Version of 31.5.03

Chapter 25

Product Measures

I come now to another chapter on 'pure' measure theory, discussing a fundamental construction – or, as you may prefer to consider it, two constructions, since the problems involved in forming the product of two arbitrary measure spaces (§251) are rather different from those arising in the product of arbitrarily many probability spaces (§254). This work is going to stretch our technique to the utmost, for while the fundamental theorems to which we are moving are natural aims, the proofs are lengthy and there are many pitfalls beside the true paths.

The central idea is that of 'repeated integration'. You have probably already seen formulae of the type '$\iint f(x,y)dxdy$' used to calculate the integral of a function of two real variables over a region in the plane. One of the basic techniques of advanced calculus is reversing the order of integration; for instance, we expect $\int_0^1(\int_y^1 f(x,y)dx)dy$ to be equal to $\int_0^1(\int_0^x f(x,y)dy)dx$. As I have developed the subject, we already have a third calculation to compare with these two: $\int_D f$, where $D=\{(x,y):0\le y\le x\le 1\}$ and the integral is taken with respect to Lebesgue measure on the plane. The first two sections of this chapter are devoted to an analysis of the relationship between one- and two-dimensional Lebesgue measure which makes these operations valid – some of the time; part of the work has to be devoted to a careful description of the exact conditions which must be imposed on f and D if we are to be safe.

Repeated integration, in one form or another, appears everywhere in measure theory, and it is therefore necessary sooner or later to develop the most general possible expression of the idea. The standard method is through the theory of products of general measure spaces. Given measure spaces (X,Σ,μ) and (Y,T,ν), the aim is to find a measure λ on $X\times Y$ which will, at least, give the right measure $\mu E\cdot\nu F$ to a 'rectangle' $E\times F$ where $E\in\Sigma$ and $F\in\mathrm{T}$. It turns out that there are already difficulties in deciding what 'the' product measure is, and to do the job properly I find I need, even at this stage, to describe two related but distinguishable constructions. These constructions and their elementary properties take up the whole of §251. In §252 I turn to integration over the product, with Fubini's and Tonelli's theorems relating $\int f d\lambda$ with $\iint f(x,y)\mu(dx)\nu(dy)$. Because the construction of λ is symmetric between the two factors, this automatically provides theorems relating $\iint f(x,y)\mu(dx)\nu(dy)$ with $\iint f(x,y)\nu(dy)\mu(dx)$. §253 looks at the space $L^1(\lambda)$ and its relationship with $L^1(\mu)$ and $L^1(\nu)$.

For general measure spaces, there are obstacles in the way of forming an infinite product; to start with, if $\langle(X_n,\mu_n)\rangle_{n\in\mathbb{N}}$ is a sequence of measure spaces, then a product measure λ on $X=\prod_{n\in\mathbb{N}}X_n$ ought to set $\lambda X=\prod_{n=0}^\infty\mu_nX_n$, and there is no guarantee that the product will converge, or behave well when it does. But for probability spaces, when $\mu_nX_n=1$ for every n, this problem at least evaporates. It is possible to define the product of any family of probability spaces; this is the burden of §254.

I end the chapter with three sections which are a preparation for Chapters 27 and 28, but are also important in their own right as an investigation of the way in which the group structure of $\mathbb{R}^r$ interacts with Lebesgue and other measures. §255 deals with the 'convolution' $f*g$ of two functions, where $(f*g)(x)=\int f(y)g(x-y)dy$ (the integration being with respect to Lebesgue measure). In §257 I show that some of the same ideas, suitably transformed, can be used to describe a convolution $\nu_1*\nu_2$ of two measures on $\mathbb{R}^r$; in preparation for this I include a section on Radon measures on $\mathbb{R}^r$ (§256).

Version of 10.11.06

251 Finite products

The first construction to set up is the product of a pair of measure spaces. It turns out that there are already substantial technical difficulties in the way of finding a canonical universally applicable method. I find myself therefore describing two related, but distinct, constructions, the 'primitive' and 'c.l.d.' product measures (251C, 251F). After listing the fundamental properties of the c.l.d product measure (251I-251J), I work through the identification of the product of Lebesgue measure with itself (251N) and a fairly thorough discussion of subspaces (251O-251S).

251A Definition Let (X,Σ,μ) and (Y,T,ν) be two measure spaces. For $A\subseteq X\times Y$ set

$$\theta A=\inf\{\sum_{n=0}^{\infty}\mu E_n\cdot\nu F_n : E_n\in\Sigma,\ F_n\in\mathrm{T}\ \forall\ n\in\mathbb{N},\ A\subseteq\bigcup_{n\in\mathbb{N}}E_n\times F_n\}.$$

Remark In the products $\mu E_n\cdot\nu F_n$, $0\cdot\infty$ is to be taken as 0, as in §135.

251B Lemma In the context of 251A, θ is an outer measure on $X\times Y$.

proof (a) Setting $E_n=F_n=\emptyset$ for every $n\in\mathbb{N}$, we see that $\theta\emptyset=0$.

(b) If $A\subseteq B\subseteq X\times Y$, then whenever $B\subseteq\bigcup_{n\in\mathbb{N}}E_n\times F_n$ we shall have $A\subseteq\bigcup_{n\in\mathbb{N}}E_n\times F_n$; so $\theta A\le\theta B$.

(c) Let $\langle A_n\rangle_{n\in\mathbb{N}}$ be a sequence of subsets of $X\times Y$, with union A. For any $\epsilon>0$, we may choose, for each $n\in\mathbb{N}$, sequences $\langle E_{nm}\rangle_{m\in\mathbb{N}}$ in Σ and $\langle F_{nm}\rangle_{m\in\mathbb{N}}$ in T such that $A_n\subseteq\bigcup_{m\in\mathbb{N}}E_{nm}\times F_{nm}$ and $\sum_{m=0}^{\infty}\mu E_{nm}\cdot\nu F_{nm}\le\theta A_n+2^{-n}\epsilon$. Because $\mathbb{N}\times\mathbb{N}$ is countable, we have a bijection $k\mapsto(n_k,m_k):\mathbb{N}\to\mathbb{N}\times\mathbb{N}$, and now

$$A\subseteq\bigcup_{n,m\in\mathbb{N}}E_{nm}\times F_{nm}=\bigcup_{k\in\mathbb{N}}E_{n_km_k}\times F_{n_km_k},$$

so that

$$\begin{aligned}\theta A&\le\sum_{k=0}^{\infty}\mu E_{n_km_k}\cdot\nu F_{n_km_k}=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\mu E_{nm}\cdot\nu F_{nm}\\&\le\sum_{n=0}^{\infty}\theta A_n+2^{-n}\epsilon=2\epsilon+\sum_{n=0}^{\infty}\theta A_n.\end{aligned}$$

As ϵ is arbitrary, $\theta A\le\sum_{n=0}^{\infty}\theta A_n$.

As $\langle A_n\rangle_{n\in\mathbb{N}}$ is arbitrary, θ is an outer measure.

251C Definition Let (X,Σ,μ) and (Y,T,ν) be measure spaces. By the **primitive product measure** on $X\times Y$ I shall mean the measure λ_0 derived by Carathéodory's method (113C) from the outer measure θ defined in 251A.

Remark I ought to point out that there is no general agreement on what 'the' product measure on $X\times Y$ should be. Indeed in 251F below I will introduce an alternative one, and in the notes to this section I will mention a third.

251D Definition It is convenient to have a name for a natural construction for σ-algebras. If X and Y are sets with σ-algebras $\Sigma\subseteq\mathcal{P}X$ and $\mathrm{T}\subseteq\mathcal{P}Y$, I will write $\Sigma\widehat{\otimes}\mathrm{T}$ for the σ-algebra of subsets of $X\times Y$ generated by $\{E\times F:E\in\Sigma,\ F\in\mathrm{T}\}$.

251E Proposition Let (X,Σ,μ) and (Y,T,ν) be measure spaces; let λ_0 be the primitive product measure on $X\times Y$, and Λ its domain. Then $\Sigma\widehat{\otimes}\mathrm{T}\subseteq\Lambda$ and $\lambda_0(E\times F)=\mu E\cdot\nu F$ for all $E\in\Sigma$ and $F\in\mathrm{T}$.

proof (a) Suppose that $E\in\Sigma$ and $A\subseteq X\times Y$. For any $\epsilon>0$, there are sequences $\langle E_n\rangle_{n\in\mathbb{N}}$ in Σ and $\langle F_n\rangle_{n\in\mathbb{N}}$ in T such that $A\subseteq\bigcup_{n\in\mathbb{N}}E_n\times F_n$ and $\sum_{n=0}^{\infty}\mu E_n\cdot\nu F_n\le\theta A+\epsilon$. Now

$$A\cap(E\times Y)\subseteq\bigcup_{n\in\mathbb{N}}(E_n\cap E)\times F_n,\quad A\setminus(E\times Y)\subseteq\bigcup_{n\in\mathbb{N}}(E_n\setminus E)\times F_n,$$

so

$$\theta(A\cap(E\times Y))+\theta(A\setminus(E\times Y))\le\sum_{n=0}^{\infty}\mu(E_n\cap E)\cdot\nu F_n+\sum_{n=0}^{\infty}\mu(E_n\setminus E)\cdot\nu F_n$$
$$=\sum_{n=0}^{\infty}\mu E_n\cdot\nu F_n\le\theta A+\epsilon.$$

As ϵ is arbitrary, $\theta(A\cap(E\times Y))+\theta(A\setminus(E\times Y))\le\theta A$. And this is enough to ensure that $E\times Y\in\Lambda$ (see 113D).

(b) Similarly, $X\times F\in\Lambda$ for every $F\in\mathrm{T}$, so $E\times F=(E\times Y)\cap(X\times F)\in\Lambda$ for every $E\in\Sigma$, $F\in\mathrm{T}$. Because Λ is a σ-algebra, it must include the smallest σ-algebra containing all the products $E\times F$, that is, $\Lambda\supseteq\Sigma\widehat{\otimes}\mathrm{T}$.

(c) Take $E\in\Sigma$, $F\in\mathrm{T}$. We know that $E\times F\in\Lambda$; setting $E_0=E$, $F_0=F$, $E_n=F_n=\emptyset$ for $n\ge1$ in the definition of θ, we have

$$\lambda_0(E\times F)=\theta(E\times F)\le\mu E\cdot\nu F.$$

We have come to the central idea of the construction. In fact $\theta(E\times F)=\mu E\cdot\nu F$. **P** Suppose that $E\times F\subseteq\bigcup_{n\in\mathbb{N}}E_n\times F_n$ where $E_n\in\Sigma$ and $F_n\in\mathrm{T}$ for every n. Set $u=\sum_{n=0}^{\infty}\mu E_n\cdot\nu F_n$. If $u=\infty$ or $\mu E=0$ or $\nu F=0$ then of course $\mu E\cdot\nu F\le u$. Otherwise, set

$$I=\{n:n\in\mathbb{N},\,\mu E_n=0\},\quad J=\{n:n\in\mathbb{N},\,\nu F_n=0\},\quad K=\mathbb{N}\setminus(I\cup J),$$

$$E'=E\setminus\bigcup_{n\in I}E_n,\quad F'=F\setminus\bigcup_{n\in J}F_n.$$

Then $\mu E'=\mu E$ and $\nu F'=\nu F$; $E'\times F'\subseteq\bigcup_{n\in K}E_n\times F_n$; and for $n\in K$, $\mu E_n<\infty$ and $\nu F_n<\infty$, since $\mu E_n\cdot\nu F_n\le u<\infty$ and neither μE_n nor νF_n is zero. Set

$$f_n=\nu F_n\chi E_n:X\to\mathbb{R}$$

if $n\in K$, and $f_n=\mathbf{0}:X\to\mathbb{R}$ if $n\in I\cup J$. Then f_n is a simple function and $\int f_n=\nu F_n\mu E_n$ for $n\in K$, 0 otherwise, so

$$\sum_{n=0}^{\infty}\int f_n(x)\mu(dx)=\sum_{n=0}^{\infty}\mu E_n\cdot\nu F_n\le u.$$

By B.Levi's theorem (123A), applied to $\langle\sum_{k=0}^{n}f_k\rangle_{n\in\mathbb{N}}$, $g=\sum_{n=0}^{\infty}f_n$ is integrable and $\int g\,d\mu\le u$. Write E'' for $\{x:x\in E',\,g(x)<\infty\}$, so that $\mu E''=\mu E'=\mu E$. Now take any $x\in E''$ and set $K_x=\{n:n\in K,\,x\in E_n\}$. Because $E'\times F'\subseteq\bigcup_{n\in K}E_n\times F_n$, $F'\subseteq\bigcup_{n\in K_x}F_n$ and

$$\nu F=\nu F'\le\sum_{n\in K_x}\nu F_n=\sum_{n=0}^{\infty}f_n(x)=g(x).$$

Thus $g(x)\ge\nu F$ for every $x\in E''$. We are supposing that $0<\mu E=\mu E''$ and $0<\nu F$, so we must have $\nu F<\infty$, $\mu E''<\infty$. Now $g\ge\nu F\chi E''$, so

$$\mu E\cdot\nu F=\mu E''\cdot\nu F=\int\nu F\chi E''\le\int g\le u=\sum_{n=0}^{\infty}\mu E_n\cdot\nu F_n.$$

As $\langle E_n\rangle_{n\in\mathbb{N}}$, $\langle F_n\rangle_{n\in\mathbb{N}}$ are arbitrary, $\theta(E\times F)\ge\mu E\cdot\nu F$ and $\theta(E\times F)=\mu E\cdot\nu F$. **Q**

Thus

$$\lambda_0(E\times F)=\theta(E\times F)=\mu E\cdot\nu F$$

for all $E\in\Sigma$, $F\in\mathrm{T}$.

251F Definition Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and λ_0 the primitive product measure defined in 251C. By the **c.l.d. product measure** on $X\times Y$ I shall mean the function $\lambda:\operatorname{dom}\lambda_0\to[0,\infty]$ defined by setting

$$\lambda W=\sup\{\lambda_0(W\cap(E\times F)):E\in\Sigma,\,F\in\mathrm{T},\,\mu E<\infty,\,\nu F<\infty\}$$

for $W\in\operatorname{dom}\lambda_0$.

251G Remark I had better show at once that λ is a measure. **P** Of course its domain $\Lambda=\operatorname{dom}\lambda_0$ is a σ-algebra, and $\lambda\emptyset=\lambda_0\emptyset=0$. If $\langle W_n\rangle_{n\in\mathbb{N}}$ is a disjoint sequence in Λ, then for any $E\in\Sigma$, $F\in\mathrm{T}$ of finite measure

$$\lambda_0(\bigcup_{n\in\mathbb{N}} W_n \cap (E\times F)) = \sum_{n=0}^\infty \lambda_0(W_n\cap(E\times F)) \le \sum_{n=0}^\infty \lambda W_n,$$

so $\lambda(\bigcup_{n\in\mathbb{N}} W_n) \le \sum_{n=0}^\infty \lambda W_n$. On the other hand, if $a < \sum_{n=0}^\infty \lambda W_n$, then we can find $m\in\mathbb{N}$ and $a_0,\dots,a_m$ such that $a\le\sum_{n=0}^m a_n$ and $a_n<\lambda W_n$ for each $n\le m$; now there are $E_0,\dots,E_m\in\Sigma$ and $F_0,\dots,F_m\in\mathrm{T}$, all of finite measure, such that $a_n\le\lambda_0(W_n\cap(E_n\times F_n))$ for each n. Setting $E=\bigcup_{n\le m}E_n$ and $F=\bigcup_{n\le m}F_n$, we have $\mu E<\infty$ and $\nu F<\infty$, so

$$\begin{aligned}\lambda(\bigcup_{n\in\mathbb{N}} W_n) &\ge \lambda_0(\bigcup_{n\in\mathbb{N}} W_n\cap(E\times F)) = \sum_{n=0}^\infty \lambda_0(W_n\cap(E\times F))\\ &\ge \sum_{n=0}^m \lambda_0(W_n\cap(E_n\times F_n)) \ge \sum_{n=0}^m a_n \ge a.\end{aligned}$$

As a is arbitrary, $\lambda(\bigcup_{n\in\mathbb{N}} W_n)\ge\sum_{n=0}^\infty\lambda W_n$ and $\lambda(\bigcup_{n\in\mathbb{N}} W_n)=\sum_{n=0}^\infty\lambda W_n$. As $\langle W_n\rangle_{n\in\mathbb{N}}$ is arbitrary, λ is a measure. **Q**

251H We need a simple property of the measure λ_0.

Lemma Let (X,Σ,μ) and (Y,T,ν) be two measure spaces; let λ_0 be the primitive product measure on $X\times Y$, and Λ its domain. If $H\subseteq X\times Y$ and $H\cap(E\times F)\in\Lambda$ whenever $\mu E<\infty$ and $\nu F<\infty$, then $H\in\Lambda$.

proof Let θ be the outer measure described in 251A. Suppose that $A\subseteq X\times Y$ and $\theta A<\infty$. Let $\epsilon>0$. Let $\langle E_n\rangle_{n\in\mathbb{N}}$, $\langle F_n\rangle_{n\in\mathbb{N}}$ be sequences in Σ, T respectively such that $A\subseteq\bigcup_{n\in\mathbb{N}}E_n\times F_n$ and $\sum_{n=0}^\infty\mu E_n\cdot\nu F_n\le\theta A+\epsilon$. Now, for each n, the product of the measures μE_n, νE_n is finite, so either one is zero or both are finite. If $\mu E_n=0$ or $\nu F_n=0$ then of course

$$\mu E_n\cdot\nu F_n = 0 = \theta((E_n\times F_n)\cap H)+\theta((E_n\times F_n)\setminus H).$$

If $\mu E_n<\infty$ and $\nu F_n<\infty$ then

$$\begin{aligned}\mu E_n\cdot\nu F_n &= \lambda_0(E_n\times F_n)\\ &= \lambda_0((E_n\times F_n)\cap H)+\lambda_0((E_n\times F_n)\setminus H)\\ &= \theta((E_n\times F_n)\cap H)+\theta((E_n\times F_n)\setminus H).\end{aligned}$$

Accordingly, because θ is an outer measure,

$$\begin{aligned}\theta(A\cap H)+\theta(A\setminus H) &\le \sum_{n=0}^\infty\theta((E_n\times F_n)\cap H)+\sum_{n=0}^\infty\theta((E_n\times F_n)\setminus H)\\ &= \sum_{n=0}^\infty\mu E_n\cdot\nu F_n\le\theta A+\epsilon.\end{aligned}$$

As ϵ is arbitrary, $\theta(A\cap H)+\theta(A\setminus H)\le\theta A$. As A is arbitrary, $H\in\Lambda$.

251I Now for the fundamental properties of the c.l.d. product measure.

Theorem Let (X,Σ,μ) and (Y,T,ν) be measure spaces; let λ be the c.l.d. product measure on $X\times Y$, and Λ its domain. Then

(a) $\Sigma\widehat{\otimes}\mathrm{T}\subseteq\Lambda$ and $\lambda(E\times F)=\mu E\cdot\nu F$ whenever $E\in\Sigma$, $F\in\mathrm{T}$ and $\mu E\cdot\nu F<\infty$;

(b) for every $W\in\Lambda$ there is a $V\in\Sigma\widehat{\otimes}\mathrm{T}$ such that $V\subseteq W$ and $\lambda V=\lambda W$;

(c) $(X\times Y,\Lambda,\lambda)$ is complete and locally determined, and in fact is the c.l.d. version of $(X\times Y,\Lambda,\lambda_0)$ as described in 213D-213E; in particular, $\lambda W=\lambda_0 W$ whenever $\lambda_0 W<\infty$;

(d) if $W\in\Lambda$ and $\lambda W>0$ then there are $E\in\Sigma$, $F\in\mathrm{T}$ such that $\mu E<\infty$, $\nu F<\infty$ and $\lambda(W\cap(E\times F))>0$;

(e) if $W \in \Lambda$ and $\lambda W < \infty$, then for every $\epsilon > 0$ there are $E_0, \dots, E_n \in \Sigma$, $F_0, \dots, F_n \in \mathrm{T}$, all of finite measure, such that $\lambda(W \triangle \bigcup_{i \le n}(E_i \times F_i)) \le \epsilon$.

proof Take θ to be the outer measure of 251A and λ_0 the primitive product measure of 251C. Set $\Sigma^f = \{E : E \in \Sigma,\ \mu E < \infty\}$ and $\mathrm{T}^f = \{F : F \in \mathrm{T},\ \nu F < \infty\}$.

(a) By 251E, $\Sigma \widehat{\otimes} \mathrm{T} \subseteq \Lambda$. If $E \in \Sigma$ and $F \in \mathrm{T}$ and $\mu E \cdot \nu F < \infty$, either $\mu E \cdot \nu F = 0$ and $\lambda(E \times F) = \lambda_0(E \times F) = 0$ or both μE and νF are finite and again $\lambda(E \times F) = \lambda_0(E \times F) = \mu E \cdot \nu F$.

(b)(i) Take any $a < \lambda W$. Then there are $E \in \Sigma^f$, $F \in \mathrm{T}^f$ such that $\lambda_0(W \cap (E \times F)) > a$ (251F); now

$$\begin{aligned}\theta((E \times F) \setminus W) &= \lambda_0((E \times F) \setminus W)\\ &= \lambda_0(E \times F) - \lambda_0(W \cap (E \times F)) < \lambda_0(E \times F) - a.\end{aligned}$$

Let $\langle E_n \rangle_{n \in \mathbb{N}}$, $\langle F_n \rangle_{n \in \mathbb{N}}$ be sequences in Σ, T respectively such that $(E \times F) \setminus W \subseteq \bigcup_{n \in \mathbb{N}} E_n \times F_n$ and $\sum_{n=0}^{\infty} \mu E_n \cdot \nu F_n \le \lambda_0(E \times F) - a$. Consider

$$V = (E \times F) \setminus \bigcup_{n \in \mathbb{N}} E_n \times F_n \in \Sigma \widehat{\otimes} \mathrm{T};$$

then $V \subseteq W$, and

$$\begin{aligned}\lambda V = \lambda_0 V &= \lambda_0(E \times F) - \lambda_0((E \times F) \setminus V)\\ &\ge \lambda_0(E \times F) - \lambda_0(\bigcup_{n \in \mathbb{N}} E_n \times F_n)\end{aligned}$$

(because $(E \times F) \setminus V \subseteq \bigcup_{n \in \mathbb{N}} E_n \times F_n$)

$$\ge \lambda_0(E \times F) - \sum_{n=0}^{\infty} \mu E_n \cdot \nu F_n \ge a$$

(by the choice of the E_n, F_n).

