f \rceil\!\rceil_{L^0(\mathbb{P})} \le \Phi(\lceil\!\lceil f \rceil\!\rceil_{L^0(\mathbb{P}')})$ for all $f \in \mathcal{F}$.

A.8.13 The Homogeneous Gauges $\|\ \|_{[\alpha]}$ on $\mathbf{L^p}$ Recall the definition of the homogeneous gauges on measurable functions f,

$$\|f\|_{[\alpha]} = \|f\|_{[\alpha;\mathbb{P}]} \stackrel{\text{def}}{=} \inf\{\lambda > 0 : \mathbb{P}[|f| > \lambda] \le \alpha\} .$$

Of course, if $\alpha < 0$, then $\|f\|_{[\alpha]} = \infty$, and if $\alpha \ge 1$, then $\|f\|_{[\alpha]} = 0$. Yet it streamlines some arguments a little to make this definition for all real α.

Exercise A.8.14 (i) $\|r\cdot f\|_{[\alpha]} = |r|\cdot\|f\|_{[\alpha]}$ for any measurable f and any $r\in\mathbb{R}$.

(ii) For any measurable function f and any $\alpha>0$, $\lceil\!\lceil f \rceil\!\rceil_{L^0} < \alpha \iff \|f\|_{[\alpha]} < \alpha$, $\|f\|_{[\alpha]} \le \lambda \iff \mathbb{P}[|f|>\lambda] \le \alpha$, and $\lceil\!\lceil f \rceil\!\rceil_{L^0} = \inf\{\alpha : \|f\|_{[\alpha]} \le \alpha\}$.

(iii) A sequence (f_n) of measurable functions converges to f in measure if and only if $\|f_n - f\|_{[\alpha]} \xrightarrow[n\to\infty]{} 0$ for all $\alpha>0$, i.e., iff $\mathbb{P}[|f-f_n|>\alpha] \xrightarrow[n\to\infty]{} 0 \quad \forall\,\alpha>0$.

(iv) The function $\alpha\mapsto\|f\|_{[\alpha]}$ is decreasing and right-continuous. Considered as a measurable function on the Lebesgue space $(0,1)$ it has the same distribution as $|f|$. It is thus often called the ***non-increasing rearrangement*** of $|f|$.

Exercise A.8.15 (i) Let f be a measurable function. Then for $0<p<\infty$

$$\big\||f|^p\big\|_{[\alpha]} = \Big(\|f\|_{[\alpha]}\Big)^p\,,\ \ \|f\|_{[\alpha]} \le \alpha^{-1/p}\cdot\|f\|_{L^p}\ ,$$

and
$$\mathbb{E}\big[|f|^p\big] = \int_0^1 \|f\|^p_{[\alpha]}\,d\alpha\ .$$

In fact, for any continuous Φ on $\mathbb{R}_+$, $\int\Phi(|f|)\,d\mathbb{P} = \int_0^1 \Phi(\|f\|_{[\alpha]})\,d\alpha$.

(ii) $f\mapsto\|f\|_{[\alpha]}$ is not subadditive, but there is a substitute:

$$\|f+g\|_{[\alpha+\beta]} \le \|f\|_{[\alpha]} + \|g\|_{[\beta]}\ .$$

Exercise A.8.16 In the proofs of 2.3.3 and theorems 3.1.6 and 4.1.12 a "Fubini–type estimate" for the gauges $\|\ \ \|_{[\alpha]}$ is needed. It is this: let $\mathbb{P},\tau$ be probabilities, and $f(\omega,t)$ a $\mathbb{P}\times\tau$-measurable function. Then for $\alpha,\beta,\gamma>0$

$$\Big\|\,\|f\|_{[\beta;\tau]}\Big\|_{[\alpha;\mathbb{P}]} \le \Big\|\,\|f\|_{[\gamma;\mathbb{P}]}\Big\|_{[\alpha\beta-\gamma;\tau]}\ .$$

Exercise A.8.17 Suppose f,g are positive random variables with $\|f\|_{L^r(\mathbb{P}/g)} \le E$, where $r>0$. Then

$$\|f\|_{[\alpha+\beta]} \le E\cdot\left(\frac{\|g\|_{[\alpha]}}{\beta}\right)^{1/r}\ ,\ \|f\|_{[\alpha]} \le E\cdot\left(\frac{2\,\|g\|_{[\alpha/2]}}{\alpha}\right)^{1/r} \tag{A.8.1}$$

and
$$\|fg\|_{[\alpha]} \le E\cdot\|g\|_{[\alpha/2]}\cdot\left(\frac{2\|g\|_{[\alpha/2]}}{\alpha}\right)^{1/r}\ . \tag{A.8.2}$$

Bounded Subsets of L^p Recall from page 379 that a subset $\mathcal{C}$ of a topological vector space $\mathcal{V}$ is ***bounded*** if it can be absorbed by any neighborhood V of zero; that is to say, if for any neighborhood V of zero there is a scalar r such that $\mathcal{C}\subset r\cdot V$.

Exercise A.8.18 (Cf. A.2.28) Let $0\le p\le\infty$. A set $\mathcal{C}\subset L^p$ is bounded if and only if

$$\sup\{\lceil\!\lceil \lambda\cdot f \rceil\!\rceil_p : f\in\mathcal{C}\} \xrightarrow[\lambda\to 0]{} 0\ , \tag{A.8.3}$$

which is the same as saying $\sup\{\lceil\!\lceil f \rceil\!\rceil_p : f\in\mathcal{C}\} < \infty$ in the case that p is strictly positive. If $p=0$, then the previous supremum is always less than or equal to 1 and equation (A.8.3) describes boundedness. Namely, for $\mathcal{C}\subset L^0(\mathbb{P})$, the following are equivalent: (i) $\mathcal{C}$ is bounded in $L^0(\mathbb{P})$; (ii) $\sup\{\lceil\!\lceil \lambda\cdot f \rceil\!\rceil_{L^0(\mathbb{P})} : f\in\mathcal{C}\} \xrightarrow[\lambda\to 0]{} 0$; (iii) for every $\alpha>0$ there exists $C_\alpha<\infty$ such that

$$\|f\|_{[\alpha;\mathbb{P}]} \le C_\alpha \quad \forall\, f\in\mathcal{C}\ .$$

Exercise A.8.19 Let $\mathbb{P}' \ll \mathbb{P}$. Then the natural injection of $L^0(\mathbb{P})$ into $L^0(\mathbb{P}')$ is continuous and thus maps bounded sets of $L^0(\mathbb{P})$ into bounded sets of $L^0(\mathbb{P}')$.

The elementary stochastic integral is a linear map from a space $\mathcal{E}$ of functions to one of the spaces L^p. It is well to study the continuity of such a map. Since both the domain and range are metric spaces, continuity and boundedness coincide – recall that a linear map is ***bounded*** if it maps bounded sets of its domain to bounded sets of its range.

Exercise A.8.20 Let $\mathcal{I}$ be a linear map from the normed linear space $(\mathcal{E}, \|\ \|_{\mathcal{E}})$ to $L^p(\mathbb{P})$, $0 \le p \le \infty$. The following are equivalent: (i) $\mathcal{I}$ is continuous. (ii) $\mathcal{I}$ is continuous at zero. (iii) $\mathcal{I}$ is bounded.

$$\text{(iv)} \qquad \sup\Big\{ \lceil\!\lceil \mathcal{I}(\lambda\cdot\phi) \rceil\!\rceil_p :\ \phi \in \mathcal{E},\ \|\phi\|_{\mathcal{E}} \le 1 \Big\} \xrightarrow[\lambda \to 0]{} 0\ .$$

If $p = 0$, then $\mathcal{I}$ is continuous if and only if for every $\alpha > 0$ the number

$$\|\mathcal{I}\|_{[\alpha;\mathbb{P}]} \stackrel{\text{def}}{=} \sup\Big\{ \|\mathcal{I}(X)\|_{[\alpha;\mathbb{P}]} :\ X \in \mathcal{E}\ ,\ \|X\|_{\mathcal{E}} \le 1 \Big\} \quad \text{is finite.}$$

A.8.21 Occasionally one wishes to estimate the size of a function or set without worrying about its measurability. In this case one argues with the upper integral $\int^*$ or the outer measure $\mathbb{P}^*$. The corresponding constructs for $\|\ \|_p$, $\lceil\!\lceil\ \rceil\!\rceil_p$, and $\|\ \|_{[\alpha]}$ are

$$\left.\begin{aligned} \|f\|_p^* = \|f\|_{L^p(\mathbb{P})}^* &\stackrel{\text{def}}{=} \Big(\int^* |f|^p \, d\mathbb{P}\Big)^{1/p} \\ \lceil\!\lceil f \rceil\!\rceil_p^* = \lceil\!\lceil f \rceil\!\rceil_{L^p(\mathbb{P})}^* &\stackrel{\text{def}}{=} \Big(\int^* |f|^p \, d\mathbb{P}\Big)^{1\wedge 1/p} \end{aligned}\right\} \quad \text{for } 0 < p < \infty\ ,$$

$$\left.\begin{aligned} \lceil\!\lceil f \rceil\!\rceil_0^* = \lceil\!\lceil f \rceil\!\rceil_{L^0(\mathbb{P})} &\stackrel{\text{def}}{=} \inf\{\lambda :\ \mathbb{P}^*[|f| \ge \lambda] \le \lambda\} \\ \|f\|_{[\alpha]}^* = \|f\|_{[\alpha;\mathbb{P}]}^* &\stackrel{\text{def}}{=} \inf\{\lambda :\ \mathbb{P}^*[|f| \ge \lambda] \le \alpha\} \end{aligned}\right\} \quad \text{for } p = 0\ .$$

Exercise A.8.22 It is well known that $\int^*$ is ***continuous along arbitrary increasing sequences***:

$$0 \le f_n \uparrow f \implies \int^* f_n \uparrow \int^* f\ .$$

Show that the starred constructs above all share this property.

Exercise A.8.23 Set $\|f\|_{1,\infty} = \|f\|_{L^{1,\infty}(\mathbb{P})} = \sup_{\lambda>0} \lambda \cdot \mathbb{P}[|f| > \lambda]$. Then

$$\|f\|_p \le p\Big(\frac{2-p}{1-p}\Big)^{1/p} \cdot \|f\|_{1,\infty} \qquad \text{for } 0 < p < 1\ ,$$

$$\|f+g\|_{1,\infty} \le 2 \cdot \Big(\|f\|_{1,\infty} + \|g\|_{1,\infty}\Big)\ ,$$

and

$$\|rf\|_{1,\infty} = |r| \cdot \|f\|_{1,\infty}\ .$$

Marcinkiewicz Interpolation

Interpolation is a large field, and we shall establish here only the one rather elementary result that is needed in the proof of theorem 2.5.30 on page 85. Let $U : L^\infty(\mathbb{P}) \to L^\infty(\mathbb{P})$ be a subadditive map and $1 \le p \le \infty$. U is said to be of ***strong type*** $p{-}p$ if there is a constant A_p with

$$\|U(f)\|_p \le A_p \cdot \|f\|_p ,$$

in other words if it is continuous from $L^p(\mathbb{P})$ to $L^p(\mathbb{P})$. It is said to be of ***weak type*** $p{-}p$ if there is a constant A'_p such that

$$\mathbb{P}[|U(f)| \ge \lambda] \le \Big(\frac{A'_p}{\lambda} \cdot \|f\|_p\Big)^p .$$

"Weak type $\infty{-}\infty$" is to mean "strong type $\infty{-}\infty$." Chebyscheff's inequality shows immediately that a map of strong type $p{-}p$ is of weak type $p{-}p$.

Proposition A.8.24 (An Interpolation Theorem) *If U is of weak types $p_1{-}p_1$ and $p_2{-}p_2$ with constants A'_{p_1}, A'_{p_2}, respectively, then it is of strong type $p{-}p$ for $p_1 < p < p_2$:*

$$\|U(f)\|_p \le A_p \cdot \|f\|_p$$

with constant
$$A_p \le p^{1/p} \cdot \Big(\frac{(2A'_{p_1})^{p_1}}{p-p_1} + \frac{(2A'_{p_2})^{p_2}}{p_2-p}\Big)^{1/p} .$$

Proof. By the subadditivity of U we have for every $\lambda > 0$

$$|U(f)| \le |U(f \cdot [|f| \ge \lambda])| + |U(f \cdot [|f| < \lambda])| ,$$

and consequently

$$\begin{aligned}\mathbb{P}[|U(f)| \ge \lambda] &\le \mathbb{P}[|U(f \cdot [|f| \ge \lambda])| \ge \lambda/2] + \mathbb{P}[|U(f \cdot [|f| < \lambda])| \ge \lambda/2] \\ &\le \Big(\frac{A'_{p_1}}{\lambda/2}\Big)^{p_1} \int_{[|f|\ge\lambda]} |f|^{p_1}\, d\mathbb{P} + \Big(\frac{A'_{p_2}}{\lambda/2}\Big)^{p_2} \int_{[|f|<\lambda]} |f|^{p_2}\, d\mathbb{P} .\end{aligned}$$

We multiply this with λ^{p-1} and integrate against $d\lambda$, using Fubini's theorem A.3.18:

$$\begin{aligned}
\frac{\mathbb{E}(|U(f)|^p)}{p} &= \int\int_0^{|U(f)|} \lambda^{p-1}\, d\lambda d\mathbb{P} = \int\int_0^{\infty} \mathbb{P}[|U(f)| \ge \lambda]\lambda^{p-1}\, d\lambda \\
&\le \int \Big(\frac{A'_{p_1}}{\lambda/2}\Big)^{p_1} \int_{[|f|\ge\lambda]} |f|^{p_1} \lambda^{p-1}\, d\lambda d\mathbb{P} \\
&\quad + \int \Big(\frac{A'_{p_2}}{\lambda/2}\Big)^{p_2} \int_{[|f|<\lambda]} |f|^{p_2} \lambda^{p-1}\, d\lambda d\mathbb{P} \\
&= \big(2A'_{p_1}\big)^{p_1} \int\int_0^{|f|} |f|^{p_1}\lambda^{p-p_1-1}\, d\lambda d\mathbb{P} \\
&\quad + \big(2A'_{p_2}\big)^{p_2} \int\int_{|f|}^{\infty} |f|^{p_2}\lambda^{p-p_2-1}\, d\lambda d\mathbb{P} \\
&= \frac{(2A'_{p_1})^{p_1}}{p-p_1} \int |f|^{p_1}|f|^{p-p_1}\, d\mathbb{P} + \frac{(2A'_{p_2})^{p_2}}{p_2-p} \int |f|^{p_2}|f|^{p-p_2}\, d\mathbb{P} \\
&= \Big(\frac{(2A'_{p_1})^{p_1}}{p-p_1} + \frac{(2A'_{p_2})^{p_2}}{p_2-p}\Big) \cdot \int |f|^p\, d\mathbb{P}\ .
\end{aligned}$$

We multiply both sides by p and take p^{th} roots; the claim follows. ∎

Note that the constant A_p blows up as $p \downarrow p_1$ or $p \uparrow p_2$. In the one application we make, in the proof of theorem 2.5.30, the map U happens to be ***self-adjoint***, which means that

$$\mathbb{E}[U(f)\cdot g] = \mathbb{E}[f \cdot U(g)]\ , \qquad f, g \in L^2\ .$$

Corollary A.8.25 *If U is self-adjoint and of weak types $1{-}1$ and $2{-}2$, then U is of strong type $p{-}p$ for all $p \in (1,\infty)$.*

Proof. Let $2 < p < \infty$. The conjugate exponent $p' = p/(p-1)$ then lies in the open interval $(1,2)$, and by proposition A.8.24

$$\mathbb{E}[U(f)\cdot g] = \mathbb{E}[f\cdot U(g)] \le \|f\|_p \cdot \|U(g)\|_{p'} \le \|f\|_p \cdot A_{p'}\|g\|_{p'}\ .$$

We take the supremum over g with $\|g\|_{p'} \le 1$ and arrive at

$$\|U(f)\|_p \le A_{p'} \cdot \|f\|_p\ .$$

The claim is satisfied with $A_p = A_{p'}$. Note that, by interpolating now between $3/2$ and 3, say, the estimate of the constants A_p can be improved so no pole at $p = 2$ appears. (It suffices to assume that U is of weak types $1{-}1$ and $(1{+}\epsilon) - (1{+}\epsilon)$). ∎

Khintchine's Inequalities

Let T be the product of countably many two-point sets $\{1,-1\}$. Its elements are sequences $t=(t_1,t_2,\ldots)$ with $t_\nu=\pm1$. Let $\epsilon_\nu : t\mapsto t_\nu$ be the ν^{th} coordinate function. If we equip T with the product τ of uniform measure on $\{1,-1\}$, then the ϵ_ν are independent and identically distributed Bernoulli variables that take the values ±1 with τ-probability $1/2$ each. The ϵ_ν form an orthonormal set in $L^2(\tau)$ that is far from being a basis; in fact, they form so sparse or *lacunary* a collection that on their linear span

$$\mathcal{R}=\Big\{\sum_{\nu=1}^{n}a_\nu\epsilon_\nu :\ n\in\mathbb{N},\ a_\nu\in\mathbb{R}\Big\}$$

all of the $L^p(\tau)$-topologies coincide:

Theorem A.8.26 (Khintchine) *Let $0<p<\infty$. There are universal constants k_p, K_p such that for any natural number n and reals $a_1,\ldots,a_n$*

$$\Big\|\sum_{\nu=1}^{n}a_\nu\epsilon_\nu\Big\|_{L^p(\tau)}\le k_p\cdot\Big(\sum_{\nu=1}^{n}a_\nu^2\Big)^{1/2} \tag{A.8.4}$$

and

$$\Big(\sum_{\nu=1}^{n}a_\nu^2\Big)^{1/2}\le K_p\cdot\Big\|\sum_{\nu=1}^{n}a_\nu\epsilon_\nu\Big\|_{L^p(\tau)} . \tag{A.8.5}$$

For $p=0$, there are universal constants $\kappa_0>0$ and $K_0<\infty$ such that

$$\Big(\sum_{\nu=1}^{n}a_\nu^2\Big)^{1/2}\le K_0\cdot\Big\|\sum_{\nu=1}^{n}a_\nu\epsilon_\nu\Big\|_{[\kappa_0;\tau]} . \tag{A.8.6}$$

In particular, a subset of $\mathcal{R}$ that is bounded in the topology induced by any of the $L^p(\tau)$, $0\le p<\infty$, is bounded in $L^2(\tau)$. The completions of $\mathcal{R}$ in these various topologies being the same, the inequalities above stay if the finite sums are replaced with infinite ones. Bounds for the universal constants k_p, K_p, K_0, κ_0 are discussed in remark A.8.28 and exercise A.8.29 below.