(ii) Thus for every $a < \lambda W$ there is a $V \in \Sigma \widehat{\otimes} \mathrm{T}$ such that $V \subseteq W$ and $\lambda V \ge a$. Now choose a sequence $\langle a_n \rangle_{n \in \mathbb{N}}$ strictly increasing to λW, and for each a_n a corresponding V_n; then $V = \bigcup_{n \in \mathbb{N}} V_n$ belongs to the σ-algebra $\Sigma \widehat{\otimes} \mathrm{T}$, is included in W, and has measure at least $\sup_{n \in \mathbb{N}} \lambda V_n$ and at most λW; so $\lambda V = \lambda W$, as required.

(c)(i) If $H \subseteq X \times Y$ is λ-negligible, there is a $W \in \Lambda$ such that $H \subseteq W$ and $\lambda W = 0$. If $E \in \Sigma$, $F \in \mathrm{T}$ are of finite measure, $\lambda_0(W \cap (E \times F)) = 0$; but λ_0, being derived from the outer measure θ by Carathéodory's method, is complete (212A), so $H \cap (E \times F) \in \Lambda$ and $\lambda_0(H \cap (E \times F)) = 0$. Because E and F are arbitrary, $H \in \Lambda$, by 251H. As H is arbitrary, λ is complete.

(ii) If $W \in \Lambda$ and $\lambda W = \infty$, then there must be $E \in \Sigma$, $F \in \mathrm{T}$ such that $\mu E < \infty$, $\nu F < \infty$ and $\lambda_0(W \cap (E \times F)) > 0$; now

$$0 < \lambda(W \cap (E \times F)) \le \mu E \cdot \nu F < \infty.$$

Thus λ is semi-finite.

(iii) If $H \subseteq X \times Y$ and $H \cap W \in \Lambda$ whenever $\lambda W < \infty$, then, in particular, $H \cap (E \times F) \in \Lambda$ whenever $\mu E < \infty$ and $\nu F < \infty$; by 251H again, $H \in \Lambda$. Thus λ is locally determined.

(iv) If $W \in \Lambda$ and $\lambda_0 W < \infty$, then we have sequences $\langle E_n \rangle_{n \in \mathbb{N}}$ in Σ, $\langle F_n \rangle_{n \in \mathbb{N}}$ in T such that $W \subseteq \bigcup_{n \in \mathbb{N}}(E_n \times F_n)$ and $\sum_{n=0}^{\infty} \mu E_n \cdot \nu F_n < \infty$. Set

$$I = \{n : \mu E_n = \infty\}, \quad J = \{n : \nu F_n = \infty\}, \quad K = \mathbb{N} \setminus (I \cup J);$$

then $\nu(\bigcup_{n \in I} F_n) = \mu(\bigcup_{n \in J} E_n) = 0$, so $\lambda_0(W \setminus W') = 0$, where

$$W' = W \cap \bigcup_{n \in K}(E_n \times F_n) \supseteq W \setminus ((\bigcup_{n \in J} E_n \times Y) \cup (X \times \bigcup_{n \in I} F_n)).$$

Now set $E'_n = \bigcup_{i \in K, i \le n} E_i$, $F'_n = \bigcup_{i \in K, i \le n} F_i$ for each n. We have $W' = \bigcup_{n \in \mathbb{N}} W' \cap (E'_n \times F'_n)$, so

$$\lambda W \le \lambda_0 W = \lambda_0 W' = \lim_{n\to\infty} \lambda_0(W' \cap (E'_n \times F'_n)) \le \lambda W' \le \lambda W,$$

and $\lambda W = \lambda_0 W$.

(v) Following the terminology of 213D, let us write

$$\tilde\Lambda = \{W : W \subseteq X \times Y,\ W \cap V \in \Lambda \text{ whenever } V \in \Lambda \text{ and } \lambda_0 V < \infty\},$$

$$\tilde\lambda W = \sup\{\lambda_0(W \cap V) : V \in \Lambda,\ \lambda_0 V < \infty\}.$$

Because $\lambda_0(E \times F) < \infty$ whenever $\mu E < \infty$ and $\nu F < \infty$, $\tilde\Lambda \subseteq \Lambda$ and $\tilde\Lambda = \Lambda$.

Now for any $W \in \Lambda$ we have

$$\begin{aligned}
\tilde\lambda W &= \sup\{\lambda_0(W \cap V) : V \in \Lambda,\ \lambda_0 V < \infty\}\\
&\ge \sup\{\lambda_0(W \cap (E \times F)) : E \in \Sigma^f,\ F \in \mathrm{T}^f\}\\
&= \lambda W\\
&\ge \sup\{\lambda(W \cap V) : V \in \Lambda,\ \lambda_0 V < \infty\}\\
&= \sup\{\lambda_0(W \cap V) : V \in \Lambda,\ \lambda_0 V < \infty\},
\end{aligned}$$

using (iv) just above, so that $\lambda = \tilde\lambda$ is the c.l.d. version of λ_0.

(d) If $W \in \Lambda$ and $\lambda W > 0$, there are $E \in \Sigma^f$ and $F \in \mathrm{T}^f$ such that $\lambda(W \cap (E \times F)) = \lambda_0(W \cap (E \times F)) > 0$.

(e) There are $E \in \Sigma^f$, $F \in \mathrm{T}^f$ such that $\lambda_0(W \cap (E \times F)) \ge \lambda W - \frac{1}{3}\epsilon$; set $V_1 = W \cap (E \times F)$; then

$$\lambda(W \setminus V_1) = \lambda W - \lambda V_1 = \lambda W - \lambda_0 V_1 \le \frac{1}{3}\epsilon.$$

There are sequences $\langle E'_n\rangle_{n\in\mathbb{N}}$ in Σ, $\langle F'_n\rangle_{n\in\mathbb{N}}$ in T such that $V_1 \subseteq \bigcup_{n\in\mathbb{N}} E'_n \times F'_n$ and $\sum_{n=0}^{\infty} \mu E'_n \cdot \nu F'_n \le \lambda_0 V_1 + \frac{1}{3}\epsilon$. Replacing E'_n, F'_n by $E'_n \cap E$, $F'_n \cap F$ if necessary, we may suppose that $E'_n \in \Sigma^f$ and $F'_n \in \mathrm{T}^f$ for every n. Set $V_2 = \bigcup_{n\in\mathbb{N}} E'_n \times F'_n$; then

$$\lambda(V_2 \setminus V_1) \le \lambda_0(V_2 \setminus V_1) \le \sum_{n=0}^{\infty} \mu E'_n \cdot \nu F'_n - \lambda_0 V_1 \le \frac{1}{3}\epsilon.$$

Let $m \in \mathbb{N}$ be such that $\sum_{n=m+1}^{\infty} \mu E'_n \cdot \nu F'_n \le \frac{1}{3}\epsilon$, and set

$$V = \bigcup_{n=0}^{m} E'_n \times F'_n.$$

Then

$$\lambda(V_2 \setminus V) \le \sum_{n=m+1}^{\infty} \mu E'_n \cdot \nu F'_n \le \frac{1}{3}\epsilon.$$

Putting these together, we have $W \triangle V \subseteq (W \setminus V_1) \cup (V_2 \setminus V_1) \cup (V_2 \setminus V)$, so

$$\lambda(W \triangle V) \le \lambda(W \setminus V_1) + \lambda(V_2 \setminus V_1) + \lambda(V_2 \setminus V) \le \frac{1}{3}\epsilon + \frac{1}{3}\epsilon + \frac{1}{3}\epsilon = \epsilon.$$

And V is of the required form.

251J Proposition If (X, Σ, μ) and (Y, T, ν) are semi-finite measure spaces and λ is the c.l.d. product measure on $X \times Y$, then $\lambda(E \times F) = \mu E \cdot \nu F$ for all $E \in \Sigma$, $F \in \mathrm{T}$.

proof Setting $\Sigma^f = \{E : E \in \Sigma,\ \mu E < \infty\}$, $\mathrm{T}^f = \{F : F \in \mathrm{T},\ \nu F < \infty\}$, we have

$$\begin{aligned}
\lambda(E \times F) &= \sup\{\lambda_0((E \cap E_0) \times (F \cap F_0)) : E_0 \in \Sigma^f,\ F_0 \in \mathrm{T}^f\}\\
&= \sup\{\mu(E \cap E_0) \cdot \nu(F \cap F_0)) : E_0 \in \Sigma^f,\ F_0 \in \mathrm{T}^f\}\\
&= \sup\{\mu(E \cap E_0) : E_0 \in \Sigma^f\} \cdot \sup\{\nu(F \cap F_0) : F_0 \in \mathrm{T}^f\} = \mu E \cdot \nu F
\end{aligned}$$

(using 213A).

251K σ-finite spaces Of course most of the measure spaces we shall apply these results to are σ-finite, and in this case there are some useful simplifications.

Proposition Let (X,Σ,μ) and (Y,T,ν) be σ-finite measure spaces. Then the c.l.d. product measure on $X\times Y$ is equal to the primitive product measure, and is the completion of its restriction to $\Sigma\widehat{\otimes}\mathrm{T}$; moreover, this common product measure is σ-finite.

proof Write λ_0, λ for the primitive and c.l.d. product measures, as usual, and Λ for their domain. Let $\langle E_n\rangle_{n\in\mathbb{N}}$, $\langle F_n\rangle_{n\in\mathbb{N}}$ be non-decreasing sequences of sets of finite measure covering X, Y respectively (see 211D).

(a) For each $n\in\mathbb{N}$, $\lambda(E_n\times F_n)=\mu E_n\cdot\nu F_n$ is finite, by 251Ia. Since $X\times Y=\bigcup_{n\in\mathbb{N}}E_n\times F_n$, λ is σ-finite.

(b) For any $W\in\Lambda$,

$$\lambda_0 W=\lim_{n\to\infty}\lambda_0(W\cap(E_n\times F_n))=\lim_{n\to\infty}\lambda(W\cap(E_n\times F_n))=\lambda W.$$

So $\lambda=\lambda_0$.

(c) Write $\lambda_{\mathcal{B}}$ for the restriction of $\lambda=\lambda_0$ to $\Sigma\widehat{\otimes}\mathrm{T}$, and $\hat\lambda_{\mathcal{B}}$ for its completion.

(i) Suppose that $W\in\operatorname{dom}\hat\lambda_{\mathcal{B}}$. Then there are W', $W''\in\Sigma\widehat{\otimes}\mathrm{T}$ such that $W'\subseteq W\subseteq W''$ and $\lambda_{\mathcal{B}}(W''\setminus W')=0$ (212C). In this case, $\lambda(W''\setminus W')=0$; as λ is complete, $W\in\Lambda$ and

$$\lambda W=\lambda W'=\lambda_{\mathcal{B}}W'=\hat\lambda_{\mathcal{B}}W.$$

Thus λ extends $\hat\lambda_{\mathcal{B}}$.

(ii) If $W\in\Lambda$, then there is a $V\in\Sigma\widehat{\otimes}\mathrm{T}$ such that $V\subseteq W$ and $\lambda(W\setminus V)=0$. **P** For each $n\in\mathbb{N}$ there is a $V_n\in\Sigma\widehat{\otimes}\mathrm{T}$ such that $V_n\subseteq W\cap(E_n\times F_n)$ and $\lambda V_n=\lambda(W\cap(E_n\times F_n))$ (251Ib). But as $\lambda(E_n\times F_n)=\mu E_n\cdot\nu F_n$ is finite, this means that $\lambda(W\cap(E_n\times F_n)\setminus V_n)=0$. So if we set $V=\bigcup_{n\in\mathbb{N}}V_n$, we shall have $V\in\Sigma\widehat{\otimes}\mathrm{T}$, $V\subseteq W$ and

$$W\setminus V=\bigcup_{n\in\mathbb{N}}W\cap(E_n\times F_n)\setminus V\subseteq\bigcup_{n\in\mathbb{N}}W\cap(E_n\times F_n)\setminus V_n$$

is λ-negligible. **Q**

Similarly, there is a $V'\in\Sigma\widehat{\otimes}\mathrm{T}$ such that $V'\subseteq(X\times Y)\setminus W$ and $\lambda(((X\times Y)\setminus W)\setminus V')=0$. Setting $V''=(X\times Y)\setminus V'$, $V''\in\Sigma\widehat{\otimes}\mathrm{T}$, $W\subseteq V''$ and $\lambda(V''\setminus W)=0$. So

$$\lambda_{\mathcal{B}}(V''\setminus V)=\lambda(V''\setminus V)=\lambda(V''\setminus W)+\lambda(W\setminus V)=0,$$

and W is measured by $\hat\lambda_{\mathcal{B}}$, with $\hat\lambda_{\mathcal{B}}W=\lambda_{\mathcal{B}}V=\lambda W$. As W is arbitrary, $\hat\lambda_{\mathcal{B}}=\lambda$.

***251L** The following result fits in naturally here; I star it because it will be used seldom (there is a more important version of the same idea in 254G) and the proof can be skipped until you come to need it.

Proposition Let (X_1,Σ_1,μ_1), (X_2,Σ_2,μ_2), (Y_1,T_1,ν_1) and (Y_2,T_2,ν_2) be σ-finite measure spaces; let λ_1, λ_2 be the product measures on $X_1\times Y_1$, $X_2\times Y_2$ respectively. Suppose that $f:X_1\to X_2$ and $g:Y_1\to Y_2$ are inverse-measure-preserving functions, and that $h(x,y)=(f(x),g(y))$ for $x\in X_1$, $y\in Y_1$. Then h is inverse-measure-preserving.

proof Write Λ_1, Λ_2 for the domains of λ_1, λ_2 respectively.

(a) Suppose that $E\in\Sigma_2$ and $F\in\mathrm{T}_2$ have finite measure. Then $\lambda_1 h^{-1}[W\cap(E\times F)]$ is defined and equal to $\lambda_2(W\cap(E\times F))$ for every $W\in\Lambda_2$. **P**

$$\begin{aligned}\lambda_1 h^{-1}[E\times F]&=\lambda_1(f^{-1}[E]\times g^{-1}[F])=\mu_1 f^{-1}[E]\cdot\nu_1 g^{-1}[F]\\&=\mu_2E\cdot\nu_2F=\lambda_2(E\times F)\end{aligned}$$

by 251E/251J. **Q**

(b) Take $E_0\in\Sigma_2$ and $F_0\in\mathrm{T}_2$ of finite measure. Let $\tilde\lambda_1$, $\tilde\lambda_2$ be the subspace measures on $f^{-1}[E_0]\times g^{-1}[F_0]$ and $E_0\times F_0$ respectively. Set $\tilde h=h\restriction f^{-1}[E_0]\times g^{-1}[F_0]$, and write ι for the identity map from $E_0\times F_0$ to $X_2\times Y_2$; let $\lambda=\tilde\lambda_1\tilde h^{-1}$ and $\lambda'=\tilde\lambda_2\iota^{-1}$ be the image measures on $X_2\times Y_2$. Then (a) tells us that

$$\lambda(E\times F)=\lambda_1(h^{-1}[(E\cap E_0)\times(F\cap F_0)]) \\ =\lambda_2((E\cap E_0)\times(F\cap F_0))=\lambda'(E\times F)$$

whenever $E\in\Sigma_2$ and $F\in\mathrm{T}_2$. By the Monotone Class Theorem (136C), λ and λ' agree on $\Sigma_2\widehat{\otimes}\mathrm{T}_2$, that is, $\lambda_1(h^{-1}[W\cap(E_0\times F_0)])=\lambda_2(W\cap(E_0\times F_0))$ for every $W\in\Sigma_2\widehat{\otimes}\mathrm{T}_2$.

If W is any member of Λ_2, there are W', $W''\in\Sigma_2\widehat{\otimes}\mathrm{T}_2$ such that $W'\subseteq W\subseteq W''$ and $\lambda_2(W''\setminus W')=0$ (251K). Now we must have

$$h^{-1}[W'\cap(E_0\times F_0)]\subseteq h^{-1}[W\cap(E_0\times F_0)]\subseteq h^{-1}[W''\cap(E_0\times F_0)],$$

$$\lambda_1(h^{-1}[W''\cap(E_0\times F_0)]\setminus h^{-1}[W'\cap(E_0\times F_0)])=\lambda_2((W''\setminus W')\cap(E_0\times F_0))=0;$$

because λ_1 is complete, $\lambda_1h^{-1}[W\cap(E_0\times F_0)]$ is defined and equal to

$$\lambda_1h^{-1}[W'\cap(E_0\times F_0)]=\lambda_2(W'\cap(E_0\times F_0))=\lambda_2(W\cap(E_0\times F_0)).$$

(c) Now suppose that $\langle E_n\rangle_{n\in\mathbb{N}}$, $\langle F_n\rangle_{n\in\mathbb{N}}$ are non-decreasing sequences of sets of finite measure with union X_2, Y_2 respectively. If $W\in\Lambda_2$,

$$\lambda_1h^{-1}[W]=\sup_{n\in\mathbb{N}}\lambda_1h^{-1}[W\cap(E_n\times F_n)]=\sup_{n\in\mathbb{N}}\lambda_2(W\cap(E_n\times F_n))=\lambda_2W.$$

So h is inverse-measure-preserving, as claimed.

251M It is time that I gave some examples. Of course the central example is Lebesgue measure. In this case we have the only reasonable result. I pause to describe the leading example of the product $\Sigma\widehat{\otimes}\mathrm{T}$ introduced in 251D.

Proposition Let r, $s\geq1$ be integers. Then we have a natural bijection $\phi:\mathbb{R}^r\times\mathbb{R}^s\to\mathbb{R}^{r+s}$, defined by setting

$$\phi((\xi_1,\dots,\xi_r),(\eta_1,\dots,\eta_s))=(\xi_1,\dots,\xi_r,\eta_1,\dots,\eta_s)$$

for $\xi_1,\dots,\xi_r,\eta_1,\dots,\eta_s\in\mathbb{R}$. If we write $\mathcal{B}_r$, $\mathcal{B}_s$ and $\mathcal{B}_{r+s}$ for the Borel σ-algebras of $\mathbb{R}^r$, $\mathbb{R}^s$ and $\mathbb{R}^{r+s}$ respectively, then ϕ identifies $\mathcal{B}_{r+s}$ with $\mathcal{B}_r\widehat{\otimes}\mathcal{B}_s$.

proof (a) Write $\mathcal{B}$ for the σ-algebra $\{\phi^{-1}[W]:W\in\mathcal{B}_{r+s}\}$ copied onto $\mathbb{R}^r\times\mathbb{R}^s$ by the bijection ϕ; we are seeking to prove that $\mathcal{B}=\mathcal{B}_r\widehat{\otimes}\mathcal{B}_s$. We have maps $\pi_1:\mathbb{R}^{r+s}\to\mathbb{R}^r$, $\pi_2:\mathbb{R}^{r+s}\to\mathbb{R}^s$ defined by setting $\pi_1(\phi(x,y))=x$, $\pi_2(\phi(x,y))=y$. Each co-ordinate of π_1 is continuous, therefore Borel measurable (121Db), so $\pi_1^{-1}[E]\in\mathcal{B}_{r+s}$ for every $E\in\mathcal{B}_r$, by 121K. Similarly, $\pi_2^{-1}[F]\in\mathcal{B}_{r+s}$ for every $F\in\mathcal{B}_s$. So $\phi[E\times F]=\pi_1^{-1}[E]\cap\pi_1^{-1}[F]$ belongs to $\mathcal{B}_{r+s}$, that is, $E\times F\in\mathcal{B}$, whenever $E\in\mathcal{B}_r$ and $F\in\mathcal{B}_s$. Because $\mathcal{B}$ is a σ-algebra, $\mathcal{B}_r\widehat{\otimes}\mathcal{B}_s\subseteq\mathcal{B}$.

(b) Now examine sets of the form

$$\{(x,y):\xi_i\leq\alpha\}=\{x:\xi_i\leq\alpha\}\times\mathbb{R}^s,$$

$$\{(x,y):\eta_j\leq\alpha\}=\mathbb{R}^r\times\{y:\eta_j\leq\alpha\}$$

for $\alpha\in\mathbb{R}$, $i\leq r$ and $j\leq s$, taking $x=(\xi_1,\dots,\xi_r)$ and $y=(\eta_1,\dots,\eta_s)$. All of these belong to $\mathcal{B}_r\widehat{\otimes}\mathcal{B}_s$. But the σ-algebra they generate is just $\mathcal{B}$, by 121J. So $\mathcal{B}\subseteq\mathcal{B}_r\widehat{\otimes}\mathcal{B}_s$ and $\mathcal{B}=\mathcal{B}_r\widehat{\otimes}\mathcal{B}_s$.

251N Theorem Let r, $s\geq1$ be integers. Then the bijection $\phi:\mathbb{R}^r\times\mathbb{R}^s\to\mathbb{R}^{r+s}$ described in 251M identifies Lebesgue measure on $\mathbb{R}^{r+s}$ with the c.l.d. product λ of Lebesgue measure on $\mathbb{R}^r$ and Lebesgue measure on $\mathbb{R}^s$.

proof Write μ_r, μ_s, μ_{r+s} for the three versions of Lebesgue measure, μ_r^*, μ_s^* and μ_{r+s}^* for the corresponding outer measures, and θ for the outer measure on $\mathbb{R}^r\times\mathbb{R}^s$ derived from μ_r and μ_s by the formula of 251A.

(a) If $I\subseteq\mathbb{R}^r$ and $J\subseteq\mathbb{R}^s$ are half-open intervals, then $\phi[I\times J]\subseteq\mathbb{R}^{r+s}$ is also a half-open interval, and

$$\mu_{r+s}(\phi[I\times J])=\mu_rI\cdot\mu_sJ;$$

this is immediate from the definition of the Lebesgue measure of an interval. (I speak of 'half-open' intervals here, that is, intervals of the form $\prod_{1\le j\le r}[\alpha_j,\beta_j[$, because I used them in the definition of Lebesgue measure in §115. If you prefer to work with open intervals or closed intervals it makes no difference.) Note also that every half-open interval in $\mathbb{R}^{r+s}$ is expressible as $\phi[I\times J]$ for suitable I, J.