Proof. Let us write $\|f\|_p$ for $\|f\|_{L^p(\tau)}$. Note that for $f=\sum_{\nu=1}^{n}a_\nu\epsilon_\nu\in\mathcal{R}$

$$\|f\|_2=\Big(\sum_\nu a_\nu^2\Big)^{1/2} .$$

In proving (A.8.4)–(A.8.6) we may by homogeneity assume that $\|f\|_2=1$. Let $\lambda>0$. Since the ϵ_ν are independent and $\cosh x\le e^{x^2/2}$ (as a term–by–term comparison of the power series shows),

$$\int e^{\lambda f(t)}\,\tau(dt)=\prod_{\nu=1}^{N}\cosh(\lambda a_\nu)\le\prod_{\nu=1}^{N}e^{\lambda^2a_\nu^2/2}=e^{\lambda^2/2} ,$$

and consequently

$$\int e^{\lambda|f(t)|}\,\tau(dt) \le \int e^{\lambda f(t)}\,\tau(dt) + \int e^{-\lambda f(t)}\,\tau(dt) \le 2e^{\lambda^2/2}\ .$$

We apply Chebysheff's inequality to this and obtain

$$\tau([|f| \ge \lambda]) = \tau\Big(\Big[e^{\lambda|f|} \ge e^{\lambda^2}\Big]\Big) \le e^{-\lambda^2}\cdot 2e^{\lambda^2/2} = 2e^{-\lambda^2/2}\ .$$

Therefore, if $p \ge 2$, then

$$\begin{aligned}
\int |f(t)|^p\,\tau(dt) &= p\cdot\int_0^\infty \lambda^{p-1}\,\tau([|f|\ge\lambda])\,d\lambda \\
&\le 2p\cdot\int_0^\infty \lambda^{p-1}e^{-\lambda^2/2}\,d\lambda \\
\text{with } z=\lambda^2/2: \quad &= 2p\cdot\int_0^\infty (2z)^{(p-2)/2}e^{-z}\,dz \\
&= 2^{\frac{p}{2}+1}\cdot\frac{p}{2}\Gamma(\frac{p}{2}) = 2^{\frac{p}{2}+1}\Gamma(\frac{p}{2}+1)\ .
\end{aligned}$$

We take p^{th} roots and arrive at (A.8.4) with

$$k_p = \begin{cases} 2^{\frac{1}{p}+\frac{1}{2}}\big(\Gamma(\frac{p}{2}+1)\big)^{1/p} \le \sqrt{2p} & \text{if } p \ge 2 \\ 1 & \text{if } p < 2. \end{cases}$$

As to inequality (A.8.5), which is only interesting if $0 < p < 2$, write

$$\begin{aligned}
\|f\|_2^2 = \int |f|^2(t)\,\tau(dt) &= \int |f|^{p/2}(t)\cdot|f|^{2-p/2}(t)\,\tau(dt) \\
\text{using Hölder:} \quad &\le \Big(\int |f|^p(t)\,\tau(dt)\Big)^{1/2}\cdot\Big(\int |f|^{4-p}(t)\,\tau(dt)\Big)^{1/2} \\
&= \|f\|_p^{p/2}\cdot\|f\|_{4-p}^{2-p/2} \le \|f\|_p^{p/2}\cdot k_{4-p}^{2-\,p/2}\|f\|_2^{2-\,p/2}\ ,
\end{aligned}$$

and thus $\|f\|_2^{p/2} \le k_{4-p}^{2-\,p/2}\cdot\|f\|_p^{p/2}$ and $\|f\|_2 \le k_{4-p}^{4/p\,-1}\cdot\|f\|_p$.

The estimate

$$k_{4-p} \le 2\cdot\Big(\Gamma\Big(\frac{4-p}{2}+1\Big)\Big)^{1/(4-p)} \le 2\cdot\big(\Gamma(3)\big)^{1/4} = 2^{5/4}$$

leads to (A.8.5) with $K_p \le 2^{5/p\,-5/4}$. In summary:

$$K_p \le \begin{cases} 1 & \text{if } p \ge 2, \\ 2^{5/p\,-5/4} < 2^{5/p} & \text{if } 0 < p < 2 \\ \le 16 & \text{if } 1 \le p < \infty. \end{cases}$$

Finally let us prove inequality (A.8.6). From

$$\begin{aligned}\|f\|_1 = \int |f(t)|\,\tau(dt) &= \int_{[|f|\ge\|f\|_1/2]} |f(t)|\,\tau(dt) \;+\; \int_{[|f|<\|f\|_1/2]} |f(t)|\,\tau(dt)\\ &\le \|f\|_2\cdot\Big(\tau\big[|f|\ge\|f\|_1/2\big]\Big)^{1/2} + \|f\|_1/2\\ &\le K_1\cdot\|f\|_1\cdot\Big(\tau\big[|f|\ge\|f\|_1/2\big]\Big)^{1/2} + \|f\|_1/2\end{aligned}$$

we deduce that $$1/2 \le K_1\cdot\Big(\tau\big[|f|\ge\|f\|_1/2\big]\Big)^{1/2}$$

and thus $$\frac{1}{(2K_1)^2} \le \tau\big([|f|\ge\|f\|_1/2]\big) \le \tau\Big[|f|\ge\frac{\|f\|_2}{2K_1}\Big]\,. \tag{$*$}$$

Recalling that $$\|f\|_{[\alpha;\tau]} = \inf\{\lambda : \tau[|f|\ge\lambda] < \alpha\}\,,$$

we rewrite $(*)$ as $$\|f\|_2 \le 2K_1\cdot\|f\|_{[1/(4K_1^2);\tau]}\,,$$

which is inequality (A.8.6) with $K_0 = 2K_1$ and $\kappa_0 = 1/(4K_1^2)$. ∎

Exercise A.8.27 $\|f\|_2 \le \big(\sqrt{1/2}-\sqrt{\kappa}\big)^{-1}\cdot\|f\|_{[\kappa]}$ for $0<\kappa<1/2$.

Remark A.8.28 Szarek [106] found the smallest possible value for K_1: $K_1=\sqrt{2}$. Haagerup [39] found the best possible constants k_p, K_p for all $p>0$:

$$k_p = \begin{cases} 1 & \text{for } 0<p\le 2,\\ \sqrt{2}\cdot\Big(\dfrac{\Gamma\big((p+1/2)\big)}{\sqrt{\pi}}\Big)^{1/p} & \text{for } 2\le p<\infty;\end{cases} \tag{A.8.7}$$

$$K_p = \begin{cases} 2^{-1/2}\cdot\Big(\dfrac{\sqrt{\pi}}{\Gamma\big((p+1)/2\big)}\vee 2\Big)^{1/p} & \text{for } 0<p\le 2,\\ 1 & \text{for } 2\le p<\infty.\end{cases} \tag{A.8.8}$$

In the text we shall use the following values to estimate the K_p and κ_0 (they are only slightly worse but rather simpler to read):

Exercise A.8.29

$$K_0^{(\text{A.8.6})} \le 2\sqrt{2} \;\text{ and }\; \kappa_0^{(\text{A.8.6})} \ge 1/8 \;\text{ for }\; p=0\,;$$
$$K_p^{(\text{A.8.5})} \le \begin{cases} 2^{1/p\,-\,1/2} & \text{for } 0<p\le 1.8\\ 1.00037\cdot 2^{1/p\,-\,1/2} & \text{for } 1.8<p\le 2\\ 1 & \text{for } 2\le p<\infty.\end{cases} \tag{A.8.9}$$

Exercise A.8.30 The values of k_p and K_p in (A.8.7) and (A.8.8) are best possible.

Exercise A.8.31 The *Rademacher functions* r_n are defined on the unit interval by

$$r_n(x) = \operatorname{sgn}\sin(2^n\pi x)\,, \qquad x\in(0,1),\ n=1,2,\ldots$$

The sequences (ϵ_ν) and (r_ν) have the same distribution.

Stable Type

In the proof of the general factorization theorem 4.1.2 on page 191, other special sequences of random variables are needed, sequences that show a behavior similar to that of the Rademacher functions, in the sense that on their linear span all L^p-topologies coincide. They are the sequences of independent symmetric q-stable random variables.

Symmetric Stable Laws Let $0 < q \le 2$. There exists a random variable $\gamma^{(q)}$ whose characteristic function is

$$\mathbb{E}\Big[e^{i\alpha\gamma^{(q)}}\Big] = e^{-|\alpha|^q} .$$

This can be proved using Bochner's theorem: $\alpha \mapsto |\alpha|^q$ is of negative type, so $\alpha \mapsto e^{-|\alpha|^q}$ is of positive type and thus is the characteristic function of a probability on $\mathbb{R}$. The random variable $\gamma^{(q)}$ is said to have a ***symmetric stable law of order q*** or to be ***symmetric q-stable***. For instance, $\gamma^{(2)}$ is evidently a centered normal random variable with variance 2; a Cauchy random variable, i.e., one that has density $1/\pi(1+x^2)$ on $(-\infty,\infty)$, is symmetric 1-stable. It has moments of orders $0 < p < 1$.

We derive the existence of $\gamma^{(q)}$ without appeal to Bochner's theorem in a computational way, since some estimates of their size are needed anyway. For our purposes it suffices to consider the case $0 < q \le 1$.

The function $\alpha \mapsto e^{-|\alpha|^q}$ is continuous and has very fast decay at $\pm\infty$, so its inverse Fourier transform

$$f^{(q)}(x) \stackrel{\text{def}}{=} \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ix\alpha}\cdot e^{-|\alpha|^q}\, d\alpha$$

is a smooth square integrable function with the right Fourier transform $\alpha \mapsto e^{-|\alpha|^q}$. It has to be shown, of course, that $f^{(q)}$ is positive, so that it qualifies as the probability density of a random variable – its integral is automatically 1, as it equals the value of its Fourier transform at $\alpha = 0$.

Lemma A.8.32 *Let $0 < q \le 1$. (i) The function $f^{(q)}$ is positive: it is indeed the density of a probability measure. (ii) If $0 < p < q$, then $\left|\gamma^{(q)}\right|$ has p^{th} moment*

$$\|\gamma^{(q)}\|_p^p = \Gamma\big((q-p)/q\big)\cdot b(p) , \tag{A.8.10}$$

where
$$b(p) \stackrel{\text{def}}{=} \frac{2}{\pi}\int_0^\infty \xi^{p-1}\sin\xi\, d\xi$$

is a strictly decreasing function with $\lim_{p\to 0} = 1$ *and*

$$2/\pi < (3-p)/\pi \le b(p) \le 1-(1{-}2/\pi)p \le 1 .$$

Proof (Charlie Friedman). For strictly positive x write

$$
\begin{aligned}
f^{(q)}(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ix\alpha}\cdot e^{-|\alpha|^q}\,d\alpha = \frac{1}{\pi}\int_0^{\infty}\cos(\alpha x)\cdot e^{-\alpha^q}\,d\alpha \\
&= \frac{1}{\pi}\sum_{k=0}^{\infty}\int_{k\pi/x}^{(k+1)\pi/x}\cos(\alpha x)\cdot e^{-\alpha^q}\,d\alpha \\
\text{with } \alpha x = \xi + k\pi: \quad &= \frac{1}{\pi x}\sum_{k=0}^{\infty}\int_0^{\pi}\cos(\xi+k\pi)\cdot e^{-\left(\frac{\xi+k\pi}{x}\right)^q}\,d\xi \\
&= \frac{1}{\pi x}\sum_{k=0}^{\infty}(-1)^k\int_0^{\pi}\cos\xi\cdot e^{-\left(\frac{\xi+k\pi}{x}\right)^q}\,d\xi \\
\text{integrating by parts:} \quad &= \sum_{k=0}^{\infty}\frac{q(-1)^k}{\pi x^{1+q}}\int_0^{\pi}\sin\xi\cdot e^{-\left(\frac{\xi+k\pi}{x}\right)^q}\cdot(\xi+k\pi)^{q-1}\,d\xi \\
&= \frac{q}{\pi x^{1+q}}\int_0^{\infty}\sin\xi\cdot e^{-(\frac{\xi}{x})^q}\cdot\xi^{q-1}\,d\xi\ .
\end{aligned}
$$

The penultimate line represents $f^{(q)}(x)$ as the sum of an alternating series. In the present case $0 < q \le 1$ it is evident that $e^{-((\xi+k\pi)/x)^q}\,(\xi+k\pi)^{q-1}$ decreases as k increases. The series' terms decrease in absolute value, so the sum converges. Since the first term is positive, so is the sum: $f^{(q)}$ is indeed a probability density.

(ii) To prove the existence of p^{th} moments write $\|\gamma^{(q)}\|_p^p$ as

$$
\begin{aligned}
\int_{-\infty}^{\infty}|x|^p\, f^{(q)}(x)\,dx &= 2\int_0^{\infty}x^p\, f^{(q)}(x)\,dx \\
&= \frac{2q}{\pi}\int_0^{\infty}\int_0^{\infty}x^{p-1-q}\sin\xi\, e^{-(\xi/x)^q}\,\xi^{q-1}\,dx\,d\xi \\
\text{with } (\xi/x)^q = y: \quad &= \frac{2}{\pi}\int_0^{\infty}\int_0^{\infty}y^{-p/q}e^{-y}\xi^{p-1}\sin\xi\,dy\,d\xi \\
&= \Gamma\big((q-p)/q\big)\cdot b(p)\ .
\end{aligned}
$$

The estimates of $b(p)$ are left as an exercise. ∎

Lemma A.8.33 *Let $0 < q \le 2$, let $\big(\gamma_1^{(q)}, \gamma_2^{(q)}, \ldots\big)$ be a sequence of independent symmetric q-stable random variables defined on a probability space $(X, \mathcal{X}, dx)$, and let $c_1, c_2, \ldots$ be a sequence of reals. (i) For any $p \in (0, q)$*

$$
\Big\|\sum_{\nu=1}^{\infty}c_\nu\,\gamma_\nu^{(q)}\Big\|_{L^p(dx)} = \|\gamma^{(q)}\|_p\cdot\Big(\sum_{\nu=1}^{\infty}|c_\nu|^q\Big)^{1/q}. \tag{A.8.11}
$$

In other words, the map $\ell^q \ni (c_\nu) \mapsto \sum_\nu c_\nu\gamma_\nu^{(q)}$ is, up to the factor $\|\gamma^{(q)}\|_p$, an isometry from ℓ^q into $L^p(dx)$.

(ii) This map is also a homeomorphism from ℓ^q into $L^0(dx)$. To be precise: for every $\beta \in (0,1)$ there exist universal constants $A_{[\beta],q}$ and $B_{[\beta],q}$ with

$$A_{[\beta],q} \cdot \Big\| \sum_\nu c_\nu \gamma_\nu^{(q)} \Big\|_{[\beta;dx]} \le \|\,(c_\nu)\,\|_{\ell^q} \le B_{[\beta],q} \cdot \Big\| \sum_\nu c_\nu \gamma_\nu^{(q)} \Big\|_{[\beta;dx]} . \tag{A.8.12}$$

Proof. (i) The functions $f \stackrel{\text{def}}{=} \sum_\nu c_\nu \gamma_\nu^{(q)}$ and $\|\,(c_\nu)\,\|_{\ell^q} \cdot \gamma_1^{(q)}$ are easily seen to have the same characteristic function, so their L^p-norms coincide.

(ii) In view of exercise A.8.15 and equation (A.8.11),

$$\|\,f\,\|_{[\beta]} \le \beta^{-1/p}\,\|\,f\,\|_{L^p(dx)} = \beta^{-1/p}\,\|\gamma^{(q)}\|_p \cdot \|\,(c_\nu)\,\|_{\ell^q}$$

for any $p \in (0,q)$: $A_{[\beta],q} \stackrel{\text{def}}{=} \sup\Big\{\beta^{1/p}\cdot\|\gamma^{(q)}\|_p^{-1} : 0<p<q\Big\}$ answers the left-hand side of inequality (A.8.12).

The constant $B_{[\beta],q}$ is somewhat harder to come by. Let $0<p_1<p_2<q$ and $r>1$ be so that $r \cdot p_1 = p_2$. Then the conjugate exponent of r is $r' = p_2/(p_2-p_1)$. For $\lambda>0$ write

$$\|f\|_{p_1}^{p_1} = \int_{[|f|>\lambda\|f\|_{p_1}]} |f|^{p_1} + \int_{[|f|\le\lambda\|f\|_{p_1}]} |f|^{p_1}$$

using Hölder:
$$\le \Big(\int |f|^{p_2}\Big)^{\frac{1}{r}} \cdot \big(dx[|f|>\lambda\|f\|_{p_1}]\big)^{\frac{1}{r'}} + \lambda^{p_1}\|f\|_{p_1}^{p_1} ,$$

and so $\quad \big(1-\lambda^{p_1}\big)\|f\|_{p_1}^{p_1} \le \|f\|_{p_2}^{p_1} \cdot \big(dx[|f|>\lambda\|f\|_{p_1}]\big)^{\frac{1}{r'}}$

and $\quad dx[|f|>\lambda\|f\|_{p_1}] \ge \left(\dfrac{\big(1-\lambda^{p_1}\big)\|f\|_{p_1}^{p_1}}{\|f\|_{p_2}^{p_1}} \vee 0\right)^{\frac{p_2}{p_2-p_1}}$

by equation (A.8.11):
$$= \left(\frac{\big(1-\lambda^{p_1}\big)\|\gamma^{(q)}\|_{p_1}^{p_1}}{\|\gamma^{(q)}\|_{p_2}^{p_1}} \vee 0\right)^{\frac{p_2}{p_2-p_1}} .$$