(b) For any $A\subseteq\mathbb{R}^{r+s}$, $\theta(\phi^{-1}[A])\le\mu^*_{r+s}(A)$. **P** For any $\epsilon>0$, there is a sequence $\langle K_n\rangle_{n\in\mathbb{N}}$ of half-open intervals in $\mathbb{R}^{r+s}$ such that $A\subseteq\bigcup_{n\in\mathbb{N}}K_n$ and $\sum_{n=0}^\infty\mu_{r+s}(K_n)\le\mu^*_{r+s}(A)+\epsilon$. Express each K_n as $\phi[I_n\times J_n]$, where I_n and J_n are half-open intervals in $\mathbb{R}^r$ and $\mathbb{R}^s$ respectively; then $\phi^{-1}[A]\subseteq\bigcup_{n\in\mathbb{N}}I_n\times J_n$, so that

$$\theta(\phi^{-1}[A])\le\sum_{n=0}^\infty\mu_rI_n\cdot\mu_sJ_n=\sum_{n=0}^\infty\mu_{r+s}(K_n)\le\mu^*_{r+s}(A)+\epsilon.$$

As ϵ is arbitrary, we have the result. **Q**

(c) If $E\subseteq\mathbb{R}^r$ and $F\subseteq\mathbb{R}^s$ are measurable, then $\mu^*_{r+s}(\phi[E\times F])\le\mu_rE\cdot\mu_sF$.

P (i) Consider first the case $\mu_rE<\infty$, $\mu_sF<\infty$. In this case, given $\epsilon>0$, there are sequences $\langle I_n\rangle_{n\in\mathbb{N}}$, $\langle J_n\rangle_{n\in\mathbb{N}}$ of half-open intervals such that $E\subseteq\bigcup_{n\in\mathbb{N}}I_n$, $F\subseteq\bigcup_{n\in\mathbb{N}}F_n$,

$$\sum_{n=0}^\infty\mu_rI_n\le\mu^*_rE+\epsilon=\mu_rE+\epsilon,$$

$$\sum_{n=0}^\infty\mu_sJ_n\le\mu^*_sF+\epsilon=\mu_sF+\epsilon.$$

Accordingly $E\times F\subseteq\bigcup_{m,n\in\mathbb{N}}I_m\times J_n$ and $\phi[E\times F]\subseteq\bigcup_{m,n\in\mathbb{N}}\phi[I_m\times J_n]$, so that

$$\begin{aligned}\mu^*_{r+s}(\phi[E\times F])&\le\sum_{m,n=0}^\infty\mu_{r+s}(\phi[I_m\times J_n])=\sum_{m,n=0}^\infty\mu_rI_m\cdot\mu_sJ_n\\&=\sum_{m=0}^\infty\mu_rI_m\cdot\sum_{n=0}^\infty\mu_sJ_n\le(\mu_rE+\epsilon)(\mu_sF+\epsilon).\end{aligned}$$

As ϵ is arbitrary, we have the result.

(ii) Next, if $\mu_rE=0$, there is a sequence $\langle F_n\rangle_{n\in\mathbb{N}}$ of sets of finite measure covering $\mathbb{R}^s\supseteq F$, so that

$$\mu^*_{r+s}(\phi[E\times F])\le\sum_{n=0}^\infty\mu^*_{r+s}(\phi[E\times F_n])\le\sum_{n=0}^\infty\mu_rE\cdot\mu_sF_n=0=\mu_rE\cdot\mu_sF.$$

Similarly, $\mu^*_{r+s}(\phi[E\times F])\le\mu_rE\cdot\mu_sF$ if $\mu_sF=0$. The only remaining case is that in which both of μ_rE, μ_sF are strictly positive and one is infinite; but in this case $\mu_rE\cdot\mu_sF=\infty$, so surely $\mu^*_{r+s}(\phi[E\times F])\le\mu_rE\cdot\mu_sF$. **Q**

(d) If $A\subseteq\mathbb{R}^{r+s}$, then $\mu^*_{r+s}(A)\le\theta(\phi^{-1}[A])$. **P** Given $\epsilon>0$, there are sequences $\langle E_n\rangle_{n\in\mathbb{N}}$, $\langle F_n\rangle_{n\in\mathbb{N}}$ of measurable sets in $\mathbb{R}^r$, $\mathbb{R}^s$ respectively such that $\phi^{-1}[A]\subseteq\bigcup_{n\in\mathbb{N}}E_n\times F_n$ and $\sum_{n=0}^\infty\mu_rE_n\cdot\mu_sF_n\le\theta(\phi^{-1}[A])+\epsilon$. Now $A\subseteq\bigcup_{n\in\mathbb{N}}\phi[E_n\times F_n]$, so

$$\mu^*_{r+s}(A)\le\sum_{n=0}^\infty\mu^*_{r+s}(\phi[E_n\times F_n])\le\sum_{n=0}^\infty\mu_rE_n\cdot\mu_sF_n\le\theta(\phi^{-1}[A])+\epsilon.$$

As ϵ is arbitrary, we have the result. **Q**

(e) Putting (c) and (d) together, we have $\theta(\phi^{-1}[A])=\mu^*_{r+s}(A)$ for every $A\subseteq\mathbb{R}^{r+s}$. Thus θ on $\mathbb{R}^r\times\mathbb{R}^s$ corresponds exactly to μ^*_{r+s} on $\mathbb{R}^{r+s}$. So the associated measures λ_0, μ_{r+s} must correspond in the same way, writing λ_0 for the primitive product measure. But 251K tells us that $\lambda_0=\lambda$, so we have the result.

251O In fact, a large proportion of the applications of the constructions here are to subspaces of Euclidean space, rather than to the whole product $\mathbb{R}^r\times\mathbb{R}^s$. It would not have been especially difficult to write 251N out to deal with arbitrary subspaces, but I prefer to give a more general description of the product of subspace measures, as I feel that it illuminates the method. I start with a straightforward result on strictly localizable spaces.

Proposition Let (X,Σ,μ) and (Y,T,ν) be strictly localizable measure spaces. Then the c.l.d. product measure on $X\times Y$ is strictly localizable; moreover, if $\langle X_i\rangle_{i\in I}$ and $\langle Y_j\rangle_{j\in J}$ are decompositions of X and Y respectively, $\langle X_i\times Y_j\rangle_{(i,j)\in I\times J}$ is a decomposition of $X\times Y$.

proof Let $\langle X_i\rangle_{i\in I}$ and $\langle Y_j\rangle_{j\in J}$ be decompositions of X, Y respectively. Then $\langle X_i\times Y_j\rangle_{(i,j)\in I\times J}$ is a partition of $X\times Y$ into measurable sets of finite measure. If $W\subseteq X\times Y$ and $\lambda W>0$, there are sets $E\in\Sigma$, $F\in\mathrm{T}$ such that $\mu E<\infty$, $\nu F<\infty$ and $\lambda(W\cap(E\times F))>0$. We know that $\mu E=\sum_{i\in I}\mu(E\cap X_i)$ and $\mu F=\sum_{j\in J}\mu(F\cap Y_j)$, so there must be finite sets $I_0\subseteq I$, $J_0\subseteq J$ such that

$$\mu E\cdot\nu F-(\textstyle\sum_{i\in I_0}\mu(E\cap X_i))(\sum_{j\in J_0}\nu(F\cap Y_j))<\lambda(W\cap(E\times F)).$$

Setting $E'=\bigcup_{i\in I_0}X_i$ and $F'=\bigcup_{j\in J_0}Y_j$ we have

$$\lambda((E\times F)\setminus(E'\times F'))=\lambda(E\times F)-\lambda((E\cap E')\times(F\cap F'))<\lambda(W\cap(E\times F)),$$

so that $\lambda(W\cap(E'\times F'))>0$. There must therefore be some $i\in I_0$, $j\in J_0$ such that $\lambda(W\cap(X_i\times Y_j))>0$.

This shows that $\{X_i\times Y_j:i\in I,\,j\in J\}$ satisfies the criterion of 213O, so that λ, being complete and locally determined, must be strictly localizable. Because $\langle X_i\times Y_j\rangle_{(i,j)\in I\times J}$ covers $X\times Y$, it is actually a decomposition of $X\times Y$ (213Ob).

251P Lemma Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and λ the c.l.d. product measure on $X\times Y$. Let λ^* be the corresponding outer measure (132B). Then

$$\lambda^*C=\sup\{\theta(C\cap(E\times F)):E\in\Sigma,\,F\in\mathrm{T},\,\mu E<\infty,\,\nu F<\infty\}$$

for every $C\subseteq X\times Y$, where θ is the outer measure of 251A.

proof Write Λ for the domain of λ, Σ^f for $\{E:E\in\Sigma,\,\mu E<\infty\}$, T^f for $\{F:F\in\mathrm{T},\,\nu F<\infty\}$; set $u=\sup\{\theta(C\cap(E\times F)):E\in\Sigma^f,\,F\in\mathrm{T}^f\}$.

(a) If $C\subseteq W\in\Lambda$, $E\in\Sigma^f$ and $F\in\mathrm{T}^f$, then

$$\theta(C\cap(E\times F))\le\theta(W\cap(E\times F))=\lambda_0(W\cap(E\times F))$$

(where λ_0 is the primitive product measure)

$$\le\lambda W.$$

As E and F are arbitrary, $u\le\lambda W$; as W is arbitrary, $u\le\lambda^*C$.

(b) If $u=\infty$, then of course $\lambda^*C=u$. Otherwise, let $\langle E_n\rangle_{n\in\mathbb{N}}$, $\langle F_n\rangle_{n\in\mathbb{N}}$ be sequences in Σ^f, T^f respectively such that

$$u=\sup_{n\in\mathbb{N}}\theta(C\cap(E_n\times F_n)).$$

Consider $C'=C\setminus(\bigcup_{n\in\mathbb{N}}E_n\times\bigcup_{n\in\mathbb{N}}F_n)$. If $E\in\Sigma^f$ and $F\in\mathrm{T}^f$, then for every $n\in\mathbb{N}$ we have

$$\begin{aligned}u&\ge\theta(C\cap((E\cup E_n)\times(F\cup F_n)))\\&=\theta(C\cap((E\cup E_n)\times(F\cup F_n))\cap(E_n\times F_n))\\&\qquad+\theta(C\cap((E\cup E_n)\times(F\cup F_n))\setminus(E_n\times F_n))\end{aligned}$$

(because $E_n\times F_n\in\Lambda$, by 251E)

$$\ge\theta(C\cap(E_n\times F_n))+\theta(C'\cap(E\times F)).$$

Taking the supremum of the right-hand expression as n varies, we have $u\ge u+\theta(C'\cap(E\times F))$ so

$$\lambda(C'\cap(E\times F))=\theta(C'\cap(E\times F))=0.$$

As E and F are arbitrary, $\lambda C'=0$.

But this means that

$$\begin{aligned}\lambda^*C&\le\lambda^*(C\cap(\bigcup_{n\in\mathbb{N}}E_n\times\bigcup_{n\in\mathbb{N}}F_n))+\lambda^*C'\\&=\lim_{n\to\infty}\lambda^*(C\cap(\bigcup_{i\le n}E_i\times\bigcup_{i\le n}F_i))\end{aligned}$$

(using 132Ae)

$\leq u,$

as required.

251Q Proposition Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and $A\subseteq X$, $B\subseteq Y$ subsets; write μ_A, ν_B for the subspace measures on A, B respectively. Let λ be the c.l.d. product measure on $X\times Y$, and $\lambda^\#$ the subspace measure it induces on $A\times B$. Let $\tilde\lambda$ be the c.l.d. product measure of μ_A and ν_B on $A\times B$. Then

(i) $\tilde\lambda$ extends $\lambda^\#$.

(ii) If

either (α) $A\in\Sigma$ and $B\in\mathrm{T}$

or (β) A and B can both be covered by sequences of sets of finite measure

or (γ) μ and ν are both strictly localizable,

then $\tilde\lambda=\lambda^\#$.

proof Let θ be the outer measure on $X\times Y$ defined from μ and ν by the formula of 251A, and $\tilde\theta$ the outer measure on $A\times B$ similarly defined from μ_A and ν_B. Write Λ for the domain of λ, $\tilde\Lambda$ for the domain of $\tilde\lambda$, and $\Lambda^\#=\{W\cap(A\times B):W\in\Lambda\}$ for the domain of $\lambda^\#$. Set $\Sigma^f=\{E:\mu E<\infty\}$, $\mathrm{T}^f=\{F:\nu F<\infty\}$.

(a) The first point to observe is that $\tilde\theta C=\theta C$ for every $C\subseteq A\times B$. **P** (i) If $\langle E_n\rangle_{n\in\mathbb{N}}$ and $\langle F_n\rangle_{n\in\mathbb{N}}$ are sequences in Σ, T respectively such that $C\subseteq\bigcup_{n\in\mathbb{N}}E_n\times F_n$, then

$$C=C\cap(A\times B)\subseteq\bigcup_{n\in\mathbb{N}}(E_n\cap A)\times(F_n\cap B),$$

so

$$\begin{aligned}\tilde\theta C&\leq\sum_{n=0}^\infty\mu_A(E_n\cap A)\cdot\nu_B(F_n\cap B)\\&=\sum_{n=0}^\infty\mu^*(E_n\cap A)\cdot\nu^*(F_n\cap B)\leq\sum_{n=0}^\infty\mu E_n\cdot\nu F_n.\end{aligned}$$

As $\langle E_n\rangle_{n\in\mathbb{N}}$ and $\langle F_n\rangle_{n\in\mathbb{N}}$ are arbitrary, $\tilde\theta C\leq\theta C$. (ii) If $\langle\tilde E_n\rangle_{n\in\mathbb{N}}$, $\langle\tilde F_n\rangle_{n\in\mathbb{N}}$ are sequences in $\Sigma_A=\operatorname{dom}\mu_A$, $\mathrm{T}_B=\operatorname{dom}\nu_B$ respectively such that $C\subseteq\bigcup_{n\in\mathbb{N}}\tilde E_n\times\tilde F_n$, then for each $n\in\mathbb{N}$ we can choose $E_n\in\Sigma$, $F_n\in\mathrm{T}$ such that

$$\tilde E_n\subseteq E_n,\quad\mu E_n=\mu^*\tilde E_n=\mu_A\tilde E_n,$$

$$\tilde F_n\subseteq F_n,\quad\nu F_n=\nu^*\tilde F_n=\nu_B\tilde F_n,$$

and now

$$\theta C\leq\sum_{n=0}^\infty\mu E_n\cdot\nu F_n=\sum_{n=0}^\infty\mu_A\tilde E_n\cdot\nu_B\tilde F_n.$$

As $\langle\tilde E_n\rangle_{n\in\mathbb{N}}$, $\langle\tilde F_n\rangle_{n\in\mathbb{N}}$ are arbitrary, $\theta C\leq\tilde\theta C$. **Q**

(b) It follows that $\Lambda^\#\subseteq\tilde\Lambda$. **P** Suppose that $V\in\Lambda^\#$ and that $C\subseteq A\times B$. In this case there is a $W\in\Lambda$ such that $V=W\cap(A\times B)$. So

$$\tilde\theta(C\cap V)+\tilde\theta(C\setminus V)=\theta(C\cap W)+\theta(C\setminus W)=\theta C=\tilde\theta C.$$

As C is arbitrary, $V\in\tilde\Lambda$. **Q**

Accordingly, for $V\in\Lambda^\#$,

$$\begin{aligned}\lambda^\#V=\lambda^*V&=\sup\{\theta(V\cap(E\times F)):E\in\Sigma^f,\,F\in\mathrm{T}^f\}\\&=\sup\{\theta(V\cap(\tilde E\times\tilde F)):\tilde E\in\Sigma_A,\,\tilde F\in\mathrm{T}_B,\,\mu_A\tilde E<\infty,\,\nu_B\tilde F<\infty\}\\&=\sup\{\tilde\theta(V\cap(\tilde E\times\tilde F)):\tilde E\in\Sigma_A,\,\tilde F\in\mathrm{T}_B,\,\mu_A\tilde E<\infty,\,\nu_B\tilde F<\infty\}=\tilde\lambda V,\end{aligned}$$

using 251P twice.

This proves part (i) of the proposition.

(c) The next thing to check is that if $V \in \tilde{\Lambda}$ and $V \subseteq E \times F$ where $E \in \Sigma^f$ and $F \in \mathrm{T}^f$, then $V \in \Lambda^\#$. **P** Let $W \subseteq E \times F$ be a measurable envelope of V with respect to λ (132Ee). Then

$$\theta(W \cap (A \times B) \setminus V) = \tilde{\theta}(W \cap (A \times B) \setminus V) = \tilde{\lambda}(W \cap (A \times B) \setminus V)$$

(because $W \cap (A \times B) \in \Lambda^\# \subseteq \tilde{\Lambda}$, $V \in \tilde{\Lambda}$)

$$\begin{aligned} &= \tilde{\lambda}(W \cap (A \times B)) - \tilde{\lambda}V = \tilde{\theta}(W \cap (A \times B)) - \tilde{\theta}V \\ &= \theta(W \cap (A \times B)) - \theta V = \lambda^*(W \cap (A \times B)) - \lambda^* V \\ &\le \lambda W - \lambda^* V = 0. \end{aligned}$$

But this means that $W' = W \cap (A \times B) \setminus V \in \Lambda$ and $V = (A \times B) \cap (W \setminus W')$ belongs to $\Lambda^\#$. **Q**

(d) Now fix any $V \in \tilde{\Lambda}$, and look at the conditions (α)-(γ) of part (ii) of the proposition.

(α) If $A \in \Sigma$ and $B \in \mathrm{T}$, and $C \subseteq X \times Y$, then $A \times B \in \Lambda$ (251E), so

$$\begin{aligned} \theta(C \cap V) + \theta(C \setminus V) &= \theta(C \cap V) + \theta((C \setminus V) \cap (A \times B)) + \theta((C \setminus V) \setminus (A \times B)) \\ &= \tilde{\theta}(C \cap V) + \tilde{\theta}(C \cap (A \times B) \setminus V) + \theta(C \setminus (A \times B)) \\ &= \tilde{\theta}(C \cap (A \times B)) + \theta(C \setminus (A \times B)) \\ &= \theta(C \cap (A \times B)) + \theta(C \setminus (A \times B)) = \theta C. \end{aligned}$$

As C is arbitrary, $V \in \Lambda$, so $V = V \cap (A \times B)$ belongs to $\Lambda^\#$.

(β) If $A \subseteq \bigcup_{n \in \mathbb{N}} E_n$ and $B \subseteq \bigcup_{n \in \mathbb{N}} F_n$ where all the E_n, F_n are of finite measure, then $V = \bigcup_{m,n \in \mathbb{N}} V \cap (E_m \times F_n) \in \Lambda^\#$, by (c).

(γ) If $\langle X_i \rangle_{i \in I}$, $\langle Y_j \rangle_{j \in J}$ are decompositions of X, Y respectively, then for each $i \in I$, $j \in J$ we have $V \cap (X_i \times Y_j) \in \Lambda^\#$, that is, there is a $W_{ij} \in \Lambda$ such that $V \cap (X_i \times Y_j) = W_{ij} \cap (A \times B)$. Now $\langle X_i \times Y_j \rangle_{(i,j) \in I \times J}$ is a decomposition of $X \times Y$ for λ (251O), so that

$$W = \bigcup_{i \in I, j \in J} W_{ij} \cap (X_i \times Y_j) \in \Lambda,$$

and $V = W \cap (A \times B) \in \Lambda^\#$.

(e) Thus any of the three conditions is sufficient to ensure that $\tilde{\Lambda} = \Lambda^\#$, in which case (a) tells us that $\tilde{\lambda} = \lambda^\#$.

251R Corollary Let r, $s \ge 1$ be integers, and $\phi : \mathbb{R}^r \times \mathbb{R}^s \to \mathbb{R}^{r+s}$ the natural bijection. If $A \subseteq \mathbb{R}^r$ and $B \subseteq \mathbb{R}^s$, then the restriction of ϕ to $A \times B$ identifies the product of Lebesgue measure on A and Lebesgue measure on B with Lebesgue measure on $\phi[A \times B] \subseteq \mathbb{R}^{r+s}$.

Remark Note that by 'Lebesgue measure on A' I mean the subspace measure μ_{rA} on A induced by r-dimensional Lebesgue measure μ_r on $\mathbb{R}^r$, whether or not A is itself a measurable set.

proof By 251Q, using either of the conditions (ii-β) or (ii-γ), the product measure $\tilde{\lambda}$ on $A \times B$ is just the subspace measure $\lambda^\#$ on $A \times B$ induced by the product measure λ on $\mathbb{R}^r \times \mathbb{R}^s$. But by 251N we know that ϕ is an isomorphism between $(\mathbb{R}^r \times \mathbb{R}^s, \lambda)$ and $(\mathbb{R}^{r+s}, \mu_{r+s})$; so it must also identify $\tilde{\lambda}$ with the subspace measure on $\phi[A \times B]$.

251S Corollary Let (X, Σ, μ) and (Y, T, ν) be measure spaces, and λ the c.l.d. product measure on $X \times Y$. If $A \subseteq X$ and $B \subseteq Y$ can be covered by sequences of sets of finite measure, then $\lambda^*(A \times B) = \mu^* A \cdot \nu^* B$.

proof In the language of 251Q,

$$\lambda^*(A \times B) = \lambda^\#(A \times B) = \tilde{\lambda}(A \times B) = \mu_A A \cdot \nu_B B$$

(by 251K and 251E)

$$= \mu^* A \cdot \nu^* B.$$

251T The next proposition gives an idea of how the technical definitions here fit together.