Therefore, setting $\lambda = \big(B \cdot \|\gamma^{(q)}\|_{p_1}\big)^{-1}$ and using equation (A.8.11):

$$dx[|f| > \|(c_\nu)\|_{\ell^q}/B] \ge \left(\frac{\big(\|\gamma^{(q)}\|_{p_1}^{p_1} - B^{-p_1}\big)}{\|\gamma^{(q)}\|_{p_2}^{p_1}} \vee 0\right)^{\frac{p_2}{p_2-p_1}} . \tag{$*$}$$

This inequality means

$$\|\,(c_\nu)\,\|_{\ell^q} \le B \cdot \|\,f\,\|_{[\beta(B,p_1,p_2,q)]} ,$$

where $\beta(B,p_1,p_2,q)$ denotes the right-hand side of $(*)$. The question is whether $\beta(B,p_1,p_2,q)$ can be made larger than the $\beta<1$ given in the statement, by a suitable choice of B. To see that it can be we solve the inequality $\beta(B, p_1, p_2, q) \geq \beta$ for B:

$$B \geq \left(\left(\|\gamma^{(q)}\|_{p_1}^{p_1} - \beta^{\frac{p_2-p_1}{p_2}} \|\gamma^{(q)}\|_{p_2}^{p_1} \right) \vee 0 \right)^{\frac{-1}{p_1}}$$

$$\text{by (A.8.10):} \quad = \left(\left(b(p_1) \cdot \Gamma\left(\frac{q-p_1}{q}\right) - \beta^{\frac{p_2-p_1}{p_2}} \left(b(p_2) \cdot \Gamma\left(\frac{q-p_2}{q}\right) \right)^{\frac{p_1}{p_2}} \right) \vee 0 \right)^{\frac{-1}{p_1}} .$$

Fix $p_2 < q$. As $p_1 \to 0$, $b(p_1) \to 1$ and $\Gamma\left(\frac{q-p_1}{q}\right) \to 1$, so the expression in the outermost parentheses will eventually be strictly positive, and $B < \infty$ satisfying this inequality can be found. We get the estimate:

$$B_{[\beta],q} \leq \inf \left(\left(\|\gamma^{(q)}\|_{p_1} - \beta^{\frac{p_2-p_1}{p_2}} \|\gamma^{(q)}\|_{p_2}^{\frac{p_1}{p_2}} \right) \vee 0 \right)^{\frac{-1}{p_1}} , \qquad \text{(A.8.13)}$$

the infimum being taken over all p_1, p_2 with $0 < p_1 < p_2 < q$. ∎

Maps and Spaces of Type (p, q) One-half of equation (A.8.11) holds even when the c_ν belong to a Banach space E, provided $0 < p < q < 1$: in this range there are universal constants $T_{p,q}$ so that for $x_1, \ldots, x_n \in E$

$$\Big\| \sum_{\nu=1}^{n} x_\nu \, \gamma_\nu^{(q)} \Big\|_{L^p(dx)} \leq T_{p,q} \cdot \Big(\sum_{\nu=1}^{n} \| x_\nu \|_E^q \Big)^{1/q} .$$

In order to prove this inequality it is convenient to extend it to maps and to make it into a definition:

Definition A.8.34 *Let $u : E \to F$ be a linear map between quasinormed vector spaces and $0 < p < q \leq 2$. u is a* ***map of type*** $(\boldsymbol{p}, \boldsymbol{q})$ *if there exists a constant $C < \infty$ such that for every finite collection $\{x_1, \ldots, x_n\} \subset E$*

$$\Big\| \Big\| \sum_{\nu=1}^{n} u(x_\nu) \, \gamma_\nu^{(q)} \Big\|_F \Big\|_{L^p(dx)} \leq C \cdot \Big(\sum_{\nu=1}^{n} \| x_\nu \|_E^q \Big)^{1/q} . \qquad \text{(A.8.14)}$$

Here $\gamma_1^{(q)}, \ldots, \gamma_n^{(q)}$ are independent symmetric q-stable random variables defined on some probability space $(X, \mathcal{X}, dx)$. The smallest constant C satisfying (A.8.14) is denoted by $T_{p,q}(u)$. A quasinormed space E (see page 381) is said to be a ***space of type*** $(\boldsymbol{p}, \boldsymbol{q})$ *if its identity map id_E is of type (p, q), and then we write $T_{p,q}(E) = T_{p,q}(id_E)$.*

Exercise A.8.35 The continuous linear maps of type (p, q) form a two-sided operator ideal: if $u : E \to F$ and $v : F \to G$ are continuous linear maps between quasinormed spaces, then $T_{p,q}(v \circ u) \leq \|u\| \cdot T_{p,q}(v)$ and $T_{p,q}(v \circ u) \leq \|v\| \cdot T_{p,q}(u)$.

Example A.8.36 [67] Let $(Y, \mathcal{Y}, dy)$ be a probability space. Then the natural injection j of $L^q(dy)$ into $L^p(dy)$ has type (p, q) with

$$T_{p,q}(j) \leq \|\gamma^{(q)}\|_p . \qquad \text{(A.8.15)}$$

Indeed, if $f_1, \ldots, f_n \in L^q(dy)$, then

$$\Big\| \Big\| \sum_{\nu=1}^{n} j(f_\nu)\, \gamma_\nu^{(q)} \Big\|_{L^p(dy)} \Big\|_{L^p(dx)} = \Big(\int \int \Big| \sum_{\nu=1}^{n} f_\nu(y)\, \gamma_\nu^{(q)}(x) \Big|^p \, dx\, dy \Big)^{1/p}$$

by equation (A.8.11):
$$= \|\gamma^{(q)}\|_p \cdot \Big(\int \Big(\sum_{\nu=1}^{n} |f_\nu(y)|^q \Big)^{p/q} dy \Big)^{1/p}$$

$$\le \|\gamma^{(q)}\|_p \cdot \Big(\int \Big(\sum_{\nu=1}^{n} |f_\nu(y)|^q \Big) dy \Big)^{1/q}$$

$$= \|\gamma^{(q)}\|_p \cdot \Big(\sum_{\nu=1}^{n} \|f_\nu\|_{L^q(dy)}^q \Big)^{1/q} .$$

Exercise A.8.37 Equality obtains in inequality (A.8.15).

Example A.8.38 [67] For $0 < p < q < 1$,

$$T_{p,q}(\ell^1) \le \|\gamma^{(q)}\|_p \cdot \frac{\|\gamma^{(1)}\|_q}{\|\gamma^{(1)}\|_p} .$$

To see this let $\gamma_1^{(1)}, \gamma_2^{(1)}, \ldots$ be a sequence of independent symmetric 1-stable (i.e., Cauchy) random variables defined on a probability space $(X, \mathcal{X}, dx)$, and consider the map u that associates with $(a_i) \in \ell^1$ the random variable $\sum_i a_i \gamma_i^{(1)}$. According to equation (A.8.11), u has norm $\|\gamma^{(1)}\|_q$ if considered as a map $u_q : \ell^1 \to L^q(dx)$, and has norm $\|\gamma^{(1)}\|_p$ if considered as a map $u_p : \ell^1 \to L^p(dx)$. Let j denote the injection of $L^q(dx)$ into $L^p(dx)$. Then by equation (A.8.11)

$$v \stackrel{\text{def}}{=} \|\gamma^{(1)}\|_p^{-1} \cdot j \circ u_q$$

is an isometry of ℓ^1 onto a subspace of $L^p(dx)$. Consequently

$$T_{p,q}(\ell^1) = T_{p,q}\big(id_{\ell^1}\big) = T_{p,q}(v^{-1} \circ v)$$

by exercise A.8.35:
$$\le T_{p,q}(v) \le \|\gamma^{(1)}\|_p^{-1} \cdot \big\| u_q \big\| \cdot T_{p,q}(j)$$

by example A.8.36:
$$\le \|\gamma^{(1)}\|_p^{-1} \cdot \|\gamma^{(1)}\|_q \cdot \|\gamma^{(q)}\|_p .$$

Proposition A.8.39 [67] *For $0 < p < q < 1$ every normed space E is of type (p, q):*

$$T_{p,q}(E) \le T_{p,q}(\ell^1) \le \|\gamma^{(q)}\|_p \cdot \frac{\|\gamma^{(1)}\|_q}{\|\gamma^{(1)}\|_p} . \tag{A.8.16}$$

Proof. Let $x_1, \ldots, x_n \in E$, and let E_0 be the finite-dimensional subspace spanned by these vectors. Next let $x_1', x_2', \ldots \in E_0$ be a sequence dense in the unit ball of E_0 and consider the map $\pi : \ell^1 \to E_0$ defined by

$$\pi\big((a_i)\big) = \sum_{i=1}^{\infty} a_i\, x_i' , \qquad (a_i) \in \ell^1 .$$

It is easily seen to be a contraction: $\|\pi(\boldsymbol{a})\|_{E_0} \le \|\boldsymbol{a}\|_{\ell^1}$. Also, given $\epsilon > 0$, we can find elements $\boldsymbol{a}_\nu \in \ell^1$ with $\pi(\boldsymbol{a}_\nu) = x_\nu$ and $\|\boldsymbol{a}_\nu\|_{\ell^1} \le \|x_\nu\|_E + \epsilon$. Using the independent symmetric q-stable random variables $\gamma_\nu^{(q)}$ from definition A.8.34 we get

$$\begin{aligned}
\Big\|\Big\|\sum_{\nu=1}^n x_\nu\,\gamma_\nu^{(q)}\Big\|_E\Big\|_{L^p(dx)} &= \Big\|\Big\|\sum_{\nu=1}^n \pi\circ id_{\ell^1}(\boldsymbol{a}_\nu)\,\gamma_\nu^{(q)}\Big\|_E\Big\|_{L^p(dx)} \\
&\le \|\pi\|\cdot T_{p,q}(\ell^1)\cdot\Big(\sum_{\nu=1}^n \|\boldsymbol{a}_\nu\|_{\ell^1}^q\Big)^{1/q} \\
&\le T_{p,q}(\ell^1)\cdot\Big(\sum_{\nu=1}^n \|x_\nu\|_E^q + \epsilon^q\Big)^{1/q} .
\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, inequality (A.8.16) follows. ∎

A.9 Semigroups of Operators

Definition A.9.1 *A family $\{T_t : 0 \le t < \infty\}$ of bounded linear operators on a Banach space C is a **semigroup** if $T_{s+t} = T_s \circ T_t$ for $s,t \ge 0$ and T_0 is the **identity operator** I. We shall need to consider only **contraction semigroups**: the **operator norm** $\|T_t\| \overset{\text{def}}{=} \sup\{\|T_t\phi\|_C : \|\phi\|_C \le 1\}$ is bounded by 1 for all $t \in [0,\infty)$. Such $T_.$ is **(strongly) continuous** if $t \mapsto T_t\phi$ is continuous from $[0,\infty)$ to C, for all $\phi \in C$.*

Exercise A.9.2 Then $T_.$ is, in fact, uniformly strongly continuous. That is to say, $\|T_t\phi - T_s\phi\| \xrightarrow[(t-s)\to 0]{} 0$ for every $\phi \in C$.

Resolvent and Generator

The resolvent or Laplace transform of a continuous contraction semigroup $T_.$ is the family $U_.$ of bounded linear operators U_α, defined on $\phi \in C$ as

$$U_\alpha\phi \overset{\text{def}}{=} \int_0^\infty e^{-\alpha t}\cdot T_t\phi\,dt\,, \qquad \alpha > 0\,.$$

This can be read as an improper Riemann integral or a Bochner integral (see A.3.15). U_α is evidently linear and has $\|\alpha U_\alpha\| \le 1$. The ***resolvent, identity***

$$U_\alpha - U_\beta = (\beta - \alpha)U_\alpha U_\beta\,, \tag{A.9.1}$$

is a straightforward consequence of a variable substitution and implies that all of the U_α have the same range $\mathcal{U} \overset{\text{def}}{=} U_1 C$. Since evidently $\alpha U_\alpha\phi \xrightarrow[\alpha\to\infty]{} \phi$ for all $\phi \in C$, $\mathcal{U}$ is dense in C.

The generator of a continuous contraction semigroup $T_.$ is the linear operator $\mathcal{A}$ defined by

$$\mathcal{A}\psi \overset{\text{def}}{=} \lim_{t\downarrow 0}\frac{T_t\psi - \psi}{t}\,. \tag{A.9.2}$$

It is not, in general, defined for all $\psi \in C$, so there is need to talk about its domain $\mathrm{dom}(\mathcal{A})$. This is the set of all $\psi \in C$ for which the limit (A.9.2) exists in C. It is very easy to see that T_t maps $\mathrm{dom}(\mathcal{A})$ to itself, and that $\mathcal{A}T_t\psi = T_t\mathcal{A}\psi$ for $\psi \in \mathrm{dom}(\mathcal{A})$ and $t \geq 0$. That is to say, $t \mapsto T_t\psi$ has a continuous right derivative at all $t \geq 0$; it is then actually differentiable at all $t > 0$ ([116, page 237 ff.]). In other words, $u_t \stackrel{\text{def}}{=} T_t\psi$ solves the C-valued initial value problem

$$\frac{du_t}{dt} = \mathcal{A}u_t \ , \ u_0 = \psi \ .$$

For an example pick a $\phi \in C$ and set $\psi \stackrel{\text{def}}{=} \int_0^s T_\sigma\phi \, d\sigma$. Then $\psi \in \mathrm{dom}(\mathcal{A})$ and a simple computation results in

$$\mathcal{A}\psi = T_s\phi - \phi \ . \tag{A.9.3}$$

The Fundamental Theorem of Calculus gives

$$T_t\psi - T_s\psi = \int_s^t T_\tau \, \mathcal{A}\psi \, d\tau \ . \tag{A.9.4}$$

for $\psi \in \mathrm{dom}(\mathcal{A})$ and $0 \leq s < t$.

If $\phi \in C$ and $\psi \stackrel{\text{def}}{=} U_\alpha\phi$, then the curve $t \mapsto T_t\psi = e^{\alpha t}\int_t^\infty e^{-\alpha s}T_s\phi \, ds$ is plainly differentiable at every $t \geq 0$, and a simple calculation produces

$$\mathcal{A}[T_t\psi] = T_t[\mathcal{A}\psi] = T_t[\alpha\psi - \phi] \ ,$$

and so at $t = 0$ $\qquad \mathcal{A}U_\alpha\phi = \alpha U_\alpha\phi - \phi \ ,$

or, equivalently, $\quad (\alpha I - \mathcal{A})U_\alpha = I \ \text{ or } \ (I - \mathcal{A}/\alpha)^{-1} = \alpha U_\alpha \ . \qquad$ (A.9.5)

This implies $\quad \left\|(I - \epsilon\mathcal{A})^{-1}\right\| \leq 1 \ \text{ for all } \epsilon > 0 \ . \qquad$ (A.9.6)

Exercise A.9.3 From this it is easy to read off these properties of the generator $\mathcal{A}$: (i) The domain of $\mathcal{A}$ contains the common range $\mathcal{U}$ of the resolvent operators U_α. In fact, $\mathrm{dom}(\mathcal{A}) = \mathcal{U}$, and therefore ***the domain of*** $\mathcal{A}$ ***is dense*** [54, p. 316]. (ii) Equation (A.9.3) also shows easily that $\mathcal{A}$ ***is a closed operator***, meaning that its graph $G_\mathcal{A} \stackrel{\text{def}}{=} \{(\psi, \mathcal{A}\psi) : \psi \in \mathrm{dom}(\mathcal{A})\}$ is a closed subset of $C \times C$. Namely, if $\mathrm{dom}(\mathcal{A}) \ni \psi_n \to \psi$ and $\mathcal{A}\psi_n \to \phi$, then by equation (A.9.4) $T_s\psi - \psi = \lim \int_0^s T_\sigma\mathcal{A}\psi_n \, d\sigma = \int_0^s T_\sigma\phi \, d\sigma$; dividing by s and letting $s \to 0$ shows that $\psi \in \mathrm{dom}(\mathcal{A})$ and $\phi = \mathcal{A}\psi$. (iii) $\mathcal{A}$ is ***dissipative***. This means that

$$\|(I - \epsilon\mathcal{A})\psi\|_C \geq \|\psi\|_C \tag{A.9.7}$$

for all $\epsilon > 0$ and all $\psi \in \mathrm{dom}(\mathcal{A})$ and follows directly from (A.9.6).

A subset $\mathcal{D}_0 \subset \mathrm{dom}\,\mathcal{A}$ is a ***core*** for $\mathcal{A}$ if the restriction $\mathcal{A}_0$ of $\mathcal{A}$ to $\mathcal{D}_0$ has closure $\mathcal{A}$ (meaning of course that the closure of $G_{\mathcal{A}_0}$ in $C \times C$ is $G_\mathcal{A}$).

Exercise A.9.4 $\mathcal{D}_0 \subset \mathrm{dom}\,\mathcal{A}$ is a core if and only if $(\alpha - \mathcal{A})\mathcal{D}_0$ is dense in C for some, and then for all, $\alpha > 0$. A dense ***invariant***[47] subspace $\mathcal{D}_0 \subset \mathrm{dom}\,\mathcal{A}$ is a core.

[47] I.e., $T_t\mathcal{D}_0 \subseteq \mathcal{D}_0 \quad \forall\, t \geq 0$.