Proposition Let (X,Σ,μ) and (Y,T,ν) be measure spaces. Write $(X,\hat{\Sigma},\hat{\mu})$ and $(X,\tilde{\Sigma},\tilde{\mu})$ for the completion and c.l.d. version of (X,Σ,μ) (212C, 213E). Let λ, $\hat{\lambda}$ and $\tilde{\lambda}$ be the three c.l.d. product measures on $X\times Y$ obtained from the pairs (μ,ν), $(\hat{\mu},\nu)$ and $(\tilde{\mu},\nu)$ of factor measures. Then $\lambda=\hat{\lambda}=\tilde{\lambda}$.

proof Write Λ, $\hat{\Lambda}$ and $\tilde{\Lambda}$ for the domains of λ, $\hat{\lambda}$, $\tilde{\lambda}$ respectively; and θ, $\hat{\theta}$, $\tilde{\theta}$ for the outer measures on $X\times Y$ obtained by the formula of 251A from the three pairs of factor measures.

(a) If $E\in\Sigma$ and $\mu E<\infty$, then θ, $\hat{\theta}$ and $\tilde{\theta}$ agree on subsets of $E\times Y$. **P** Take $A\subseteq E\times Y$ and $\epsilon>0$.

(i) There are sequences $\langle E_n\rangle_{n\in\mathbb{N}}$ in Σ, $\langle F_n\rangle_{n\in\mathbb{N}}$ in T such that $A\subseteq\bigcup_{n\in\mathbb{N}}E_n\times F_n$ and $\sum_{n=0}^\infty\mu E_n\cdot\nu F_n\le\theta A+\epsilon$. Now $\tilde{\mu}E_n\le\mu E_n$ for every n (213Fb), so

$$\tilde{\theta}A\le\sum_{n=0}^\infty\tilde{\mu}E_n\cdot\nu F_n\le\sum_{n=0}^\infty\mu E_n\cdot\nu F_n\le\theta A+\epsilon.$$

(ii) There are sequences $\langle\hat{E}_n\rangle_{n\in\mathbb{N}}$ in $\hat{\Sigma}$, $\langle\hat{F}_n\rangle_{n\in\mathbb{N}}$ in T such that $A\subseteq\bigcup_{n\in\mathbb{N}}\hat{E}_n\times\hat{F}_n$ and $\sum_{n=0}^\infty\hat{\mu}\hat{E}_n\cdot\nu\hat{F}_n\le\hat{\theta}A+\epsilon$. Now for each n there is an $E'_n\in\Sigma$ such that $\hat{E}_n\subseteq E'_n$ and $\mu E'_n=\hat{\mu}\hat{E}_n$, so that

$$\theta A\le\sum_{n=0}^\infty\mu E'_n\cdot\nu\hat{F}_n=\sum_{n=0}^\infty\hat{\mu}\hat{E}_n\cdot\nu\hat{F}_n\le\hat{\theta}A+\epsilon.$$

(iii) There are sequences $\langle\tilde{E}_n\rangle_{n\in\mathbb{N}}$ in $\tilde{\Sigma}$, $\langle\tilde{F}_n\rangle_{n\in\mathbb{N}}$ in T such that $A\subseteq\bigcup_{n\in\mathbb{N}}\tilde{E}_n\times\tilde{F}_n$ and $\sum_{n=0}^\infty\tilde{\mu}\tilde{E}_n\cdot\nu\tilde{F}_n\le\tilde{\theta}A+\epsilon$. Now for each n, $\tilde{E}_n\cap E\in\hat{\Sigma}$, so

$$\hat{\theta}A\le\sum_{n=0}^\infty\hat{\mu}(\tilde{E}_n\cap E)\cdot\nu\tilde{F}_n\le\sum_{n=0}^\infty\tilde{\mu}\tilde{E}_n\cdot\nu\tilde{F}_n\le\tilde{\theta}A+\epsilon.$$

(iv) Since A and ϵ are arbitrary, $\theta=\hat{\theta}=\tilde{\theta}$ on $\mathcal{P}(E\times Y)$. **Q**

(b) Consequently, the outer measures λ^*, $\hat{\lambda}^*$ and $\tilde{\lambda}^*$ are identical. **P** Use 251P. Take $A\subseteq X\times Y$, $E\in\Sigma$, $\hat{E}\in\hat{\Sigma}$, $\tilde{E}\in\tilde{\Sigma}$, $F\in\mathrm{T}$ such that μE, $\hat{\mu}\hat{E}$, $\tilde{\mu}\tilde{E}$ and νF are all finite. Then

(i)

$$\theta(A\cap(E\times F))=\hat{\theta}(A\cap(E\times F))\le\hat{\lambda}^*A,\quad\theta(A\cap(E\times F))=\tilde{\theta}(A\cap(E\times F))\le\tilde{\lambda}^*A$$

because $\hat{\mu}E$ and $\tilde{\mu}E$ are both finite.

(ii) There is an $E'\in\Sigma$ such that $\hat{E}\subseteq E'$ and $\mu E'<\infty$, so that

$$\hat{\theta}(A\cap(\hat{E}\times F))\le\hat{\theta}(A\cap(E'\times F))=\theta(A\cap(E'\times F))\le\lambda^*A.$$

(iii) There is an $E''\in\Sigma$ such that $E''\subseteq\tilde{E}$ and $\tilde{\mu}(\tilde{E}\setminus E'')=0$ (213Fc), so that $\tilde{\theta}((\tilde{E}\setminus E'')\times Y)=0$ and $\mu E''<\infty$; accordingly

$$\tilde{\theta}(A\cap(\tilde{E}\times F))=\tilde{\theta}(A\cap(E''\times F))=\theta(A\cap(E''\times F))\le\lambda^*A.$$

(iv) Taking the suprema over E, $\hat{E}$, $\tilde{E}$ and F, we get

$$\lambda^*A\le\hat{\lambda}^*A,\quad\lambda^*A\le\tilde{\lambda}^*A,\quad\hat{\lambda}^*A\le\lambda^*A,\quad\tilde{\lambda}^*A\le\lambda^*A.$$

As A is arbitrary, $\lambda^*=\hat{\lambda}^*=\tilde{\lambda}^*$. **Q**

(c) Now λ, $\hat{\lambda}$ and $\tilde{\lambda}$ are all complete and locally determined, so by 213C are the measures defined by Carathéodory's method from their own outer measures, and are therefore identical.

251U It is 'obvious', and an easy consequence of theorems so far proved, that the set $\{(x,x):x\in\mathbb{R}\}$ is negligible for Lebesgue measure on $\mathbb{R}^2$. The corresponding result is true in the square of any *atomless* measure space.

Proposition Let (X,Σ,μ) be an atomless measure space, and let λ be the c.l.d. measure on $X\times X$. Then $\Delta=\{(x,x):x\in X\}$ is λ-negligible.

proof Let E, $F \in \Sigma$ be sets of finite measure, and $n \in \mathbb{N}$. Applying 215D repeatedly, we can find a disjoint family $\langle F_i \rangle_{i<n}$ of measurable subsets of F such that $\mu F_i = \frac{\mu F}{n+1}$ for each i; setting $F_n = F \setminus \bigcup_{i<n} F_i$, we also have $\mu F_n = \frac{\mu F}{n+1}$. Now

$$\Delta \cap (E \times F) \subseteq \bigcup_{i \le n} (E \cap F_i) \times F_i,$$

so

$$\lambda^*(\Delta \cap (E \times F)) \le \sum_{i=0}^{n} \mu(E \cap F_i) \cdot \mu F_i = \frac{\mu F}{n+1} \sum_{i=0}^{n} \mu(E \cap F_i) \le \frac{1}{n+1} \mu E \cdot \mu F.$$

As n is arbitrary, $\lambda(\Delta \cap (E \times F)) = 0$; as E and F are arbitrary, $\lambda \Delta = 0$.

***251W Products of more than two spaces** The whole of this section can be repeated for arbitrary finite products. The labour is substantial but no new ideas are required. By the time we need the general construction in any formal way, it should come very naturally, and I do not think it is necessary to work through the next page before proceeding, especially as products of *probability* spaces are dealt with in §254. However, for completeness, and to help locate results when applications do appear, I list them here. They do of course constitute a very instructive set of exercises. The most important fragments are repeated in 251Xe-251Xf.

Let $\langle (X_i, \Sigma_i, \mu_i) \rangle_{i \in I}$ be a finite family of measure spaces, and set $X = \prod_{i \in I} X_i$. Write $\Sigma_i^f = \{E : E \in \Sigma_i,\ \mu_i E < \infty\}$ for each $i \in I$.

(a) For $A \subseteq X$ set

$$\theta A = \inf\{\sum_{n=0}^{\infty} \prod_{i \in I} \mu_i E_{ni} : E_{ni} \in \Sigma_i \ \forall\, i \in I,\ n \in \mathbb{N},\ A \subseteq \bigcup_{n \in \mathbb{N}} \prod_{i \in I} E_{ni}\}.$$

Then θ is an outer measure on X. Let λ_0 be the measure on X derived by Carathéodory's method from θ, and Λ its domain.

(b) If $\langle X_i \rangle_{i \in I}$ is a finite family of sets and Σ_i is a σ-algebra of subsets of X_i for each $i \in I$, then $\widehat{\bigotimes}_{i \in I} \Sigma_i$ is the σ-algebra of subsets of $X = \prod_{i \in I} X_i$ generated by $\{\prod_{i \in I} E_i : E_i \in \Sigma_i$ for every $i \in I\}$. (For the corresponding construction when I is infinite, see 254E.)

(c) $\lambda_0(\prod_{i \in I} E_i)$ is defined and equal to $\prod_{i \in I} \mu_i E_i$ whenever $E_i \in \Sigma_i$ for each $i \in I$.

(d) The **c.l.d. product measure** on X is the measure λ defined by setting

$$\lambda W = \sup\{\lambda_0(W \cap \prod_{i \in I} E_i) : E_i \in \Sigma_i^f \text{ for each } i \in I\}$$

for $W \in \Lambda$. If I is empty, so that $X = \{\emptyset\}$, then the appropriate convention is to set $\lambda X = 1$.

(e) If $H \subseteq X$, then $H \in \Lambda$ iff $H \cap \prod_{i \in I} E_i \in \Lambda$ whenever $E_i \in \Sigma_i^f$ for each $i \in I$.

(f)(i) $\widehat{\bigotimes}_{i \in I} \Sigma_i \subseteq \Lambda$ and $\lambda(\prod_{i \in I} E_i) = \prod_{i \in I} \mu_i E_i$ whenever $E_i \in \Sigma_i^f$ for each i.

(ii) For every $W \in \Lambda$ there is a $V \in \widehat{\bigotimes}_{i \in I} \Sigma_i$ such that $V \subseteq W$ and $\lambda V = \lambda W$.

(iii) λ is complete and locally determined, and is the c.l.d. version of λ_0.

(iv) If $W \in \Lambda$ and $\lambda W > 0$ then there are $E_i \in \Sigma_i^f$, for $i \in I$, such that $\lambda(W \cap \prod_{i \in I} E_i) > 0$.

(v) If $W \in \Lambda$ and $\lambda W < \infty$, then for every $\epsilon > 0$ there are $n \in \mathbb{N}$ and $E_{0i}, \dots, E_{ni} \in \Sigma_i^f$, for each $i \in I$, such that $\lambda(W \triangle \bigcup_{k \le n} \prod_{i \in I} E_{ki}) \le \epsilon$.

(g) If each μ_i is σ-finite, so is λ, and $\lambda = \lambda_0$ is the completion of its restriction to $\widehat{\bigotimes}_{i \in I} \Sigma_i$.

(h) If $\langle I_j \rangle_{j \in J}$ is any partition of I, then λ can be identified with the c.l.d. product of $\langle \lambda_j \rangle_{j \in J}$, where λ_j is the c.l.d. product of $\langle \mu_i \rangle_{i \in I_j}$. (See the arguments in 251N and also in 254N below.)

(i) If $I = \{1,\dots,n\}$ and each μ_i is Lebesgue measure on $\mathbb{R}$, then λ can be identified with Lebesgue measure on $\mathbb{R}^n$.

(j) If, for each $i \in I$, we have a decomposition $\langle X_{ij}\rangle_{j\in J_i}$ of X_i, then $\langle\prod_{i\in I} X_{i,f(i)}\rangle_{f\in\prod_{i\in I} J_i}$ is a decomposition of X.

(k) For any $C \subseteq X$,

$$\lambda^*C = \sup\{\theta(C \cap \prod_{i\in I} E_i) : E_i \in \Sigma_i^f \text{ for every } i \in I\}.$$

(l) Suppose that $A_i \subseteq X_i$ for each $i \in I$. Write $\lambda^\#$ for the subspace measure on $A = \prod_{i\in I} A_i$, and $\tilde\lambda$ for the c.l.d. product of the subspace measures on the A_i. Then $\tilde\lambda$ extends $\lambda^\#$, and if

either $A_i \in \Sigma_i$ for every i

or every A_i can be covered by a sequence of sets of finite measure

or every μ_i is strictly localizable,

then $\tilde\lambda = \lambda^\#$.

(m) If $A_i \subseteq X_i$ can be covered by a sequence of sets of finite measure for each $i \in I$, then $\lambda^*(\prod_{i\in I} A_i) = \prod_{i\in I}\mu_i^* A_i$.

(n) Writing $\hat\mu_i$, $\tilde\mu_i$ for the completion and c.l.d. version of each μ_i, λ is the c.l.d. product of $\langle\hat\mu_i\rangle_{i\in I}$ and also of $\langle\tilde\mu_i\rangle_{i\in I}$.

(o) If all the (X_i, Σ_i, μ_i) are the same atomless measure space (X, Σ, μ), then $\{x : x \in X,\ i \mapsto x(i) \text{ is injective}\}$ is λ-conegligible.

(p) Now suppose that we have another family $\langle(Y_i, \mathrm{T}_i, \nu_i)\rangle_{i\in I}$ of measure spaces, with product (Y, Λ', λ'), and inverse-measure-preserving functions $f_i : X_i \to Y_i$ for each i; define $f : X \to Y$ by saying that $f(x)(i) = f_i(x(i))$ for $x \in X$ and $i \in I$. If all the ν_i are σ-finite, then f is inverse-measure-preserving for λ and λ'.

251X Basic exercises (a) Let X and Y be sets, $\mathcal{A} \subseteq \mathcal{P}X$ and $\mathcal{B} \subseteq \mathcal{P}Y$. Let Σ be the σ-algebra of subsets of X generated by $\mathcal{A}$ and T the σ-algebra of subsets of Y generated by $\mathcal{B}$. Show that $\Sigma\widehat{\otimes}\mathrm{T}$ is the σ-algebra of subsets of $X \times Y$ generated by $\{A \times B : A \in \mathcal{A},\ B \in \mathcal{B}\}$.

(b) Let (X, Σ, μ) and (Y, T, ν) be measure spaces; let λ_0 be the primitive product measure on $X \times Y$, and λ the c.l.d. product measure. Show that $\lambda_0 W < \infty$ iff $\lambda W < \infty$ and W is included in a set of the form

$$(E \times Y) \cup (X \times F) \cup \bigcup_{n\in\mathbb{N}} E_n \times F_n$$

where $\mu E = \nu F = 0$ and $\mu E_n < \infty$, $\nu F_n < \infty$ for every n.

>(c) Show that if X and Y are any sets, with their respective counting measures, then the primitive and c.l.d. product measures on $X \times Y$ are both counting measure on $X \times Y$.

(d) Let (X, Σ, μ) and (Y, T, ν) be measure spaces; let λ_0 be the primitive product measure on $X \times Y$, and λ the c.l.d. product measure. Show that

$$\begin{aligned}&\lambda_0 \text{ is locally determined}\\ &\quad\iff \lambda_0 \text{ is semi-finite}\\ &\quad\iff \lambda_0 = \lambda\\ &\quad\iff \lambda_0 \text{ and } \lambda \text{ have the same negligible sets.}\end{aligned}$$

>(e) (See 251W.) Let $\langle(X_i, \Sigma_i, \mu_i)\rangle_{i\in I}$ be a family of measure spaces, where I is a non-empty finite set. Set $X = \prod_{i\in I} X_i$. For $A \subseteq X$, set

$$\theta A = \inf\{\sum_{n=0}^{\infty} \prod_{i\in I} \mu_i E_{ni} : E_{ni} \in \Sigma_i \ \forall\, n \in \mathbb{N},\, i \in I,\, A \subseteq \bigcup_{n\in\mathbb{N}} \prod_{i\in I} E_{ni}\}.$$

Show that θ is an outer measure on X. Let λ_0 be the measure defined from θ by Carathéodory's method, and for $W \in \operatorname{dom}\lambda_0$ set

$$\lambda W = \sup\{\lambda_0(W \cap \prod_{i\in I} E_i) : E_i \in \Sigma_i,\ \mu_i E_i < \infty \text{ for every } i \in I\}.$$

Show that λ is a measure on X, and is the c.l.d. version of λ_0.

>(f) (See 251W.) Let I be a non-empty finite set and $\langle(X_i, \Sigma_i, \mu_i)\rangle_{i\in I}$ a family of measure spaces. For non-empty $K \subseteq I$ set $X^{(K)} = \prod_{i\in K} X_i$ and let $\lambda_0^{(K)}$, $\lambda^{(K)}$ be the measures on $X^{(K)}$ constructed as in 251Xe. Show that if K is a non-empty proper subset of I, then the natural bijection between $X^{(I)}$ and $X^{(K)} \times X^{(I\setminus K)}$ identifies $\lambda_0^{(I)}$ with the primitive product measure of $\lambda_0^{(K)}$ and $\lambda_0^{(I\setminus K)}$, and $\lambda^{(I)}$ with the c.l.d. product measure of $\lambda^{(K)}$ and $\lambda^{(I\setminus K)}$.

>(g) Using 251Xe-251Xf above, or otherwise, show that if (X_1, Σ_1, μ_1), (X_2, Σ_2, μ_2), (X_3, Σ_3, μ_3) are measure spaces then the primitive and c.l.d. product measures λ_0, λ of $(X_1 \times X_2) \times X_3$, constructed by first taking the appropriate product measure on $X_1 \times X_2$ and then taking the product of this with the measure of X_3, are identified with the corresponding product measures on $X_1 \times (X_2 \times X_3)$ by the canonical bijection between the sets $(X_1 \times X_2) \times X_3$ and $X_1 \times (X_2 \times X_3)$.

(h)(i) What happens in 251Xe when I is a singleton? (ii) Devise an appropriate convention to make 251Xe-251Xf remain valid when one or more of the sets I, K, $I \setminus K$ there is empty.

>(i) Let (X, Σ, μ) be a complete locally determined measure space, and I any non-empty set; let ν be counting measure on I. Show that the c.l.d. product measure on $X \times I$ is equal to (or at any rate identifiable with) the direct sum measure of the family $\langle(X_i, \Sigma_i, \mu_i)\rangle_{i\in I}$, if we set $(X_i, \Sigma_i, \mu_i) = (X, \Sigma, \mu)$ for every i.

>(j) Let $\langle(X_i, \Sigma_i, \mu_i)\rangle_{i\in I}$ be a family of measure spaces, with direct sum (X, Σ, μ) (214L). Let (Y, T, ν) be any measure space, and give $X \times Y$, $X_i \times Y$ their c.l.d. product measures. Show that the natural bijection between $X \times Y$ and $Z = \bigcup_{i\in I}(X_i \times Y) \times \{i\}$ is an isomorphism between the measure of $X \times Y$ and the direct sum measure on Z.

>(k) Let (X, Σ, μ) be any measure space, and Y a singleton set $\{y\}$; let ν be the measure on Y such that $\nu Y = 1$. Show that the natural bijection between $X \times \{y\}$ and X identifies the primitive product measure on $X \times \{y\}$ with $\check\mu$ as defined in 213Xa, and the c.l.d. product measure with the c.l.d. version of μ. Explain how to put this together with 251Xg and 251Ic to prove 251T.

>(l) Let (X, Σ, μ) and (Y, T, ν) be measure spaces, and λ the c.l.d. product measure on $X \times Y$. Show that λ is the c.l.d. version of its restriction to $\Sigma\widehat{\otimes}\mathrm{T}$.

(m) Let (X, Σ, μ) and (Y, T, ν) be measure spaces, with primitive and c.l.d. product measures λ_0, λ. Let λ_1 be any measure with domain $\Sigma\widehat{\otimes}\mathrm{T}$ such that $\lambda_1(E \times F) = \mu E \cdot \nu F$ whenever $E \in \Sigma$ and $F \in \mathrm{T}$. Show that $\lambda W \le \lambda_1 W \le \lambda_0 W$ for every $W \in \Sigma\widehat{\otimes}\mathrm{T}$.

(n) Let (X, Σ, μ) and (Y, T, ν) be two measure spaces, and λ_0 the primitive product measure on $X \times Y$. Show that the corresponding outer measure λ_0^* is just the outer measure θ of 251A.

(o) Let (X, Σ, μ) and (Y, T, ν) be measure spaces, and $A \subseteq X$, $B \subseteq Y$ subsets; write μ_A, ν_B for the subspace measures. Let λ_0 be the primitive product measure on $X \times Y$, and $\lambda_0^{\#}$ the subspace measure it induces on $A \times B$. Let $\tilde\lambda_0$ be the primitive product measure of μ_A and ν_B on $A \times B$. Show that $\tilde\lambda_0$ extends $\lambda_0^{\#}$. Show that if *either* (α) $A \in \Sigma$ and $B \in \mathrm{T}$ *or* (β) A and B can both be covered by sequences of sets of finite measure *or* (γ) μ and ν are both strictly localizable, then $\tilde\lambda_0 = \lambda_0^{\#}$.

(p) In 251Q, show that $\tilde\lambda$ and $\lambda^{\#}$ will have the same null ideals, even if none of the conditions of 251Q(ii) are satisfied.

(q) Let (X,Σ,μ) and (Y,T,ν) be any measure spaces, and λ_0 the primitive product measure on $X\times Y$. Show that $\lambda_0^*(A\times B)=\mu^*A\cdot\nu^*B$ for any $A\subseteq X$ and $B\subseteq Y$.

(r) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and $\hat\mu$ the completion of μ. Show that μ, ν and $\hat\mu$, ν have the same primitive product measures.

(s) Let (X,Σ,μ) be a semi-finite measure space. Show that μ is atomless iff the diagonal $\{(x,x):x\in X\}$ is negligible for the c.l.d. product measure on $X\times X$.

>(t) Let (X,Σ,μ) be an atomless measure space, and (Y,T,ν) any measure space. Show that the c.l.d. product measure on $X\times Y$ is atomless.