Feller Semigroups

In this book we are interested only in the case where the Banach space C is the space $C_0(E)$ of continuous functions vanishing at infinity on some separable locally compact space E. Its norm is the supremum norm. This Banach space carries an additional structure, the order, and the semigroups of interest are those that respect it:

Definition A.9.5 ***A Feller semigroup*** *on E is a strongly continuous semigroup $T_{\bullet}$ on $C_0(E)$ of* ***positive***[48] *contractive linear operators T_t from $C_0(E)$ to itself. The Feller semigroup $T_{\bullet}$ is* ***conservative*** *if for every $x \in E$ and $t \ge 0$*

$$\sup\{T_t\phi\,(x) : \phi \in C_0(E)\ ,\ 0 \le \phi \le 1\} = 1\ .$$

The positivity and contractivity of a Feller semigroup imply that the linear functional $\phi \mapsto T_t\phi(x)$ on $C_0(E)$ is a positive Radon measure of total mass ≤ 1. It extends in any of the usual fashions (see, e.g., page 395) to a subprobability $T_t(x, \cdot)$ on the Borels of E. We may use the measure $T_t(x, \cdot)$ to write, for $\phi \in C_0(E)$,

$$T_t\phi(x) = \int T_t(x, dy)\ \phi(y)\ . \tag{A.9.8}$$

In terms of the ***transition*** *sub****probabilities*** $T_t(x, dy)$, the semigroup property of $T_{\bullet}$ reads

$$\int T_{s+t}(x, dy)\ \phi(y) = \int T_s(x, dy) \int T_t(y, dy')\ \phi(y') \tag{A.9.9}$$

and extends to all bounded Baire functions ϕ by the Monotone Class Theorem A.3.4; (A.9.9) is known as ***the Chapman–Kolmogorov equations***.

Remark A.9.6 Conservativity simply signifies that the $T_t(s, \cdot)$ are all probabilities. The study of general Feller semigroups can be reduced to that of conservative ones with the following little trick. Let us identify $C_0(E)$ with those continuous functions on the one-point compactification E^{Δ} that vanish at "the grave Δ." On any $\Phi \in C^{\Delta} \stackrel{\text{def}}{=} C(E^{\Delta})$ define the semigroup $T^{\Delta}_{\bullet}$ by

$$T^{\Delta}_t\Phi\,(x) = \begin{cases} \Phi(\Delta) + \int_E T_t(x, dy)\big(\Phi(y) - \Phi(\Delta)\big) & \text{if } x \in E, \\ \Phi(\Delta) & \text{if } x = \Delta. \end{cases} \tag{A.9.10}$$

We leave it to the reader to convince herself that $T^{\Delta}_{\bullet}$ is a strongly continuous conservative Feller semigroup on $C(E^{\Delta})$, and that "the grave" Δ is ***absorbing***: $T_t(\Delta, \{\Delta\}) = 1$. This terminology comes from the behavior of any process $X_{\bullet}$ stochastically representing $T^{\Delta}_{\bullet}$ (see definition 5.7.1); namely, once $X_{\bullet}$ has reached the grave it stays there. The compactification $T_{\bullet} \to T^{\Delta}_{\bullet}$ comes in handy even when $T_{\bullet}$ is conservative but E is not compact.

---

[48] That is to say, $\phi \ge 0$ implies $T_t\phi \ge 0$.

Examples A.9.7 (i) The simplest example of a conservative Feller semigroup perhaps is this: suppose that $\{\theta_s : 0 \le s < \infty\}$ is a semigroup under composition, of continuous maps $\theta_s : E \to E$ with $\lim_{s\downarrow 0}\theta_s(x) = x = \theta_0(x)$ for all $x \in E$, a ***flow***. Then $T_s\phi = \phi \circ \theta_s$ defines a Feller semigroup $T_.$ on $C_0(E)$, provided that the inverse image $\theta_s^{-1}(K)$ is compact whenever $K \subset E$ is.

(ii) Another example is the Gaussian semigroup of exercise 1.2.13 on $\mathbb{R}^d$:

$$\Gamma_t\phi(x) \stackrel{\text{def}}{=} \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} \phi(x+y)\, e^{-|y|^2/2t}\, dy = \overset{*}{\gamma}_t\star\phi\,(x)\ .$$

(iii) The Poisson semigroup is introduced in exercise 5.7.11, and the semigroup that comes with a Lévy process in equation (4.6.31).

(iv) ***A convolution semigroup of probabilities*** on $\mathbb{R}^n$ is a family $\{\mu_t : t \ge 0\}$ of probabilities so that $\mu_{s+t} = \mu_s\star\mu_t$ for $s, t > 0$ and $\mu_0 = \delta_0$. Such gives rise to a semigroup of bounded positive linear operators T_t on $C_0(\mathbb{R}^n)$ by the prescription

$$T_t\phi\,(\boldsymbol{z}) \stackrel{\text{def}}{=} \overset{*}{\mu}_t\star\phi\,(x) = \int_{\mathbb{R}^n} \phi(\boldsymbol{z}+\boldsymbol{z}')\ \mu_t(d\boldsymbol{z}')\ , \qquad \phi \in C_0(\mathbb{R}^n)\ ,\ \boldsymbol{z} \in \mathbb{R}^n\ .$$

It follows directly from proposition A.4.1 and corollary A.4.3 that the following are equivalent: (a) $\lim_{t\downarrow 0} T_t\phi = \phi$ for all $\phi \in C_0(\mathbb{R}^n)$; (b) $t \mapsto \widehat{\mu_t}(\boldsymbol{\zeta})$ is continuous on $\mathbb{R}_+$ for all $\boldsymbol{\zeta} \in \mathbb{R}^n$; and (c) $\mu_{t_n} \Rightarrow \mu_t$ weakly as $t_n \to t$. If any and then all of these continuity properties is satisfied, then $\{\mu_t : t \ge 0\}$ is called a ***conservative Feller convolution semigroup***.

A.9.8 Here are a few observations. They are either readily verified or substantiated in appendix C or are accessible in the concise but detailed presentation of Kallenberg [54, pages 313–326].

(i) The positivity of the T_t causes the resolvent operators to be positive as well. It causes the generator $\mathcal{A}$ to obey the ***positive maximum principle***; that is to say, whenever $\psi \in \mathrm{dom}(\mathcal{A})$ attains a positive maximum at $x \in E$, then $\mathcal{A}\psi\,(x) \le 0$.

(ii) If the semigroup $T_.$ is conservative, then its generator $\mathcal{A}$ is ***conservative*** as well. This means that there exists a sequence $\psi_n \in \mathrm{dom}(\mathcal{A})$ with $\sup_n \|\psi_n\|_\infty < \infty$; $\sup_n \|\mathcal{A}\psi_n\|_\infty < \infty$; and $\psi_n \to 1$, $\mathcal{A}\psi_n \to 0$ pointwise on E.

A.9.9 The Hille–Yosida Theorem states that the closure $\overline{\mathcal{A}}$ of a closable operator[49] $\mathcal{A}$ is the generator of a Feller semigroup – which is then unique – if and only if $\mathcal{A}$ is densely defined and satisfies the positive maximum principle, and $\alpha - \mathcal{A}$ has dense range in $C_0(E)$ for some, and then all, $\alpha > 0$.

---

[49] $\mathcal{A}$ is closable if the closure of its graph $G_{\mathcal{A}}$ in $C_0(E) \times C_0(E)$ is the graph of an operator $\overline{\mathcal{A}}$, which is then called the closure of $\mathcal{A}$. This simply means that the relation $\overline{G_{\mathcal{A}}} \subset C_0(E) \times C_0(E)$ actually is (the graph of) a function, equivalently, that $(0, \phi) \in \overline{G_{\mathcal{A}}}$ implies $\phi = 0$.

For proofs see [54, page 321], [101], and [116]. One idea is to emulate the formula $e^{at} = \lim_{n\to\infty}(1 - ta/n)^{-n}$ for real numbers a by proving that

$$T_t\phi \stackrel{\text{def}}{=} \lim_{n\to\infty}(I - t\overline{\mathcal{A}}/n)^{-n}\phi$$

exists for every $\phi \in C_0(E)$ and defines a contraction semigroup $T_.$ whose generator is $\overline{\mathcal{A}}$. This idea succeeds and we will take this for granted. It is then easy to check that the conservativity of $\mathcal{A}$ implies that of $T_.$.

The Natural Extension of a Feller Semigroup

Consider the second example of A.9.7. The Gaussian semigroup $\Gamma_.$ applies naturally to a much larger class of continuous functions than merely those vanishing at infinity. Namely, if ϕ grows at most exponentially at infinity,[50] then $\Gamma_t\phi$ is easily seen to show the same limited growth. This phenomenon is rather typical and asks for appropriate definitions. In other words, given a strongly continuous semigroup $T_.$, we are looking for a suitable extension $\breve{T}_.$ in the space $C = C(E)$ of continuous functions on E.

The natural topology of C is that of uniform convergence on compacta, which makes it a Fréchet space (examples A.2.27).

Exercise A.9.10 A curve $[0,\infty) \ni t \mapsto \psi_t$ in C is continuous if and only if the map $(s,x) \mapsto \psi_s(x)$ is continuous on every compact subset of $\mathbb{R}_+ \times E$.