>(u) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and λ the c.l.d. product measure on $X\times Y$. (i) Show that if μ and ν are purely atomic, so is λ. (ii) Show that if μ and ν are point-supported, so is λ.

251Y Further exercises (a) Let X, Y be sets with σ-algebras of subsets Σ, T. Suppose that $h:X\times Y\to\mathbb{R}$ is $\Sigma\widehat{\otimes}\mathrm{T}$-measurable and $\phi:X\to Y$ is (Σ,T)-measurable (121Yc). Show that $x\mapsto h(x,\phi(x)):X\to\mathbb{R}$ is Σ-measurable.

(b) Show that there are measure spaces (X_1,Σ_1,μ_1) and (X_2,Σ_2,μ_2), a probability space (Y,T,ν) and an inverse-measure-preserving function $f:X_1\to X_2$ such that $h:X_1\times Y\to X_2\times Y$ is not inverse-measure-preserving for the c.l.d. product measures on $X_1\times Y$ and $X_2\times Y$, where $h(x,y)=(f(x),y)$ for $x\in X_1$ and $y\in Y$.

(c) Let (X,Σ,μ) be a complete locally determined measure space with a subspace A whose measure is not locally determined (see 216Xb). Set $Y=\{0\}$, $\nu Y=1$ and consider the c.l.d. product measures on $X\times Y$ and $A\times Y$; write Λ, $\tilde\Lambda$ for their domains. Show that $\tilde\Lambda$ properly includes $\{W\cap(A\times Y):W\in\Lambda\}$.

(d) Let (X,Σ,μ) be any measure space, (Y,T,ν) an atomless measure space, and $f:X\to Y$ a (Σ,T)-measurable function. Show that $\{(x,f(x)):x\in X\}$ is negligible for the c.l.d. product measure on $X\times Y$.

251 Notes and comments There are real difficulties in deciding which construction to declare as 'the' product of two arbitrary measures. My phrase 'primitive product measure', and notation λ_0, betray a bias; my own preference is for the c.l.d. product λ, for two principal reasons. The first is that λ_0 is likely to be 'bad', in particular, not semi-finite, even if μ and ν are 'good' (251Xd, 252Yk), while λ inherits some of the most important properties of μ and ν (e.g., 251O); the second is that in the case of topological measure spaces X and Y, there is often a canonical topological measure on $X\times Y$, which is likely to be more closely related to λ than to λ_0. But for elucidation of this point I must ask you to wait until §417 in Volume 4.

It would be possible to remove the 'primitive' product measure entirely from the exposition, or at least to relegate it to the exercises. This is indeed what I expect to do in the rest of this treatise, since (in my view) all significant features of product measures on finitely many factors can be expressed in terms of the c.l.d. product measure. For the first introduction to product measures, however, a direct approach to the c.l.d. product measure (through the description of λ^* in 251P, for instance) is an uncomfortably large bite, and I have some sort of duty to present the most natural rival to the c.l.d. product measure prominently enough for you to judge for yourself whether I am right to dismiss it. There certainly are results associated with the primitive product measure (251Xn, 251Xq, 252Yc) which have an agreeable simplicity.

The clash is avoided altogether, of course, if we specialize immediately to σ-finite spaces, in which the two constructions coincide (251K). But even this does not solve all problems. There is a popular alternative measure often called 'the' product measure: the restriction $\lambda_{0\mathcal{B}}$ of λ_0 to the σ-algebra $\Sigma\widehat{\otimes}\mathrm{T}$. (See, for instance, Halmos 50.) The advantage of this is that if a function f on $X\times Y$ is $\Sigma\widehat{\otimes}\mathrm{T}$-measurable, then $x\mapsto f(x,y)$ is Σ-measurable for every $y\in Y$. (This is because

$$\{W:W\subseteq X\times Y,\,\{x:(x,y)\in W\}\in\Sigma\ \forall\ y\in Y\}$$

is a σ-algebra of subsets of $X\times Y$ containing $E\times F$ whenever $E\in\Sigma$ and $F\in\mathrm{T}$, and therefore including $\Sigma\widehat{\otimes}\mathrm{T}$.) The primary objection, to my mind, is that Lebesgue measure on $\mathbb{R}^2$ is no longer 'the' product of

Lebesgue measure on $\mathbb{R}$ with itself. Generally, it is right to seek measures which measure as many sets as possible, and I prefer to face up to the technical problems (which I acknowledge are off-putting) by seeking appropriate definitions on the approach to major theorems, rather than rely on ad hoc fixes when the time comes to apply them.

I omit further examples of product measures for the moment, because the investigation of particular examples will be much easier with the aid of results from the next section. Of course the leading example, and the one which should come always to mind in response to the words 'product measure', is Lebesgue measure on $\mathbb{R}^2$, the case $r = s = 1$ of 251N and 251R. For an indication of what can happen when one of the factors is not σ-finite, you could look ahead to 252K.

I hope that you will see that the definition of the outer measure θ in 251A corresponds to the standard definition of Lebesgue outer measure, with 'measurable rectangles' $E \times F$ taking the place of intervals, and the functional $E \times F \mapsto \mu E \cdot \nu F$ taking the place of 'length' or 'volume' of an interval; moreover, thinking of E and F as intervals, there is an obvious relation between Lebesgue measure on $\mathbb{R}^2$ and the product measure on $\mathbb{R} \times \mathbb{R}$. Of course an 'obvious relationship' is not the same thing as a proper theorem with exact hypotheses and conclusions, but Theorem 251N is clearly central. Long before that, however, there is another parallel between the construction of 251A and that of Lebesgue measure. In both cases, the proof that we have an outer measure comes directly from the defining formula (in 113Yd I gave as an exercise a general result covering 251B), and consequently a very general construction can lead us to a measure. But the measure would be of far less interest and value if it did not measure, and measure correctly, the basic sets, in this case the measurable rectangles. Thus 251E corresponds to the theorem that intervals are Lebesgue measurable, with the right measure (114Db, 114G). This is the real key to the construction, and is one of the fundamental ideas of measure theory.

Yet another parallel is in 251Xn; the outer measure defining the primitive product measure λ_0 is exactly equal to the outer measure defined from λ_0. I described the corresponding phenomenon for Lebesgue measure in 132C.

Any construction which claims the title 'canonical' must satisfy a variety of natural requirements; for instance, one expects the canonical bijection between $X \times Y$ and $Y \times X$ to be an isomorphism between the corresponding product measure spaces. 'Commutativity' of the product in this sense is I think obvious from the definitions in 251A-251C. It is obviously desirable – not, I think, obviously true – that the product should be 'associative' in that the canonical bijection between $(X \times Y) \times Z$ and $X \times (Y \times Z)$ should also be an isomorphism between the corresponding products of product measures. This is in fact valid for both the primitive and c.l.d. product measures (251Wh, 251Xe-251Xg).

Working through the classification of measure spaces presented in §211, we find that the primitive product measure λ_0 of arbitrary factor measures μ, ν is complete, while the c.l.d. product measure λ is always complete and locally determined. λ_0 may not be semi-finite, even if μ and ν are strictly localizable (252Yk); but λ will be strictly localizable if μ and ν are (251O). Of course this is associated with the fact that the c.l.d. product measure is distributive over direct sums (251Xj). If either μ or ν is atomless, so is λ (251Xt). Both λ and λ_0 are σ-finite if μ and ν are (251K). It is possible for both μ and ν to be localizable but λ not (254U).

At least if you have worked through Chapter 21, you have now done enough 'pure' measure theory for this kind of investigation, however straightforward, to raise a good many questions. Apart from direct sums, we also have the constructions of 'completion', 'subspace', 'outer measure' and (in particular) 'c.l.d. version' to integrate into the new ideas; I offer some results in 251T and 251Xk. Concerning subspaces, some possibly surprising difficulties arise. The problem is that the product measure on the product of two subspaces can have a larger domain than one might expect. I give a simple example in 251Yc and a more elaborate one in 254Yg. For strictly localizable spaces, there is no problem (251Q); but no other criterion drawn from the list of properties considered in §251 seems adequate to remove the possibility of a disconcerting phenomenon.

Version of 6.12.07

252 Fubini's theorem

Perhaps the most important feature of the concept of 'product measure' is the fact that we can use it to discuss repeated integrals. In this section I give versions of Fubini's theorem and Tonelli's theorem (252B, 252G) with a variety of corollaries, the most useful ones being versions for σ-finite spaces (252C, 252H). As applications I describe the relationship between integration and measuring ordinate sets (252N) and calculate the r-dimensional volume of a ball in $\mathbb{R}^r$ (252Q, 252Xi). I mention counter-examples showing the difficulties which can arise with non-σ-finite measures and non-integrable functions (252K-252L, 252Xf-252Xg).

252A Repeated integrals Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and f a real-valued function defined on a set $\operatorname{dom} f \subseteq X\times Y$. We can seek to form the **repeated integral**

$$\iint f(x,y)\nu(dy)\mu(dx) = \int\Bigl(\int f(x,y)\nu(dy)\Bigr)\mu(dx),$$

which should be interpreted as follows: set

$$D = \{x : x\in X,\ \int f(x,y)\nu(dy) \text{ is defined in } [-\infty,\infty]\},$$

$$g(x) = \int f(x,y)\nu(dy) \text{ for } x\in D,$$

and then write $\iint f(x,y)\nu(dy)\mu(dx) = \int g(x)\mu(dx)$ if this is defined. Of course the subset of Y on which $y\mapsto f(x,y)$ is defined may vary with x, but it must always be conegligible, as must D.

Similarly, exchanging the roles of X and Y, we can seek a repeated integral

$$\iint f(x,y)\mu(dx)\nu(dy) = \int\Bigl(\int f(x,y)\mu(dx)\Bigr)\nu(dy).$$

The point is that, under appropriate conditions on μ and ν, we can relate these repeated integrals to each other by connecting them both with the integral of f itself with respect to the product measure on $X\times Y$.

As will become apparent shortly, it is essential here to allow oneself to discuss the integral of a function which is not everywhere defined. It is of less importance whether one allows integrands and integrals to take infinite values, but for definiteness let me say that I shall be following the rules of 135F; that is, $\int f = \int f^+ - \int f^-$ provided that f is defined almost everywhere, takes values in $[-\infty,\infty]$ and is virtually measurable, and at most one of $\int f^+$, $\int f^-$ is infinite.

252B Theorem Let (X,Σ,μ) and (Y,T,ν) be measure spaces, with c.l.d. product $(X\times Y,\Lambda,\lambda)$ (251F). Suppose that ν is σ-finite and that μ is *either* strictly localizable *or* complete and locally determined. Let f be a $[-\infty,\infty]$-valued function such that $\int f d\lambda$ is defined in $[-\infty,\infty]$. Then $\iint f(x,y)\nu(dy)\mu(dx)$ is defined and is equal to $\int f d\lambda$.

proof The proof of this result involves substantial technical difficulties. If you have not seen these ideas before, you should almost certainly not go straight to the full generality of the version announced above. I will therefore start by writing out a proof in the case in which both μ and ν are totally finite; this is already lengthy enough. I will present it in such a way that only the central section (part (b) below) needs to be amended in the general case, and then, after completing the proof of the special case, I will give the alternative version of (b) which is required for the full result.

(a) Write $\mathcal{L}$ for the family of $[0,\infty]$-valued functions f such that $\int f d\lambda$ and $\iint f(x,y)\nu(dy)\mu(dx)$ are defined and equal. My aim is to show first that $f\in\mathcal{L}$ whenever f is non-negative and $\int f d\lambda$ is defined, and then to look at differences of functions in $\mathcal{L}$. To prove that enough functions belong to $\mathcal{L}$, my strategy will be to start with 'elementary' functions and work outwards through progressively larger classes. It is most efficient to begin by describing ways of building new members of $\mathcal{L}$ from old, as follows.

(i) $f_1+f_2\in\mathcal{L}$ for all f_1, $f_2\in\mathcal{L}$, and $cf\in\mathcal{L}$ for all $f\in\mathcal{L}$, $c\in[0,\infty[$; this is because

$$\int(f_1+f_2)(x,y)\nu(dy) = \int f_1(x,y)\nu(dy) + \int f_2(x,y)\nu(dy),$$

$$\int(cf)(x,y)\nu(dy) = c\int f(x,y)\nu(dy)$$

whenever the right-hand sides are defined, which we are supposing to be the case for almost every x, so that

$$\begin{aligned}\iint(f_1+f_2)(x,y)\nu(dy)\mu(dx) &= \iint f_1(x,y)\nu(dy)\mu(dx) + \iint f_2(x,y)\nu(dy)\mu(dx)\\ &= \int f_1 d\lambda + \int f_2 d\lambda = \int(f_1+f_2)d\lambda,\end{aligned}$$

$$\iint (cf)(x,y)\nu(dy)\mu(dx) = c\int f(x,y)\nu(dy)\mu(dx) = c\int f d\lambda = \int (cf)d\lambda.$$

(ii) If $\langle f_n\rangle_{n\in\mathbb{N}}$ is a sequence in $\mathcal{L}$ such that $f_n(x,y) \le f_{n+1}(x,y)$ whenever $n \in \mathbb{N}$ and $(x,y) \in \operatorname{dom} f_n \cap \operatorname{dom} f_{n+1}$, then $\sup_{n\in\mathbb{N}} f_n \in \mathcal{L}$. **P** Set $f = \sup_{n\in\mathbb{N}} f_n$; for $x \in X$, $n \in \mathbb{N}$ set $g_n(x) = \int f_n(x,y)\nu(dy)$ when the integral is defined in $[0,\infty]$. Since here I am allowing ∞ as a value of a function, it is natural to regard f as defined on $\bigcap_{n\in\mathbb{N}} \operatorname{dom} f_n$. By B.Levi's theorem, $\int f d\lambda = \sup_{n\in\mathbb{N}} \int f_n d\lambda$; write u for this common value in $[0,\infty]$. Next, because $f_n \le f_{n+1}$ wherever both are defined, $g_n \le g_{n+1}$ wherever both are defined, for each n; we are supposing that $f_n \in \mathcal{L}$, so g_n is defined μ-almost everywhere for each n, and

$$\sup_{n\in\mathbb{N}} \int g_n d\mu = \sup_{n\in\mathbb{N}} \int f_n d\lambda = u.$$

By B.Levi's theorem again, $\int g\, d\mu = u$, where $g = \sup_{n\in\mathbb{N}} g_n$. Now take any $x \in \bigcap_{n\in\mathbb{N}} \operatorname{dom} g_n$, and consider the functions f_{xn} on Y, setting $f_{xn}(y) = f_n(x,y)$ whenever this is defined. Each f_{xn} has an integral in $[0,\infty]$, and $f_{xn}(y) \le f_{x,n+1}(y)$ whenever both are defined, and

$$\sup_{n\in\mathbb{N}} \int f_{xn} d\nu = g(x);$$

so, using B.Levi's theorem for a third time, $\int(\sup_{n\in\mathbb{N}} f_{xn})d\nu$ is defined and equal to $g(x)$, that is,

$$\int f(x,y)\nu(dy) = g(x).$$

This is true for almost every x, so

$$\iint f(x,y)\nu(dy)\mu(dx) = \int g\, d\mu = u = \int f d\lambda.$$

Thus $f \in \mathcal{L}$, as claimed. **Q**

(iii) The expression of the ideas in the next section of the proof will go more smoothly if I introduce another term. Write $\mathcal{W}$ for $\{W : W \subseteq X \times Y,\ \chi W \in \mathcal{L}\}$. Then

(α) if W, $W' \in \mathcal{W}$ and $W \cap W' = \emptyset$, $W \cup W' \in \mathcal{W}$

by (i), because $\chi(W \cup W') = \chi W + \chi W'$,

(β) $\bigcup_{n\in\mathbb{N}} W_n \in \mathcal{W}$ whenever $\langle W_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in $\mathcal{W}$

because $\langle \chi W_n\rangle_{n\in\mathbb{N}} \uparrow \chi W$, and we can use (ii).

It is also helpful to note that, for any $W \subseteq X \times Y$ and any $x \in X$, $\int \chi W(x,y)\nu(dy) = \nu W[\{x\}]$, at least whenever $W[\{x\}] = \{y : (x,y) \in W\}$ is measured by ν. Moreover, because λ is complete, a set $W \subseteq X \times Y$ belongs to Λ iff χW is λ-virtually measurable iff $\int \chi W\, d\lambda$ is defined in $[0,\infty]$, and in this case $\lambda W = \int \chi W\, d\lambda$.

(iv) Finally, we need to observe that, in appropriate circumstances, the difference of two members of $\mathcal{W}$ will belong to $\mathcal{W}$: if W, $W' \in \mathcal{W}$ and $W \subseteq W'$ and $\lambda W' < \infty$, then $W' \setminus W \in \mathcal{W}$. **P** We are supposing that $g(x) = \int \chi W(x,y)\nu(dy)$ and $g'(x) = \int \chi W'(x,y)\nu(dy)$ are defined for almost every x, and that $\int g\, d\mu = \lambda W$, $\int g' d\mu = \lambda W'$. Because $\lambda W'$ is finite, g' must be finite almost everywhere, and $D = \{x : x \in \operatorname{dom} g \cap \operatorname{dom} g',\ g'(x) < \infty\}$ is conegligible. Now, for any $x \in D$, both $g(x)$ and $g'(x)$ are finite, so

$$y \mapsto \chi(W' \setminus W)(x,y) = \chi W'(x,y) - \chi W(x,y)$$

is the difference of two integrable functions, and

$$\begin{aligned}\int \chi(W' \setminus W)(x,y)\nu(dy) &= \int \chi W'(x,y) - \chi W(x,y)\nu(dy)\\ &= \int \chi W'(x,y)\nu(dy) - \int \chi W(x,y)\nu(dy) = g'(x) - g(x).\end{aligned}$$

Accordingly

$$\iint \chi(W' \setminus W)(x,y)\nu(dy)\mu(dx) = \int g'(x) - g(x)\mu(dx) = \lambda W' - \lambda W = \lambda(W' \setminus W),$$

and $W' \setminus W$ belongs to $\mathcal{W}$. **Q**

(Of course the argument just above can be shortened by a few words if we allow ourselves to assume that μ and ν are totally finite, since then $g(x)$ and $g'(x)$ will be finite whenever they are defined; but the key idea, that the difference of integrable functions is integrable, is unchanged.)

(b) Now let us examine the class $\mathcal{W}$, assuming that μ and ν are totally finite.

(i) $E \times F \in \mathcal{W}$ for all $E \in \Sigma$, $F \in \mathrm{T}$. **P** $\lambda(E \times F) = \mu E \cdot \nu F$ (251J), and

$$\int \chi(E \times F)(x, y)\nu(dy) = \nu F\, \chi E(x)$$

for each x, so

$$\iint \chi(E \times F)(x, y)\nu(dy)\mu(dx) = \int (\nu F\, \chi E(x))\mu(dx) = \mu E \cdot \nu F$$
$$= \lambda(E \times F) = \int \chi(E \times F) d\lambda. \text{ **Q**}$$

(ii) Let $\mathcal{C}$ be $\{E \times F : E \in \Sigma, F \in \mathrm{T}\}$. Then $\mathcal{C}$ is closed under finite intersections (because $(E \times F) \cap (E' \times F') = (E \cap E') \times (F \cap F')$) and is included in $\mathcal{W}$. In particular, $X \times Y \in \mathcal{W}$. But this, together with (a-iv) and (a-iii-β) above, means that $\mathcal{W}$ is a Dynkin class (definition: 136A), so includes the σ-algebra of subsets of $X \times Y$ generated by $\mathcal{C}$, by the Monotone Class Theorem (136B); that is, $\mathcal{W} \supseteq \Sigma \widehat{\otimes} \mathrm{T}$ (definition: 251D).

(iii) Next, $W \in \mathcal{W}$ whenever $W \subseteq X \times Y$ is λ-negligible. **P** By 251Ib, there is a $V \in \Sigma \widehat{\otimes} \mathrm{T}$ such that $V \subseteq (X \times Y) \setminus W$ and $\lambda V = \lambda((X \times Y) \setminus W)$. Because $\lambda(X \times Y) = \mu X \cdot \nu Y$ is finite, $V' = (X \times Y) \setminus V$ is λ-negligible, and we have $W \subseteq V' \in \Sigma \widehat{\otimes} \mathrm{T}$. Consequently

$$0 = \lambda V' = \iint \chi V'(x, y)\nu(dy)\mu(dx).$$

But this means that

$$D = \{x : \int \chi V'(x, y)\nu(dy) \text{ is defined and equal to } 0\}$$

is conegligible. If $x \in D$, then we must have $\chi V'(x, y) = 0$ for ν-almost every y, that is, $V'[\{x\}]$ is negligible; in which case $W[\{x\}] \subseteq V'[\{x\}]$ also is negligible, and $\int \chi W(x, y)\nu(dy) = 0$. And this is true for every $x \in D$, so $\int \chi W(x, y)\nu(dy)$ is defined and equal to 0 for almost every x, and

$$\iint \chi W(x, y)\nu(dy)\mu(dx) = 0 = \lambda W,$$

as required. **Q**

(iv) It follows that $\Lambda \subseteq \mathcal{W}$. **P** If $W \in \Lambda$, then, by 251Ib again, there is a $V \in \Sigma \widehat{\otimes} \mathrm{T}$ such that $V \subseteq W$ and $\lambda V = \lambda W$, so that $\lambda(W \setminus V) = 0$. Now $V \in \mathcal{W}$ by (ii) and $W \setminus V \in \mathcal{W}$ by (iii), so $W \in \mathcal{W}$ by (a-iii-α). **Q**

(c) I return to the class $\mathcal{L}$.