Given a Feller semigroup $T_.$ on E and a function $f : E \to \mathbb{R}$, set[51]

$$\|f\|\breve{}_{t,K} \stackrel{\text{def}}{=} \sup\left\{\int_E^* T_s(x,dy)\,|f(y)| : \ 0 \le s \le t,\ x \in K\right\} \qquad \text{(A.9.11)}$$

for any $t > 0$ and any compact subset $K \subset E$, and then set

$$\lceil\!\lceil f \rceil\!\rceil\breve{} \stackrel{\text{def}}{=} \sum_\nu 2^{-\nu} \wedge \|f\|\breve{}_{\nu,K_\nu}\ ,$$

and
$$\begin{aligned}\breve{\mathfrak{F}} &\stackrel{\text{def}}{=} \left\{f : E \to \overline{\mathbb{R}}\ :\ \lceil\!\lceil \lambda f \rceil\!\rceil\breve{}\ \xrightarrow[\lambda\to 0]{} 0\right\}\\ &= \left\{f : E \to \overline{\mathbb{R}}\ :\ \|f\|\breve{}_{t,K} < \infty\ \forall t < \infty,\ \forall K \text{ compact}\right\}.\end{aligned}$$

The $\|\cdot\|\breve{}_{t,K}$ and $\lceil\!\lceil \cdot \rceil\!\rceil\breve{}$ are clearly solid and countably subadditive; therefore $\breve{\mathfrak{F}}$ is complete (theorem 3.2.10 on page 98). Since the $\|\cdot\|\breve{}_{t,K}$ are seminorms, this space is also locally convex. Let us now define the ***natural domain*** $\breve{C}$ of $T_.$ as the $\lceil\!\lceil\ \rceil\!\rceil\breve{}$-closure of $C_{00}(E)$ in $\breve{\mathfrak{F}}$, and the ***natural extension*** $\breve{T}_.$ on $\breve{C}$ by

$$\breve{T}_t\breve{\phi}\,(x) \stackrel{\text{def}}{=} \int_E T_t(x,dy)\ \breve{\phi}(y)$$

[50] $|\phi(x)| \le Ce^{c\|x\|}$ for some $C, c > 0$. In fact, for $\phi = e^{c\|x\|}$, $\int \Gamma_t(x,dy)\ \phi(y) = e^{tc^2/2}e^{c\|x\|}$.

[51] $\int^*$ denotes the upper integral – see equation (A.3.1) on page 396.

for $t \ge 0$, $\breve{\phi} \in \breve{C}$, and $x \in E$. Since the injection $C_{00}(E) \hookrightarrow \breve{C}$ is evidently continuous and $C_{00}(E)$ is separable, so is $\breve{C}$; since the topology is defined by the *seminorms* (A.9.11), $\breve{C}$ is locally convex; since $\breve{\mathfrak{F}}$ is complete, so is $\breve{C}$: $\breve{C}$ is a Fréchet space under the gauge $\lceil\!\lceil \ \rceil\!\rceil\breve{}$. Since $\lceil\!\lceil \ \rceil\!\rceil\breve{}$ is solid and $C_{00}(E)$ is a vector lattice, so is $\breve{C}$. Here is a reasonably simple membership criterion:

Exercise A.9.11 (i) A continuous function ϕ belongs to $\breve{C}$ if and only if for every $t < \infty$ and compact $K \subset E$ there exists a $\psi \in C$ with $|\phi| \le \psi$ so that the function $(s,x) \mapsto \int_E^* T_s(x,dy)\ \psi(y)$ is finite and continuous on $[0,t] \times K$. In particular, when $T_{\boldsymbol{.}}$ is conservative then $\breve{C}$ contains the bounded continuous functions $C_b = C_b(E)$ and in fact is a module over C_b.

(ii) $\breve{T}_{\boldsymbol{.}}$ is a strongly continuous semigroup of positive continuous linear operators.

A.9.12 The Natural Extension of Resolvent and Generator The Bochner integral[52]

$$\breve{U}_\alpha \psi = \int_0^\infty e^{-\alpha t} \cdot \breve{T}_t \psi \, dt \tag{A.9.12}$$

may fail to exist for some functions ψ in $\breve{C}$ and some $\alpha > 0$.[50] So we introduce the ***natural domains of the extended resolvent***

$$\breve{\mathcal{D}}_\alpha = \breve{\mathcal{D}}[\breve{U}_\alpha] \stackrel{\text{def}}{=} \left\{\psi \in \breve{C} : \text{the integral (A.9.12) exists and belongs to } \breve{C}\ \right\},$$

and on this set define the ***natural extension of the resolvent operator*** $\breve{U}_\alpha$ by (A.9.12). Similarly, the ***natural extension of the generator*** is defined by

$$\breve{A}\psi \stackrel{\text{def}}{=} \lim_{t \downarrow 0} \frac{\breve{T}_t \psi - \psi}{t} \tag{A.9.13}$$

on the subspace $\breve{\mathcal{D}} = \breve{\mathcal{D}}[\breve{A}] \subset \breve{C}$ where this limit exists and lies in $\breve{C}$. It is convenient and sufficient to understand the limit in (A.9.13) as a pointwise limit.

Exercise A.9.13 $\breve{\mathcal{D}}_\alpha$ increases with α and is contained in $\breve{\mathcal{D}}$. On $\breve{\mathcal{D}}_\alpha$ we have

$$(\alpha I - \breve{A})\breve{U}_\alpha = I\ .$$

The requirement that $\breve{A}\psi \in \breve{C}$ has the effect that $\breve{T}_t \breve{A}\psi = \breve{A}\breve{T}_t\psi$ for all $t \ge 0$ and

$$\breve{T}_t\psi - \breve{T}_s\psi = \int_s^t \breve{T}_\sigma \breve{A}\psi \, d\sigma\ , \qquad 0 \le s \le t < \infty\ .$$

A.9.14 A Feller Family of Transition Probabilities is a slew $\{T_{t,s} : 0 \le s \le t\}$ of positive contractive linear operators from $C_0(E)$ to itself such that for all $\phi \in C_0(E)$ and all $0 \le s \le t \le u < \infty$

$$T_{u,s}\phi = T_{u,t} \circ T_{t,s}\phi \ \text{ and } \ T_{t,t}\phi = \phi\ ;$$

$$(s,t) \mapsto T_{t,s}\phi \quad \text{ is continuous from } \mathbb{R}_+ \times \mathbb{R}_+ \text{ to } C_0(E).$$

[52] See item A.3.15 on page 400.

It is ***conservative*** if for every pair $s \le t$ and $x \in E$ the positive Radon measure

$$\phi \mapsto T_{t,s}\phi\,(x) \ \text{ is a probability } \ T_{t,s}(x, \cdot)\ ;$$

we may then write $\quad T_{t,s}\phi\,(x) = \displaystyle\int_E T_{t,s}(x, dy)\ \phi(y)\ .$

The study of $T_{\bullet,\bullet}$ can be reduced to that of a Feller semigroup with the following little trick: let $E^{\vdash} \stackrel{\text{def}}{=} \mathbb{R}_+ \times E$, and define $T^{\vdash}_t : C^{\vdash}_0 \stackrel{\text{def}}{=} C_0(E^{\vdash}) \to C^{\vdash}_0$ by

$$T^{\vdash}_s \phi\,(t, x) = \int_E T_{s+t,t}(x, dy)\ \phi(s+t, y)\ ,$$

or $$T^{\vdash}_s \big((t, x), d\tau \times d\xi\big) = \delta_{s+t}(d\tau) \times T_{s+t,t}(x, d\xi)\ .$$

Then $T^{\vdash}_{\bullet}$ is a Feller semigroup on $E^{\vdash}$. We call it the ***time-rectification*** of $T_{\bullet,\bullet}$. This little procedure can be employed to advantage even if $T_{\bullet,\bullet}$ is already "time-rectified," that is to say, even if $T_{t,s}$ depends only on the elapsed time $t - s$: $T_{t,s} = T_{t-s,0} = T_{t-s}$, where $T_{\bullet}$ is some semigroup.

Let us define the operators

$$\mathcal{A}_\tau : \phi \mapsto \lim_{t \downarrow 0} \frac{T_{\tau+t,\tau}\phi - \phi}{t}\ , \qquad \phi \in \operatorname{dom}(\mathcal{A}_\tau)\ ,$$

on the sets $\operatorname{dom}(\mathcal{A}_\tau)$ where these limits exist, and write

$$(\mathcal{A}_\tau \phi)(\tau, x) = \big(\mathcal{A}_\tau \phi(\tau, \cdot)\big)(x)$$

for $\phi(\tau, \cdot) \in \operatorname{dom}(\mathcal{A}_\tau)$. We leave it to the reader to connect these operators with the generator $A^{\vdash}$:

Exercise A.9.15 Describe the generator $A^{\vdash}$ of $T^{\vdash}_{\bullet}$ in terms of the operators $\mathcal{A}_\tau$. Identify $\operatorname{dom}(\breve{T}^{\vdash}_{\bullet})$, $\mathcal{D}[\breve{U}^{\vdash}_{\bullet}]$, $\breve{U}^{\vdash}_{\bullet}$, $\operatorname{dom}(\breve{A}^{\vdash})$, and $\breve{A}^{\vdash}$. In particular, for $\psi \in \operatorname{dom}(\breve{A}^{\vdash})$,

$$A^{\vdash}\psi(\tau, x) = \frac{\partial\psi}{\partial t}(\tau, x) + \mathcal{A}_\tau \psi(\tau, x)\ .$$

Repeated Footnotes: 366[1] 366[2] 366[3] 367[5] 367[6] 375[12] 376[13] 376[14] 384[15] 385[16] 388[18] 395[22] 397[26] 399[27] 410[34] 421[37] 421[38] 422[40] 436[42] 467[50]