(i) If $f \in \mathcal{L}$ and g is a $[0, \infty]$-valued function defined and equal to f λ-a.e., then $g \in \mathcal{L}$. **P** Set

$$W = (X \times Y) \setminus \{(x, y) : (x, y) \in \operatorname{dom} f \cap \operatorname{dom} g,\ f(x, y) = g(x, y)\},$$

so that $\lambda W = 0$. (Remember that λ is complete.) By (b), $\iint \chi W(x, y)\nu(dy)\mu(dx) = 0$, that is, $W[\{x\}]$ is ν-negligible for μ-almost every x. Let D be $\{x : x \in X, W[\{x\}] \text{ is } \nu\text{-negligible}\}$. Then D is μ-conegligible. If $x \in D$, then

$$W[\{x\}] = Y \setminus \{y : (x, y) \in \operatorname{dom} f \cap \operatorname{dom} g,\ f(x, y) = g(x, y)\}$$

is negligible, so that $\int f(x, y)\nu(dy) = \int g(x, y)\nu(dy)$ if either is defined. Thus the functions

$$x \mapsto \int f(x, y)\nu(dy), \quad x \mapsto \int g(x, y)\nu(dy)$$

are equal almost everywhere, and

$$\iint g(x, y)\nu(dy)\mu(dx) = \iint f(x, y)\nu(dy)\mu(dx) = \int f d\lambda = \int g\, d\lambda,$$

so that $g \in \mathcal{L}$. **Q**

(ii) Now let f be any non-negative function such that $\int f d\lambda$ is defined in $[0,\infty]$. Then $f\in\mathcal{L}$. **P** For k, $n\in\mathbb{N}$ set

$$W_{nk}=\{(x,y):(x,y)\in\operatorname{dom}f,\ f(x,y)\ge 2^{-n}k\}.$$

Because λ is complete and f is λ-virtually measurable and $\operatorname{dom}f$ is conegligible, every W_{nk} belongs to Λ, so $\chi W_{nk}\in\mathcal{L}$, by (b). Set $f_n=\sum_{k=1}^{4^n}2^{-n}\chi W_{nk}$, so that

$$\begin{aligned} f_n(x,y)&=2^{-n}k \text{ if } k\le 4^n \text{ and } 2^{-n}k\le f(x,y)<2^{-n}(k+1),\\ &=2^n \text{ if } f(x,y)\ge 2^n,\\ &=0 \text{ if } (x,y)\in(X\times Y)\setminus\operatorname{dom}f. \end{aligned}$$

By (a-i), $f_n\in\mathcal{L}$ for every $n\in\mathbb{N}$, while $\langle f_n\rangle_{n\in\mathbb{N}}$ is non-decreasing, so $f'=\sup_{n\in\mathbb{N}}f_n\in\mathcal{L}$, by (a-ii). But $f=_{\text{a.e.}}f'$, so $f\in\mathcal{L}$, by (i) just above. **Q**

(iii) Finally, let f be any $[-\infty,\infty]$-valued function such that $\int f d\lambda$ is defined in $[-\infty,\infty]$. Then $\int f^+d\lambda$, $\int f^-d\lambda$ are both defined and at most one is infinite. By (ii), both f^+ and f^- belong to $\mathcal{L}$. Set $g(x)=\int f^+(x,y)\nu(dy)$, $h(x)=\int f^-(x,y)\nu(dy)$ whenever these are defined; then $\int g\,d\mu=\int f^+d\lambda$ and $\int h\,d\mu=\int f^-d\lambda$ are both defined in $[0,\infty]$.

Suppose first that $\int f^-d\lambda$ is finite. Then $\int h\,d\mu$ is finite, so h must be finite μ-almost everywhere; set

$$D=\{x:x\in\operatorname{dom}g\cap\operatorname{dom}h,\ h(x)<\infty\}.$$

For any $x\in D$, $\int f^+(x,y)\nu(dy)$ and $\int f^-(x,y)\nu(dy)$ are defined in $[0,\infty]$, and the latter is finite; so

$$\int f(x,y)\nu(dy)=\int f^+(x,y)\nu(dy)-\int f^-(x,y)\nu(dy)=g(x)-h(x)$$

is defined in $]-\infty,\infty]$. Because D is conegligible,

$$\begin{aligned}\iint f(x,y)\nu(dy)\mu(dx)&=\int g(x)-h(x)\mu(dx)=\int g\,d\mu-\int h\,d\mu\\ &=\int f^+d\lambda-\int f^-d\lambda=\int fd\lambda,\end{aligned}$$

as required.

Thus we have the result when $\int f^-d\lambda$ is finite. Similarly, or by applying the argument above to $-f$, we see that $\iint f(x,y)\nu(dy)\mu(dx)=\int fd\lambda$ if $\int f^+d\lambda$ is finite.

Thus the theorem is proved, at least when μ and ν are totally finite.

(b*) The only point in the argument above where we needed to know anything special about the measures μ and ν was in part (b), when showing that $\Lambda\subseteq\mathcal{W}$. I now return to this point under the hypotheses of the theorem as stated, that ν is σ-finite and μ is either strictly localizable or complete and locally determined.

(i) It will be helpful to note that the completion $\hat\mu$ of μ (212C) is identical with its c.l.d. version $\tilde\mu$ (213E). **P** If μ is strictly localizable, then $\hat\mu=\tilde\mu$ by 213Ha. If μ is complete and locally determined, then $\hat\mu=\mu=\tilde\mu$ (212D, 213Hf). **Q**

(ii) Write $\Sigma^f=\{G:G\in\Sigma,\ \mu G<\infty\}$, $\mathrm{T}^f=\{H:H\in\mathrm{T},\ \nu H<\infty\}$. For $G\in\Sigma^f$, $H\in\mathrm{T}^f$ let μ_G, ν_H and $\lambda_{G\times H}$ be the subspace measures on G, H and $G\times H$ respectively; then $\lambda_{G\times H}$ is the c.l.d. product measure of μ_G and ν_H (251Q(ii-α)). Now $W\cap(G\times H)\in\mathcal{W}$ for every $W\in\Lambda$. **P** $W\cap(G\times H)$ belongs to the domain of $\lambda_{G\times H}$, so by (b) of this proof, applied to the totally finite measures μ_G and ν_H,

$$\begin{aligned}\lambda(W\cap(G\times H))&=\lambda_{G\times H}(W\cap(G\times H))\\ &=\int_G\int_H\chi(W\cap(G\times H))(x,y)\nu_H(dy)\mu_G(dx)\\ &=\int_G\int_Y\chi(W\cap(G\times H))(x,y)\nu(dy)\mu_G(dx)\end{aligned}$$

(because $\chi(W\cap(G\times H))(x,y)=0$ if $y\in Y\setminus H$, so we can use 131E)

$$= \int_X \int_Y \chi(W \cap (G \times H))(x,y)\nu(dy)\mu(dx)$$

by 131E again, because $\int_Y \chi(W \cap (G \times H))(x,y)\nu(dy) = 0$ if $x \in X \setminus G$. So $W \cap (G \times H) \in \mathcal{W}$. **Q**

(iii) In fact, $W \in \mathcal{W}$ for every $W \in \Lambda$. **P** Remember that we are supposing that ν is σ-finite. Let $\langle Y_n \rangle_{n \in \mathbb{N}}$ be a non-decreasing sequence in T^f covering Y, and for each $n \in \mathbb{N}$ set $W_n = W \cap (X \times Y_n)$, $g_n(x) = \int \chi W_n(x,y)\nu(dy)$ whenever this is defined. For any $G \in \Sigma^f$,

$$\int_G g_n d\mu = \iint \chi(W \cap (G \times Y_n))(x,y)\nu(dy)\mu(dx)$$

is defined and equal to $\lambda(W \cap (G \times Y_n))$, by (ii). But this means, first, that $G \setminus \operatorname{dom} g_n$ is negligible, that is, that $\hat\mu(G \setminus \operatorname{dom} g_n) = 0$. Since this is so whenever μG is finite, $\tilde\mu(X \setminus \operatorname{dom} g_n) = 0$, and g_n is defined $\tilde\mu$-a.e.; but $\tilde\mu = \hat\mu$, so g_n is defined $\hat\mu$-a.e., that is, μ-a.e. (212Eb). Next, if we set $E_{na} = \{x : x \in \operatorname{dom} g_n,\ g_n(x) \ge a\}$ for $a \in \mathbb{R}$, then $E_{na} \cap G \in \hat\Sigma$ whenever $G \in \Sigma^f$, where $\hat\Sigma$ is the domain of $\hat\mu$; by the definition in 213D, E_{na} is measured by $\tilde\mu = \hat\mu$. As a is arbitrary, g_n is μ-virtually measurable (212Fa).

We can therefore speak of $\int g_n d\mu$. Now

$$\iint \chi W_n(x,y)\nu(dy)\mu(dx) = \int g_n d\mu = \sup_{G \in \Sigma^f} \int_G g_n$$

(213B, because μ is semi-finite)

$$= \sup_{G \in \Sigma^f} \lambda(W \cap (G \times Y_n)) = \lambda(W \cap (X \times Y_n))$$

by the definition in 251F. Thus $W \cap (X \times Y_n) \in \mathcal{W}$.

This is true for every $n \in \mathbb{N}$. Because $\langle Y_n \rangle_{n \in \mathbb{N}} \uparrow Y$, $W \in \mathcal{W}$, by (a-iii-β). **Q**

(iv) We can therefore return to part (c) of the argument above and conclude as before.

252C The theorem above is of course asymmetric, in that different hypotheses are imposed on the two factor measures μ and ν. If we want a 'symmetric' theorem we have to suppose that they are both σ-finite, as follows.

Corollary Let (X, Σ, μ) and (Y, T, ν) be two σ-finite measure spaces, and λ the c.l.d. product measure on $X \times Y$. If f is λ-integrable, then $\iint f(x,y)\nu(dy)\mu(dx)$ and $\iint f(x,y)\mu(dx)\nu(dy)$ are defined, finite and equal.

proof Since μ and ν are surely strictly localizable (211Lc), we can apply 252B from either side to conclude that

$$\iint f(x,y)\nu(dy)\mu(dx) = \int f\, d\lambda = \iint f(x,y)\mu(dx)\nu(dy).$$

252D So many applications of Fubini's theorem are to indicator functions that I take a few lines to spell out the form which 252B takes in this case, as in parts (b)-(b*) of the proof there.

Corollary Let (X, Σ, μ) and (Y, T, ν) be measure spaces and λ the c.l.d. product measure on $X \times Y$. Suppose that ν is σ-finite and that μ is either strictly localizable or complete and locally determined.

(i) If $W \in \operatorname{dom} \lambda$, then $\int \nu^* W[\{x\}]\mu(dx)$ is defined in $[0, \infty]$ and equal to λW.

(ii) If ν is complete, we can write $\int \nu W[\{x\}]\mu(dx)$ in place of $\int \nu^* W[\{x\}]\mu(dx)$.

proof The point is just that $\int \chi W(x,y)\nu(dy) = \hat\nu W[\{x\}]$ whenever either is defined, where $\hat\nu$ is the completion of ν (212F). Now 252B tells us that

$$\lambda W = \iint \chi W(x,y)\nu(dy)\mu(dx) = \int \hat\nu W[\{x\}]\mu(dx).$$

We always have $\hat\nu W[\{x\}] = \nu^* W[\{x\}]$, by the definition of $\hat\nu$ (212C); and if ν is complete, then $\hat\nu = \nu$ so $\lambda W = \int \nu W[\{x\}]\mu(dx)$.

252E Corollary Let (X,Σ,μ) and (Y,T,ν) be measure spaces, with c.l.d. product $(X\times Y,\Lambda,\lambda)$. Suppose that ν is σ-finite and that μ has locally determined negligible sets (213I). Then if f is a Λ-measurable real-valued function defined on a subset of $X\times Y$, $y\mapsto f(x,y)$ is ν-virtually measurable for μ-almost every $x\in X$.

proof Let $\tilde f$ be a Λ-measurable extension of f to a real-valued function defined everywhere in $X\times Y$ (121I), and set $\tilde f_x(y)=\tilde f(x,y)$ for all $x\in X$, $y\in Y$,

$$D=\{x:x\in X,\ \tilde f_x \text{ is } \nu\text{-virtually measurable}\}.$$

If $G\in\Sigma$ and $\mu G<\infty$, then $G\setminus D$ is negligible. **P** Let $\langle Y_n\rangle_{n\in\mathbb{N}}$ be a non-decreasing sequence of sets of finite measure covering Y respectively, and set

$$\begin{aligned}\tilde f_n(x,y)&=\tilde f(x,y) \text{ if } x\in G,\ y\in Y_n \text{ and } |\tilde f(x,y)|\le n,\\ &=0 \text{ for other } x\in X\times Y.\end{aligned}$$

Then each $\tilde f_n$ is λ-integrable, being bounded and Λ-measurable and zero off $G\times Y_n$. Consequently, setting $\tilde f_{nx}(y)=\tilde f_n(x,y)$,

$$\int(\int \tilde f_{nx}d\nu)\mu(dx) \text{ exists} =\int \tilde f_n d\lambda.$$

But this surely means that $\tilde f_{nx}$ is ν-integrable, therefore ν-virtually measurable, for almost every $x\in X$. Set

$$D_n=\{x:x\in X,\ \tilde f_{nx} \text{ is } \nu\text{-virtually measurable}\};$$

then every D_n is μ-conegligible, so $\bigcap_{n\in\mathbb{N}}D_n$ is conegligible. But for any $x\in G\cap\bigcap_{n\in\mathbb{N}}D_n$, $\tilde f_x=\lim_{n\to\infty}\tilde f_{nx}$ is ν-virtually measurable. Thus $G\setminus D\subseteq X\setminus\bigcap_{n\in\mathbb{N}}D_n$ is negligible. **Q**

This is true whenever $\mu G<\infty$. As G is arbitrary and μ has locally determined negligible sets, D is conegligible. But, for any $x\in D$, $y\mapsto f(x,y)$ is a restriction of $\tilde f_x$ and must be ν-virtually measurable.

252F As a further corollary we can get some useful information about the c.l.d. product measure for arbitrary measure spaces.

Corollary Let (X,Σ,μ) and (Y,T,ν) be two measure spaces, λ the c.l.d. product measure on $X\times Y$, and Λ its domain. Let $W\in\Lambda$ be such that the vertical section $W[\{x\}]$ is ν-negligible for μ-almost every $x\in X$. Then $\lambda W=0$.

proof Take $E\in\Sigma$, $F\in\mathrm{T}$ of finite measure. Let $\lambda_{E\times F}$ be the subspace measure on $E\times F$. By 251Q(ii-α) again, this is just the product of the subspace measures μ_E and ν_F. We know that $W\cap(E\times F)$ is measured by $\lambda_{E\times F}$. At the same time, the vertical section $(W\cap(E\times F))[\{x\}]=W[\{x\}]\cap F$ is ν_F-negligible for μ_E-almost every $x\in E$. Applying 252B to μ_E and ν_F and $\chi(W\cap(E\times F))$,

$$\lambda(W\cap(E\times F))=\lambda_{E\times F}(W\cap(E\times F))=\int_E\nu_F(W[\{x\}]\cap F)\mu_E(dx)=0.$$

But looking at the definition in 251F, we see that this means that $\lambda W=0$, as claimed.

252G Theorem 252B and its corollaries depend on the factor measures μ and ν belonging to restricted classes. There is a partial result which applies to all c.l.d. product measures, as follows.

Tonelli's theorem Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and $(X\times Y,\Lambda,\lambda)$ their c.l.d. product. Let f be a Λ-measurable $[-\infty,\infty]$-valued function defined on a member of Λ, and suppose that either $\iint|f(x,y)|\mu(dx)\nu(dy)$ or $\iint|f(x,y)|\nu(dy)\mu(dx)$ exists in $\mathbb{R}$. Then f is λ-integrable.

proof Because the construction of the product measure is symmetric in the two factors, it is enough to consider the case in which $\iint|f(x,y)|\nu(dy)\mu(dx)$ is defined and finite, as the same ideas will surely deal with the other case also.

(a) The first step is to check that f is defined and finite λ-a.e. **P** Set $W=\{(x,y):(x,y)\in\operatorname{dom}f,\ f(x,y)$ is finite$\}$. Then $W\in\Lambda$. The hypothesis

$$\iint|f(x,y)|\nu(dy)\mu(dx) \text{ is defined and finite}$$

includes the assertion

$$\int |f(x,y)|\nu(dy) \text{ is defined and finite for } \mu\text{-almost every } x,$$

which implies that

$$\text{for } \mu\text{-almost every } x,\ f(x,y) \text{ is defined and finite for } \nu\text{-almost every } y;$$

that is, that

$$\text{for } \mu\text{-almost every } x,\ W[\{x\}] \text{ is } \nu\text{-conegligible.}$$

But by 252F this implies that $(X\times Y)\setminus W$ is λ-negligible, as required. **Q**

(b) Let h be any non-negative λ-simple function such that $h\le|f|$ λ-a.e. Then $\int h$ cannot be greater than $\iint |f(x,y)|\nu(dy)\mu(dx)$. **P** Set

$$W=\{(x,y):(x,y)\in\operatorname{dom}f,\ h(x,y)\le|f(x,y)|\},\quad h'=h\times\chi W;$$

then h' is a simple function and $h'=_{\text{a.e.}}h$. Express h' as $\sum_{i=0}^n a_i\chi W_i$ where $a_i\ge 0$ and $\lambda W_i<\infty$ for each i. Let $\epsilon>0$. For each $i\le n$ there are $E_i\in\Sigma$, $F_i\in\mathrm{T}$ such that $\mu E_i<\infty$, $\nu F_i<\infty$ and $\lambda(W_i\cap(E_i\times F_i))\ge\lambda W_i-\epsilon$. Set $E=\bigcup_{i\le n}E_i$ and $F=\bigcup_{i\le n}F_i$. Consider the subspace measures μ_E and ν_F and their product $\lambda_{E\times F}$ on $E\times F$; then $\lambda_{E\times F}$ is the subspace measure on $E\times F$ defined from λ (251Q(ii-α) once more). Accordingly, applying 252B to the product $\mu_E\times\nu_F$,

$$\int_{E\times F}h'\,d\lambda=\int_{E\times F}h'\,d\lambda_{E\times F}=\int_E\int_F h'(x,y)\nu_F(dy)\mu_E(dx).$$

For any x, we know that $h'(x,y)\le|f(x,y)|$ whenever $f(x,y)$ is defined. So we can be sure that

$$\int_F h'(x,y)\nu_F(dy)=\int h'(x,y)\chi F(y)\nu(dy)\le\int|f(x,y)|\nu(dy)$$

at least whenever $\int_F h'(x,y)\nu_F(dy)$ and $\int|f(x,y)|\nu(dy)$ are both defined, which is the case for almost every $x\in E$. Consequently

$$\begin{aligned}\int_{E\times F}h'\,d\lambda&=\int_E\int_F h'(x,y)\nu_F(dy)\mu_E(dx)\\&\le\int_E\int|f(x,y)|\nu(dy)\mu(dx)\le\iint|f(x,y)|\nu(dy)\mu(dx).\end{aligned}$$

On the other hand,

$$\begin{aligned}\int h'\,d\lambda-\int_{E\times F}h'\,d\lambda&=\sum_{i=0}^n a_i\lambda(W_i\setminus(E\times F))\\&\le\sum_{i=0}^n a_i\lambda(W_i\setminus(E_i\times F_i))\le\epsilon\sum_{i=0}^n a_i.\end{aligned}$$

So

$$\int h\,d\lambda=\int h'd\lambda\le\iint|f(x,y)|\nu(dy)\mu(dx)+\epsilon\sum_{i=0}^n a_i.$$

As ϵ is arbitrary, $\int h\,d\lambda\le\iint|f(x,y)|\nu(dy)\mu(dx)$, as claimed. **Q**

(c) This is true whenever h is a λ-simple function less than or equal to $|f|$ λ-a.e. But $|f|$ is Λ-measurable and λ is semi-finite (251Ic), so this is enough to ensure that $|f|$ is λ-integrable (213B), which (because f is supposed to be Λ-measurable) in turn implies that f is λ-integrable.

252H Corollary Let (X,Σ,μ) and (Y,T,ν) be σ-finite measure spaces, λ the c.l.d. product measure on $X\times Y$, and Λ its domain.

(a) Let f be a Λ-measurable $[-\infty,\infty]$-valued function defined on a member of Λ. Then if one of

$$\int_{X\times Y}|f(x,y)|\lambda(d(x,y)),\quad\int_Y\int_X|f(x,y)|\mu(dx)\nu(dy),\quad\int_X\int_Y|f(x,y)|\nu(dy)\mu(dx)$$

exists in $\mathbb{R}$, so do the other two, and in this case

$$\int_{X\times Y} f(x,y)\lambda(d(x,y)) = \int_Y \int_X f(x,y)\mu(dx)\nu(dy) = \int_X \int_Y f(x,y)\nu(dy)\mu(dx).$$

(b) Let f be a Λ-measurable $[0,\infty]$-valued function defined on a member of Λ. Then

$$\int_{X\times Y} f(x,y)\lambda(d(x,y)) = \int_Y \int_X f(x,y)\mu(dx)\nu(dy) = \int_X \int_Y f(x,y)\nu(dy)\mu(dx)$$

in the sense that if one of the integrals is defined in $[0,\infty]$ so are the other two, and all three are then equal.

proof (a)(i) Suppose that $\int|f|d\lambda$ is finite. Because both μ and ν are σ-finite, 252B tells us that

$$\iint |f(x,y)|\mu(dx)\nu(dy), \quad \iint |f(x,y)|\nu(dy)\mu(dx)$$

both exist and are equal to $\int|f|d\lambda$, while

$$\iint f(x,y)\mu(dx)\nu(dy), \quad \iint f(x,y)\nu(dy)\mu(dx)$$

both exist and are equal to $\int f d\lambda$.

(ii) Now suppose that $\iint |f(x,y)|\nu(dy)\mu(dx)$ exists in $\mathbb{R}$. Then 252G tells us that $|f|$ is λ-integrable, so we can use (i) to complete the argument. Exchanging the coordinates, the same argument applies if $\iint |f(x,y)|\mu(dx)\nu(dy)$ exists in $\mathbb{R}$.

(b)(i) If $\int f d\lambda$ is defined, the result is immediate from 252B.

(ii) Suppose that $\iint_X f(x,y)\nu(dy)\mu(dx)$ is defined. As in part (a) of the proof of 252G, but this time setting $W = \operatorname{dom} f$, we see that $W \in \Lambda$ and that $W[\{x\}]$ is ν-conegligible for μ-almost every x, so that W is λ-conegligible. Since f is non-negative, Λ-measurable and defined almost everywhere, $\int f d\lambda$ is defined in $[0,\infty]$ and we are in case (i).

252I Corollary Let (X,Σ,μ) and (Y,T,ν) be measure spaces, λ the c.l.d. product measure on $X\times Y$, and Λ its domain. Take $W\in\Lambda$. If either of the integrals

$$\int \mu^* W^{-1}[\{y\}]\nu(dy), \quad \int \nu^* W[\{x\}]\mu(dx)$$

exists and is finite, then $\lambda W < \infty$.

proof Apply 252G with $f = \chi W$, remembering that

$$\mu^* W^{-1}[\{y\}] = \int \chi W(x,y)\mu(dx), \quad \nu^* W[\{x\}] = \int \chi W(x,y)\nu(dy)$$

whenever the integrals are defined, as in the proof of 252D.

252J Remarks 252H is the basic form of Fubini's theorem; it is not a coincidence that most authors avoid non-σ-finite spaces in this context. The next two examples exhibit some of the difficulties which can arise if we leave the familiar territory of more-or-less Borel measurable functions on σ-finite spaces. The first is a classic.

252K Example Let (X,Σ,μ) be $[0,1]$ with Lebesgue measure, and let (Y,T,ν) be $[0,1]$ with counting measure.

(a) Consider the set

$$W = \{(t,t) : t\in[0,1]\} \subseteq X\times Y.$$

We observe that W is expressible as

$$\bigcap_{n\in\mathbb{N}} \bigcup_{k=0}^{n} [\tfrac{k}{n+1}, \tfrac{k+1}{n+1}] \times [\tfrac{k}{n+1}, \tfrac{k+1}{n+1}] \in \Sigma\widehat{\otimes}\mathrm{T}.$$

If we look at the sections

$$W^{-1}[\{t\}] = W[\{t\}] = \{t\}$$

for $t\in[0,1]$, we have

$$\iint \chi W(x,y)\mu(dx)\nu(dy) = \int \mu W^{-1}[\{y\}]\nu(dy) = \int 0\,\nu(dy) = 0,$$

$$\iint \chi W(x,y)\nu(dy)\mu(dx) = \int \nu W[\{x\}]\mu(dx) = \int 1\,\mu(dx) = 1,$$

so the two repeated integrals differ. It is therefore not generally possible to reverse the order of repeated integration, even for a non-negative measurable function in which both repeated integrals exist and are finite.

(b) Because the set W of part (a) actually belongs to $\Sigma\widehat{\otimes}\mathrm{T}$, we know that it is measured by the c.l.d. product measure λ, and 252F (applied with the coordinates reversed) tells us that $\lambda W = 0$.

(c) It is in fact easy to give a full description of λ.

(i) The point is that a set $W \subseteq [0,1] \times [0,1]$ belongs to the domain Λ of λ iff every horizontal section $W^{-1}[\{y\}]$ is Lebesgue measurable. **P** (α) If $W \in \Lambda$, then, for every $b \in [0,1]$, $\lambda([0,1] \times \{b\})$ is finite, so $W \cap ([0,1] \times \{b\})$ is a set of finite measure, and

$$\lambda(W \cap ([0,1] \times \{b\})) = \int \mu(W \cap ([0,1] \times \{b\}))^{-1}[\{y\}]\nu(dy) = \mu W^{-1}[\{b\}]$$

by 252D, because μ is σ-finite, ν is both strictly localizable and complete and locally determined, and

$$\begin{aligned}(W \cap ([0,1] \times \{b\}))^{-1}[\{y\}] &= W^{-1}[\{b\}] \text{ if } y = b,\\ &= \emptyset \text{ otherwise.}\end{aligned}$$

As b is arbitrary, every horizontal section of W is measurable. (β) If every horizontal section of W is measurable, let $F \subseteq [0,1]$ be any set of finite measure for ν; then F is finite, so

$$W \cap ([0,1] \times F) = \bigcup_{y\in F} W^{-1}[\{y\}] \times \{y\} \in \Sigma\widehat{\otimes}\mathrm{T} \subseteq \Lambda.$$

But it follows that W itself belongs to Λ, by 251H. **Q**

(ii) Now some of the same calculations show that for every $W \in \Lambda$,

$$\lambda W = \sum_{y\in[0,1]} \mu W^{-1}[\{y\}].$$

P For any finite $F \subseteq [0,1]$,

$$\begin{aligned}\lambda(W \cap ([0,1] \times F)) &= \int \mu(W \cap ([0,1] \times F))^{-1}[\{y\}]\nu(dy)\\ &= \int_F \mu W^{-1}[\{y\}]\nu(dy) = \sum_{y\in F} \mu W^{-1}[\{y\}].\end{aligned}$$

So

$$\lambda W = \sup_{F\subseteq[0,1] \text{ is finite}} \sum_{y\in F} \mu W^{-1}[\{y\}] = \sum_{y\in[0,1]} \mu W^{-1}[\{y\}]. \text{ **Q**}$$

252L Example For the second example, I turn to a problem that can arise if we neglect to check that a function is measurable as a function of two variables.

Let $(X,\Sigma,\mu) = (Y,\mathrm{T},\nu)$ be ω_1, the first uncountable ordinal (2A1Fc), with the countable-cocountable measure (211R). Set

$$W = \{(\xi,\eta) : \xi \le \eta < \omega_1\} \subseteq X \times Y.$$

Then all the horizontal sections $W^{-1}[\{\eta\}] = \{\xi : \xi \le \eta\}$ are countable, so

$$\int \mu W^{-1}[\{\eta\}]\nu(d\eta) = \int 0\,\nu(d\eta) = 0,$$

while all the vertical sections $W[\{\xi\}] = \{\eta : \xi \le \eta < \omega_1\}$ are cocountable, so

$$\int \nu W[\{\xi\}]\mu(d\xi) = \int 1\,\mu(d\xi) = 1.$$

Because the two repeated integrals are different, they cannot both be equal to the measure of W, and the sole resolution is to say that W is not measured by the product measure.

252M Remark A third kind of difficulty in the formula

$$\iint f(x,y)dxdy = \iint f(x,y)dydx$$

can arise even on probability spaces with $\Sigma\widehat{\otimes}\mathrm{T}$-measurable real-valued functions defined everywhere if we neglect to check that f is integrable with respect to the product measure. In 252H, we do need the hypothesis that one of

$$\int_{X\times Y}|f(x,y)|\lambda(d(x,y)),\quad \int_Y\int_X|f(x,y)|\mu(dx)\nu(dy),\quad \int_X\int_Y|f(x,y)|\nu(dy)\mu(dx)$$

is finite. For examples to show this, see 252Xf and 252Xg.

252N Integration through ordinate sets I: Proposition Let (X,Σ,μ) be a complete locally determined measure space, and λ the c.l.d. product measure on $X\times\mathbb{R}$, where $\mathbb{R}$ is given Lebesgue measure; write Λ for the domain of λ. For any $[0,\infty]$-valued function f defined on a conegligible subset of X, write Ω_f, Ω'_f for the **ordinate sets**

$$\Omega_f = \{(x,a) : x\in\operatorname{dom} f,\ 0\le a\le f(x)\}\subseteq X\times\mathbb{R},$$

$$\Omega'_f = \{(x,a) : x\in\operatorname{dom} f,\ 0\le a< f(x)\}\subseteq X\times\mathbb{R}.$$

Then

$$\lambda\Omega_f = \lambda\Omega'_f = \int f d\mu$$

in the sense that if one of these is defined in $[0,\infty]$, so are the other two, and they are equal.

proof (a) If $\Omega_f\in\Lambda$, then

$$\int f(x)\mu(dx) = \int \nu\{y : (x,y)\in\Omega_f\}\mu(dx) = \lambda\Omega_f$$

by 252D, writing μ for Lebesgue measure, because f is defined almost everywhere. Similarly, if $\Omega'_f\in\Lambda$,

$$\int f(x)\mu(dx) = \int \nu\{y : (x,y)\in\Omega'_f\}\mu(dx) = \lambda\Omega'_f.$$

(b) If $\int f d\mu$ is defined, then f is μ-virtually measurable, therefore measurable (because μ is complete); again because μ is complete, $\operatorname{dom} f\in\Sigma$. So

$$\Omega'_f = \bigcup_{q\in\mathbb{Q},q>0}\{x : x\in\operatorname{dom} f,\ f(x)>q\}\times[0,q],$$

$$\Omega_f = \bigcap_{n\ge1}\bigcup_{q\in\mathbb{Q},q>0}\{x : x\in\operatorname{dom} f,\ f(x)\ge q-\frac{1}{n}\}\times[0,q]$$

belong to Λ, so that $\lambda\Omega_f$ and $\lambda\Omega'_f$ are defined. Now both are equal to $\int f d\mu$, by (a).

252O Integration through ordinate sets II: Proposition Let (X,Σ,μ) be a measure space, and f a non-negative μ-virtually measurable function defined on a conegligible subset of X. Then

$$\int f d\mu = \int_0^\infty \mu^*\{x : x\in\operatorname{dom} f,\ f(x)\ge t\}dt = \int_0^\infty \mu^*\{x : x\in\operatorname{dom} f,\ f(x)>t\}dt$$

in $[0,\infty]$, where the integrals $\int\ldots dt$ are taken with respect to Lebesgue measure.

proof Completing μ does not change the integral of f or the outer measure μ^* (212Fb, 212Ea), so we may suppose that μ is complete, in which case $\operatorname{dom} f$ and f will be measurable. For n, $k\in\mathbb{N}$ set $E_{nk} = \{x : x\in\operatorname{dom} f,\ f(x)>2^{-n}k\}$, $g_n(x) = 2^{-n}\sum_{k=1}^{4^n}\chi E_{nk}$. Then $\langle g_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence of measurable functions converging to f at every point of $\operatorname{dom} f$, so $\int f d\mu = \lim_{n\to\infty}\int g_n d\mu$ and $\mu\{x : f(x)>t\} = \lim_{n\to\infty}\mu\{x : g_n(x)>t\}$ for every $t\ge0$; consequently

$$\int_0^\infty \mu\{x : f(x)>t\}dt = \lim_{n\to\infty}\int_0^\infty \mu\{x : g_n(x)>t\}dt.$$

On the other hand, $\mu\{x : g_n(x)>t\} = \mu E_{nk}$ if $1\le k\le 4^n$ and $2^{-n}(k-1)\le t<2^{-n}k$, 0 if $t\ge 2^n$, so that

$$\int_0^\infty \mu\{x : g_n(x)>t\}dt = \sum_{k=1}^{4^n}2^{-n}\mu E_{nk} = \int g_n\,d\mu,$$

for every $n\in\mathbb{N}$. So $\int_0^\infty \mu\{x : f(x)>t\}dt = \int f d\mu$.

Now $\mu\{x : f(x) \ge t\} = \mu\{x : f(x) > t\}$ for almost all t. **P** Set $C = \{t : \mu\{x : f(x) > t\} < \infty\}$, $h(t) = \mu\{x : f(x) > t\}$ for $t \in C$. If C is not empty, $h : C \to [0,\infty[$ is monotonic, so is continuous almost everywhere in C (222A). But at any point of $C \setminus \{\inf C\}$ at which h is continuous,

$$\mu\{x : f(x) \ge t\} = \lim_{s\uparrow t} \mu\{x : f(x) > s\} = \mu\{x : f(x) > t\}.$$

So we have the result, since $\mu\{x : f(x) \ge t\} = \mu\{x : f(x) > t\} = \infty$ for any $t \in [0,\infty[\setminus C$. **Q**

Accordingly $\int_0^\infty \mu\{x : f(x) \ge t\}dt$ is also equal to $\int f d\mu$.

***252P** If we work through the ideas of 252B for $\Sigma\widehat{\otimes}\mathrm{T}$-measurable functions, we get the following, which is sometimes useful.

Proposition Let (X,Σ,μ) be a measure space, and (Y,T,ν) a σ-finite measure space. Then for any $\Sigma\widehat{\otimes}\mathrm{T}$-measurable function $f : X \times Y \to [0,\infty]$, $x \mapsto \int f(x,y)\nu(dy) : X \to [0,\infty]$ is Σ-measurable. If μ is semi-finite, $\iint f(x,y)\nu(dy)\mu(dx) = \int f d\lambda$, where λ is the c.l.d. product measure on $X \times Y$.

proof (a) Let $\langle Y_n\rangle_{n\in\mathbb{N}}$ be a non-decreasing sequence of subsets of Y of finite measure with union Y. Set

$$\mathcal{A} = \{W : W \subseteq X \times Y,\ W[\{x\}] \in \mathrm{T} \text{ for every } x \in X,$$
$$x \mapsto \nu(Y_n \cap W[\{x\}]) \text{ is } \Sigma\text{-measurable for every } n \in \mathbb{N}\}.$$

Then $\mathcal{A}$ is a Dynkin class of subsets of $X \times Y$ including $\{E \times F : E \in \Sigma,\ F \in \mathrm{T}\}$, so includes $\Sigma\widehat{\otimes}\mathrm{T}$, by the Monotone Class Theorem again (136B).

This means that if $W \in \Sigma\widehat{\otimes}\mathrm{T}$, then

$$\mu W[\{x\}] = \sup_{n\in\mathbb{N}} \nu(Y_n \cap W[\{x\}])$$

is defined for every $x \in X$ and is a Σ-measurable function of x.

(b) Now, for n, $k \in \mathbb{N}$, set

$$W_{nk} = \{(x,y) : f(x,y) \ge 2^{-n}k\}, \quad g_n = \sum_{k=1}^{4^n} 2^{-n}\chi W_{nk}.$$

Then if we set

$$h_n(x) = \int g_n(x,y)\nu(dy) = \sum_{k=1}^{4^n} 2^{-n}\nu W_{nk}[\{x\}]$$

for $n \in \mathbb{N}$ and $x \in X$, $h_n : X \to [0,\infty]$ is Σ-measurable, and

$$\lim_{n\to\infty} h_n(x) = \int(\lim_{n\to\infty} g_n(x,y))\nu(dy) = \int f(x,y)\nu(dy)$$

for every x, because $\langle g_n(x,y)\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence with limit $f(x,y)$ for all $x \in X$, $y \in Y$. So $x \mapsto \int f(x,y)\nu(dy)$ is defined everywhere in X and is Σ-measurable.

(c) If $E \subseteq X$ is measurable and has finite measure, then $\int_E \int f(x,y)\nu(dy)\mu(dx) = \int_{E\times Y} f d\lambda$, applying 252B to the product of the subspace measure μ_E and ν (and using 251Q to check that the product of μ_E and ν is the subspace measure on $E \times Y$). Now if λW is defined and finite, there must be a non-decreasing sequence $\langle E_n\rangle_{n\in\mathbb{N}}$ of subsets of X of finite measure such that $\lambda W = \sup_{n\in\mathbb{N}} \lambda(W \cap (E_n \times Y))$, so that $W \setminus \bigcup_{n\in\mathbb{N}}(E_n \times Y)$ is negligible, and

$$\int_W f d\lambda = \lim_{n\to\infty} \int_{W\cap(E_n\times Y)} f d\lambda$$

(by B.Levi's theorem applied to $\langle f \times \chi(W \cap (E_n \times Y))\rangle_{n\in\mathbb{N}}$)

$$\le \lim_{n\to\infty} \int_{E_n\times Y} f d\lambda = \lim_{n\to\infty} \int_{E_n} \int f(x,y)\nu(dy)\mu(dx)$$
$$\le \iint f(x,y)\nu(dy)\mu(dx).$$

By 213B once more,

$$\int f d\lambda = \sup_{\lambda W<\infty} \int_W f d\lambda \le \iint f(x,y)\nu(dy)\mu(dx).$$

But also, if μ is semi-finite,

$$\iint f(x,y)\nu(dy)\mu(dx) = \sup_{\mu E<\infty} \int_E \int f(x,y)\nu(dy)\mu(dx) \le \int f d\lambda,$$

so $\int f d\lambda = \iint f(x,y)\nu(dy)\mu(dx)$, as claimed.

252Q The volume of a ball We now have all the essential machinery to perform a little calculation which is, I suppose, desirable simply as general knowledge: the volume of the unit ball $\{x : \|x\| \le 1\} = \{(\xi_1,\dots,\xi_r) : \sum_{i=1}^r \xi_i^2 \le 1\}$ in $\mathbb{R}^r$. In fact, from a theoretical point of view, I think we could very nearly just call it β_r and leave it at that; but since there is a general formula in terms of $\beta_2 = \pi$ and factorials, it seems shameful not to present it. The calculation has nothing to do with Lebesgue integration, and I could dismiss it as mere advanced calculus; but since only a minority of mathematicians are now taught calculus to this level with reasonable rigour before being introduced to the Lebesgue integral, I do not doubt that many readers, like myself, missed some of the subtleties involved. I therefore take the space to spell the details out in the style used elsewhere in this volume, recognising that the machinery employed is a good deal more elaborate than is really necessary for this result.

(a) The first basic fact we need is that, for any $n \ge 1$,

$$I_n = \int_{-\pi/2}^{\pi/2} \cos^n t\,dt = \frac{(2k)!}{(2^k k!)^2}\pi \text{ if } n = 2k \text{ is even,}$$

$$= 2\frac{(2^k k!)^2}{(2k+1)!} \text{ if } n = 2k+1 \text{ is odd.}$$

P For $n = 0$, of course,

$$I_0 = \int_{-\pi/2}^{\pi/2} 1\,dt = \pi = \frac{0!}{(2^0 0!)^2}\pi,$$

while for $n = 1$ we have

$$I_1 = \sin\tfrac{\pi}{2} - \sin(-\tfrac{\pi}{2}) = 2 = 2\frac{(2^0 0!)^2}{1!},$$

using the Fundamental Theorem of Calculus (225L) and the fact that $\sin' = \cos$ is bounded. For the inductive step to $n+1 \ge 2$, we can use integration by parts (225F):

$$I_{n+1} = \int_{-\pi/2}^{\pi/2} \cos t \cos^n t\,dt$$

$$= \sin\frac{\pi}{2}\cos^n\frac{\pi}{2} - \sin(-\frac{\pi}{2})\cos^n(-\frac{\pi}{2}) + \int_{-\pi/2}^{\pi/2} \sin t \cdot n\cos^{n-1} t \cdot \sin t\,dt$$

$$= n\int_{-\pi/2}^{\pi/2} (1-\cos^2 t)\cos^{n-1} t\,dt = n(I_{n-1} - I_{n+1}),$$

so that $I_{n+1} = \frac{n}{n+1} I_{n-1}$. Now the given formulae follow by an easy induction. **Q**

(b) The next result is that, for any $n \in \mathbb{N}$ and any $a \ge 0$,

$$\int_{-a}^{a} (a^2 - s^2)^{n/2} ds = I_{n+1} a^{n+1}.$$

P Of course this is an integration by substitution; but the singularity of the integrand at $s = \pm a$ complicates the issue slightly. I offer the following argument. If $a = 0$ the result is trivial; take $a > 0$. For $-a \le b \le a$, set $F(b) = \int_{-a}^{b} (a^2 - s^2)^{n/2} ds$. Because the integrand is continuous, $F'(b)$ exists and is equal to $(a^2 - b^2)^{n/2}$ for $-a < b < a$ (222H). Set $G(t) = F(a\sin t)$; then G is continuous and

$$G'(t) = aF'(a\sin t)\cos t = a^{n+1}\cos^{n+1} t$$

for $-\frac{\pi}{2} < t < \frac{\pi}{2}$. Consequently

$$\int_{-a}^{a}(a^2-s^2)^{n/2}ds = F(a)-F(-a) = G(\tfrac{\pi}{2})-G(-\tfrac{\pi}{2}) = \int_{-\pi/2}^{\pi/2}G'(t)dt$$

(by 225L, as before)

$$= a^{n+1}I_{n+1},$$

as required. **Q**

(c) Now at last we are ready for the balls $B_r = \{x : x \in \mathbb{R}^r, \|x\| \le 1\}$. Let μ_r be Lebesgue measure on $\mathbb{R}^r$, and set $\beta_r = I_1 I_2 \ldots I_r$ for $r \ge 1$. I claim that, writing

$$B_r(a) = \{x : x \in \mathbb{R}^r, \|x\| \le a\},$$

we have $\mu_r(B_r(a)) = \beta_r a^r$ for every $a \ge 0$. **P** Induce on r. For $r = 1$ we have $\beta_1 = 2$, $B_1(a) = [-a, a]$, so the result is trivial. For the inductive step to $r + 1$, we have

$$\mu_{r+1}B_{r+1}(a) = \int \mu_r\{x : (x, t) \in B_{r+1}(a)\}dt$$

(putting 251N and 252D together, and using the fact that $B_{r+1}(a)$ is closed, therefore measurable)

$$= \int_{-a}^{a}\mu_r B_r(\sqrt{a^2-t^2})dt$$

(because $(x, t) \in B_{r+1}(a)$ iff $|t| \le a$ and $\|x\| \le \sqrt{a^2-t^2}$)

$$= \int_{-a}^{a}\beta_r(a^2-t^2)^{r/2}dt$$

(by the inductive hypothesis)

$$= \beta_r a^{r+1}I_{r+1}$$

(by (b) above)

$$= \beta_{r+1}a^{r+1}$$

(by the definition of β_{r+1}). Thus the induction continues. **Q**

(d) In particular, the r-dimensional Lebesgue measure of the r-dimensional ball $B_r = B_r(1)$ is just $\beta_r = I_1 \ldots I_r$. Now an easy induction on k shows that

$$\beta_r = \frac{1}{k!}\pi^k \text{ if } r = 2k \text{ is even,}$$

$$= \frac{2^{2k+1}k!}{(2k+1)!}\pi^k \text{ if } r = 2k+1 \text{ is odd.}$$

(e) Note that in part (c) of the proof we saw that $\{x : x \in \mathbb{R}^r, \|x\| \le a\}$ has measure $\beta_r a^r$ for every $a \ge 0$.

The formulae here are consistent with the assignation $\beta_0 = 1$; which corresponds to saying that $\mathbb{R}^0 = \{\emptyset\}$, that $\mu_0\mathbb{R}^0 = 1$ and that $B_0 = \{\emptyset\}$. Taking $\mu_0\mathbb{R}^0$ to be 1 is itself consistent with the idea that, following 251N, the product measure $\mu_0 \times \mu_r$ ought to match μ_{0+r} on $\mathbb{R}^{0+r}$.

252R Complex-valued functions It is easy to apply the results of 252B-252I above to complex-valued functions, by considering their real and imaginary parts. Specifically:

(a) Let (X, Σ, μ) and (Y, T, ν) be measure spaces, and λ the c.l.d. product measure on $X \times Y$. Suppose that ν is σ-finite and that μ is either strictly localizable or complete and locally determined. Let f be a λ-integrable complex-valued function. Then $\iint f(x, y)\nu(dy)\mu(dx)$ is defined and equal to $\int f d\lambda$.

(b) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, λ the c.l.d. product measure on $X\times Y$, and Λ its domain. Let f be a Λ-measurable complex-valued function defined on a member of Λ, and suppose that either $\iint|f(x,y)|\mu(dx)\nu(dy)$ or $\iint|f(x,y)|\nu(dy)\mu(dx)$ is defined and finite. Then f is λ-integrable.

(c) Let (X,Σ,μ) and (Y,T,ν) be σ-finite measure spaces, λ the c.l.d. product measure on $X\times Y$, and Λ its domain. Let f be a Λ-measurable complex-valued function defined on a member of Λ. Then if one of

$$\int_{X\times Y}|f(x,y)|\lambda(d(x,y)),\quad \int_Y\int_X|f(x,y)|\mu(dx)\nu(dy),\quad \int_X\int_Y|f(x,y)|\nu(dy)\mu(dx)$$

exists in $\mathbb{R}$, so do the other two, and in this case

$$\int_{X\times Y}f(x,y)\lambda(d(x,y))=\int_Y\int_X f(x,y)\mu(dx)\nu(dy)=\int_X\int_Y f(x,y)\nu(dy)\mu(dx).$$

252X Basic exercises (a) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and λ the c.l.d. product measure on $X\times Y$. Let f be a λ-integrable real-valued function such that $\int_{E\times F}f=0$ whenever $E\in\Sigma$, $F\in\mathrm{T}$, $\mu E<\infty$ and $\nu F<\infty$. Show that $f=0$ λ-a.e. (*Hint*: use 251Ie to show that $\int_W f=0$ whenever $\lambda W<\infty$.)

(b) Let f, $g:\mathbb{R}\to\mathbb{R}$ be two non-decreasing functions, and μ_f, μ_g the associated Lebesgue-Stieltjes measures (see 114Xa). Set

$$f(x^+)=\lim_{t\downarrow x}f(t),\quad f(x^-)=\lim_{t\uparrow x}f(t)$$

for each $x\in\mathbb{R}$, and define $g(x^+)$, $g(x^-)$ similarly. Show that whenever $a\le b$ in $\mathbb{R}$,

$$\begin{aligned}\int_{[a,b]}f(x^-)\mu_g(dx)+\int_{[a,b]}g(x^+)\mu_f(dx)&=g(b^+)f(b^+)-g(a^-)f(a^-)\\&=\int_{[a,b]}\frac{1}{2}(f(x^-)+f(x^+))\mu_g(dx)+\int_{[a,b]}\frac{1}{2}((g(x^-)+g(x^+))\mu_f(dx).\end{aligned}$$

(*Hint*: find two expressions for $(\mu_f\times\mu_g)\{(x,y):a\le x<y\le b\}$.)

>(c) Let (X,Σ,μ) and (Y,T,ν) be complete locally determined measure spaces, λ the c.l.d. product measure on $X\times Y$, and Λ its domain. Suppose that $A\subseteq X$ and $B\subseteq Y$. Show that $A\times B\in\Lambda$ iff either $\mu A=0$ or $\nu B=0$ or $A\in\Sigma$ and $B\in\mathrm{T}$. (*Hint*: if B is not negligible and $A\times B\in\Lambda$, take H such that $\nu H<\infty$ and $B\cap H$ is not negligible. Then $W=A\times(B\cap H)$ is measured by $\mu\times\nu_H$, where ν_H is the subspace measure on H. Now apply 252D to μ, ν_H and W to see that $A\in\Sigma$.)

>(d) Let (X_1,Σ_1,μ_1), (X_2,Σ_2,μ_2), (X_2,Σ_3,μ_3) be three σ-finite measure spaces, and f a real-valued function defined almost everywhere on $X_1\times X_2\times X_3$ and Λ-measurable, where Λ is the domain of the product measure described in 251W or 251Xg. Show that if $\iiint|f(x_1,x_2,x_3)|dx_1dx_2dx_3$ is defined in $\mathbb{R}$, then $\iiint f(x_1,x_2,x_3)dx_2dx_3dx_1$ and $\iiint f(x_1,x_2,x_3)dx_3dx_1dx_2$ exist and are equal.

(e) Give an example of strictly localizable measure spaces (X,Σ,μ), (Y,T,ν) and a $W\in\Sigma\widehat{\otimes}\mathrm{T}$ such that $x\mapsto\nu W[\{x\}]$ is not Σ-measurable. (*Hint*: in 252Kb, try Y a proper subset of $[0,1]$.)

>(f) Set $f(x,y)=\sin(x-y)$ if $0\le y\le x\le y+2\pi$, 0 for other x, $y\in\mathbb{R}^2$. Show that $\iint f(x,y)dx\,dy=0$ and $\iint f(x,y)dy\,dx=2\pi$, taking all integrals with respect to Lebesgue measure.

>(g) Set $f(x,y)=\frac{x^2-y^2}{(x^2+y^2)^2}$ for x, $y\in\,]0,1]$. Show that $\int_0^1\int_0^1 f(x,y)dydx=\frac{\pi}{4}$, $\int_0^1\int_0^1 f(x,y)dxdy=-\frac{\pi}{4}$.

>(h) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and f a $\Sigma\widehat{\otimes}\mathrm{T}$-measurable function defined on a subset of $X\times Y$. Show that $y\mapsto f(x,y)$ is T-measurable for every $x\in X$.

(i) Let $r\ge1$ be an integer, and write β_r for the Lebesgue measure of the unit ball in $\mathbb{R}^r$. Set $g_r(t)=r\beta_rt^{r-1}$ for $t\ge0$, $\phi(x)=\|x\|$ for $x\in\mathbb{R}^r$. (i) Writing μ_r for Lebesgue measure on $\mathbb{R}^r$, show that $\mu_r\phi^{-1}[E]=\int_E r\beta_rt^{r-1}\mu_1(dt)$ for every Lebesgue measurable set $E\subseteq[0,\infty[$. (*Hint*: start with intervals E, noting from

115Xe that $\mu_r\{x : \|x\| \le a\} = \beta_r a^r$ for $a \ge 0$, and progress to open sets, negligible sets and general measurable sets.) (ii) Using 235R, show that

$$\int e^{-\|x\|^2/2}\mu_r(dx) = r\beta_r \int_0^\infty t^{r-1}e^{-t^2/2}\mu_1(dt) = 2^{(r-2)/2}r\beta_r\Gamma(\frac{r}{2})$$
$$= 2^{r/2}\beta_r\Gamma(1+\frac{r}{2}) = (\sqrt{2}\Gamma(\frac{1}{2}))^r$$

where Γ is the Γ-function (225Xh). (iii) Show that

$$2\Gamma(\tfrac{1}{2})^2 = 2\beta_2\Gamma(2) = 2\beta_2\int_0^\infty te^{-t^2/2}dt = 2\pi,$$

and hence that $\beta_r = \frac{\pi^{r/2}}{\Gamma(1+\frac{r}{2})}$ and $\int_{-\infty}^\infty e^{-t^2/2}dt = \sqrt{2\pi}$.

>(j) Let (X,Σ,μ) be a measure space, and $f : X \to [0,\infty[$ a function. Write $\mathcal{B}$ for the Borel σ-algebra of $\mathbb{R}$. Show that the following are equiveridical: (α) f is Σ-measurable; (β) $\{(x,a) : x \in X,\ 0 \le a \le f(x)\} \in \Sigma\widehat{\otimes}\mathcal{B}$; ($\gamma$) $\{(x,a) : x \in X,\ 0 \le a < f(x)\} \in \Sigma\widehat{\otimes}\mathcal{B}$.

252Y Further exercises (a) Let (X,Σ,μ) be a measure space. Show that the following are equiveridical: (i) the completion of μ is locally determined; (ii) the completion of μ coincides with the c.l.d. version of μ; (iii) whenever (Y,T,ν) is a σ-finite measure space and λ the c.l.d. product measure on $X \times Y$ and f is a function such that $\int f d\lambda$ is defined in $[-\infty,\infty]$, then $\iint f(x,y)\nu(dy)\mu(dx)$ is defined and equal to $\int f d\lambda$.

(b) Let (X,Σ,μ) be a measure space. Show that the following are equiveridical: (i) μ has locally determined negligible sets; (ii) whenever (Y,T,ν) is a σ-finite measure space and λ the c.l.d. product measure on $X \times Y$, then $\iint f(x,y)\nu(dy)\mu(dx)$ is defined and equal to $\int f d\lambda$ for any λ-integrable function f.

(c) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and λ_0 the primitive product measure on $X\times Y$ (251C). Let f be any λ_0-integrable real-valued function. Show that $\iint f(x,y)\nu(dy)\mu(dx) = \int f d\lambda_0$. (*Hint*: show that there are sequences $\langle G_n\rangle_{n\in\mathbb{N}}$, $\langle H_n\rangle_{n\in\mathbb{N}}$ of sets of finite measure such that $f(x,y)$ is defined and equal to 0 for every $(x,y) \in (X \times Y) \setminus \bigcup_{n\in\mathbb{N}} G_n \times H_n$.)

(d) Let (X,Σ,μ) and (Y,T,ν) be measure spaces; let λ_0 be the primitive product measure on $X \times Y$, and λ the c.l.d. product measure. Show that if f is a λ_0-integrable real-valued function, it is λ-integrable, and $\int f d\lambda = \int f d\lambda_0$.

(e) Let (X,Σ,μ) be a complete locally determined measure space and $a < b$ in $\mathbb{R}$, endowed with Lebesgue measure; let Λ be the domain of the c.l.d. product measure λ on $X \times [a,b]$. Let $f : X \times]a,b[\to \mathbb{R}$ be a Λ-measurable function such that $t \mapsto f(x,t) : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on $]a,b[$ for every $x \in X$. (i) Show that the partial derivative $\frac{\partial f}{\partial t}$ with respect to the second variable is Λ-measurable. (ii) Now suppose that $\frac{\partial f}{\partial t}$ is λ-integrable and that $\int f(x,t_0)\mu(dx)$ is defined and finite for some $t_0 \in]a,b[$. Show that $F(t) = \int f(x,t)\mu(dx)$ is defined in $\mathbb{R}$ for every $t \in [a,b]$, that F is absolutely continuous, and that $F'(t) = \int \frac{\partial f}{\partial t}(x,t)\mu(dx)$ for almost every $t \in]a,b[$. (*Hint*: $F(c) = F(a) + \int_{X\times[a,c]} \frac{\partial f}{\partial t}d\lambda$ for every $c \in [a,b]$.)

(f) Show that $\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \int_0^1 t^{a-1}(1-t)^{b-1}dt$ for all a, $b > 0$. (*Hint*: show that

$$\int_0^\infty t^{a-1}\int_t^\infty e^{-x}(x-t)^{b-1}dxdt = \int_0^\infty e^{-x}\int_0^x t^{a-1}(x-t)^{b-1}dtdx.)$$

(g) Let (X,Σ,μ) and (Y,T,ν) be σ-finite measure spaces and λ the c.l.d. product measure on $X \times Y$. Suppose that $f \in \mathcal{L}^0(\lambda)$ and that $1 < p < \infty$. Show that $(\int|\int f(x,y)dx|^p dy)^{1/p} \le \int(\int|f(x,y)|^p dy)^{1/p}dx$. (*Hint*: set $q = \frac{p}{p-1}$ and consider the integral $\int|f(x,y)g(y)|\lambda(d(x,y))$ for $g \in \mathcal{L}^q(\nu)$, using 244K.)

(h) Let ν be Lebesgue measure on $[0,\infty[$; suppose that $f \in \mathcal{L}^p(\nu)$ where $1 < p < \infty$. Set $F(y) = \frac{1}{y}\int_0^y f$ for $y > 0$. Show that $\|F\|_p \le \frac{p}{p-1}\|f\|_p$. (*Hint*: $F(y) = \int_0^1 f(xy)dx$; use 252Yg with $X = [0,1]$, $Y = [0,\infty[$.)

(i) Show that if p is any non-zero (real) polynomial in r variables, then $\{x : x \in \mathbb{R}^r,\ p(x) = 0\}$ is Lebesgue negligible.

(j) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, with c.l.d. product $(X\times Y,\Lambda,\lambda)$. Let f be a non-negative Λ-measurable real-valued function defined on a λ-conegligible set, and suppose that

$$\overline{\int}\bigl(\overline{\int} f(x,y)\mu(dx)\bigr)\nu(dy)$$

is finite. Show that f is λ-integrable.

(k) Let (X,Σ,μ) be the unit interval $[0,1]$ with Lebesgue measure, and (Y,T,ν) the interval with counting measure, as in 252K; let λ_0 be the primitive product measure on $[0,1]^2$. (i) Setting $\Delta = \{(t,t) : t \in [0,1]\}$, show that $\lambda_0\Delta = \infty$. (ii) Show that λ_0 is not semi-finite. (iii) Show that if $W \in \operatorname{dom}\lambda_0$, then $\lambda_0 W = \sum_{y\in[0,1]}\mu W^{-1}[\{y\}]$ if there are a countable set $A \subseteq [0,1]$ and a Lebesgue negligible set $E \subseteq [0,1]$ such that $W \subseteq ([0,1]\times A)\cup(E\times[0,1])$, ∞ otherwise.

(l) Let (X,Σ,μ) be a measure space, and λ_0 the primitive product measure on $X\times\mathbb{R}$, where $\mathbb{R}$ is given Lebesgue measure; write Λ for its domain. For any $[0,\infty]$-valued function f defined on a conegligible subset of X, write Ω_f, Ω'_f for the corresponding ordinate sets, as in 252N. Show that if any of $\lambda_0\Omega_f$, $\lambda_0\Omega'_f$, $\int f d\mu$ is defined and finite, so are the others, and all three are equal.

(m) Let (X,Σ,μ) be a complete locally determined measure space, and f a non-negative function defined on a conegligible subset of X. Write Ω_f, Ω'_f for the corresponding ordinate sets, as in 252N. Let λ be the c.l.d. product measure on $X\times\mathbb{R}$, where $\mathbb{R}$ is given Lebesgue measure. Show that $\overline{\int} f\,d\mu = \lambda^*\Omega_f = \lambda^*\Omega'_f$.

(n) Let (X,Σ,μ) be a measure space and $f : X \to [0,\infty[$ a function. Show that $\overline{\int} f d\mu = \int_0^\infty \mu^*\{x : f(x) \ge t\}dt$.

(o) Let (X,Σ,μ) be a complete measure space and write $\mathcal{M}^{0,\infty}$ for the set $\{f : f \in \mathcal{L}^0(\mu),\ \mu\{x : |f(x)| \ge a\}$ is finite for some $a \in [0,\infty[\}$. (i) Show that for each $f \in \mathcal{M}^{0,\infty}$ there is a non-increasing $f^* :]0,\infty[\to \mathbb{R}$ such that $\mu_L\{t : f^*(t) \ge \alpha\} = \mu\{x : |f(x)| \ge \alpha\}$ for every $\alpha > 0$, writing μ_L for Lebesgue measure. (ii) Show that $\int_E |f| d\mu \le \int_0^{\mu E} f^* d\mu_L$ for every $E \in \Sigma$ (allowing ∞). (*Hint*: $(f\times\chi E)^* \le f^*$.) (iii) Show that $\|f^*\|_p = \|f\|_p$ for every $p \in [1,\infty]$, $f \in \mathcal{M}^{0,\infty}$. (*Hint*: $(|f|^p)^* = (f^*)^p$.) (iv) Show that if f, $g \in \mathcal{M}^{0,\infty}$ then $\int |f\times g| d\mu \le \int f^*\times g^* d\mu_L$. (*Hint*: look at simple functions first.) (v) Show that if μ is atomless then $\int_0^a f^* d\mu_L = \sup_{E\in\Sigma,\mu E\le a}\int_E |f|$ for every $a \ge 0$. (*Hint*: 215D.) (vi) Show that $A \subseteq \mathcal{L}^1(\mu)$ is uniformly integrable iff $\{f^* : f \in A\}$ is uniformly integrable in $\mathcal{L}^1(\mu_L)$. (f^* is called the **decreasing rearrangement** of f.)

(p) Let (X,Σ,μ) be a complete locally determined measure space, and write ν for Lebesgue measure on $[0,1]$. Show that the c.l.d. product measure λ on $X\times[0,1]$ is localizable iff μ is localizable. (*Hints*: (i) if $\mathcal{E} \subseteq \Sigma$, show that $F \in \Sigma$ is an essential supremum for $\mathcal{E}$ in Σ iff $F\times[0,1]$ is an essential supremum for $\{E\times[0,1] : E \in \mathcal{E}\}$ in $\Lambda = \operatorname{dom}\lambda$. (ii) For $W \in \Lambda$, $n \in \mathbb{N}$, $k < 2^n$ set

$$W_{nk} = \{x : x \in X,\ \nu^*\{t : (x,t) \in W,\ 2^{-n}k \le t \le 2^{-n}(k+1)\} \ge 2^{-n-1}\}.$$

Show that if $\mathcal{W} \subseteq \Lambda$ and F_{nk} is an essential supremum for $\{W_{nk} : W \in \mathcal{W}\}$ in Σ for all n, k, then

$$\bigcup_{n\in\mathbb{N}}\bigcap_{m\ge n}\bigcup_{k<2^m} F_{mk}\times[2^{-m}k, 2^{-m}(k+1)]$$

is an essential supremum for $\mathcal{W}$ in Λ.)

(q) Let (X,Σ,μ) be the space of Example 216D, and give Lebesgue measure to $[0,1]$. Show that the c.l.d. product measure on $X\times[0,1]$ is complete, locally determined, atomless and not localizable.

(r) Let (X,Σ,μ) be a complete locally determined measure space and (Y,T,ν) a semi-finite measure space with $\nu Y>0$. Show that if the c.l.d. product measure on $X\times Y$ is strictly localizable, then μ is strictly localizable. (*Hint*: take $F\in\mathrm{T}$, $0<\nu F<\infty$. Let $\langle W_i\rangle_{i\in I}$ be a decomposition of $X\times Y$. For $i\in I$, $n\in\mathbb{N}$ set $E_{in}=\{x:\nu^*\{y:y\in F,\,(x,y)\in W_i\}\ge 2^{-n}\}$. Apply 213Yf to $\{E_{in}:i\in I,\,n\in\mathbb{N}\}$.)

(s) Let (X,Σ,μ) be the space of Example 216E, and give Lebesgue measure to $[0,1]$. Show that the c.l.d. product measure on $X\times[0,1]$ is complete, locally determined, atomless and localizable, but not strictly localizable.

(t) Let (X,Σ,μ) be a measure space and f a μ-integrable complex-valued function. For $\alpha\in\,]-\pi,\pi]$ set $H_\alpha=\{x:x\in\operatorname{dom}f,\,\mathcal{R}\mathrm{e}(e^{-i\alpha}f(x))>0\}$. Show that $\int_{-\pi}^{\pi}\mathcal{R}\mathrm{e}(e^{-i\alpha}\int_{H_\alpha}f)d\alpha=2\int|f|$, and hence that there is some α such that $|\int_{H_\alpha}f|\ge\frac{1}{\pi}\int|f|$. (Compare 246F.)

(u) Set $f(t)=t-\ln(t+1)$ for $t>-1$. (i) Show that $\Gamma(a+1)=a^{a+1}e^{-a}\int_{-1}^{\infty}e^{-af(u)}du$ for every $a>0$. (*Hint*: substitute $u=\frac{t}{a}-1$ in 225Xh(iii).) (ii) Show that there is a $\delta>0$ such that $f(t)\ge\frac13t^2$ for $-1\le t\le\delta$. (iii) Setting $\alpha=\frac12f(\delta)$, show that (for $a\ge1$)

$$\sqrt{a}\int_\delta^\infty e^{-af(t)}dt\le\sqrt{a}e^{-a\alpha}\int_0^\infty e^{-f(t)/2}dt\to0$$

as $a\to\infty$. (iv) Set $g_a(t)=e^{-af(t/\sqrt{a})}$ if $-\sqrt{a}<t\le\delta\sqrt{a}$, 0 otherwise. Show that $g_a(t)\le e^{-t^2/3}$ for all a, t and that $\lim_{a\to\infty}g_a(t)=e^{-t^2/2}$ for all t, so that

$$\begin{aligned}\lim_{a\to\infty}\frac{e^a\Gamma(a+1)}{a^{a+\frac12}}&=\lim_{a\to\infty}\sqrt{a}\int_{-1}^{\infty}e^{-af(t)}dt=\lim_{a\to\infty}\sqrt{a}\int_{-1}^{\delta}e^{-af(t)}dt\\&=\lim_{a\to\infty}\int_{-\infty}^{\infty}g_a(t)dt=\int_{-\infty}^{\infty}e^{-t^2/2}dt=\sqrt{2\pi}.\end{aligned}$$

(v) Show that $\lim_{n\to\infty}\dfrac{n!}{e^{-n}n^n\sqrt{n}}=\sqrt{2\pi}$. (This is **Stirling's formula**.)

(v) Let (X,Σ,μ) be a complete locally determined measure space and f, g two real-valued, μ-virtually measurable functions defined almost everywhere in X. (i) Let λ be the c.l.d. product of μ and Lebesgue measure on $\mathbb{R}$. Setting $\Omega_f^*=\{(x,a):x\in\operatorname{dom}f,\,a\in\mathbb{R},\,a\le f(x)\}$ and $\Omega_g^*=\{(x,a):x\in\operatorname{dom}g,\,a\in\mathbb{R},\,a\le g(x)\}$, show that $\lambda(\Omega_f^*\setminus\Omega_g^*)=\int(f-g)^+d\mu$ and $\lambda(\Omega^*\triangle\Omega_g^*)=\int|f-g|d\mu$. (ii) Suppose that μ is σ-finite. Show that

$$\int|f-g|d\mu=\int_{-\infty}^{\infty}\mu(\{x:x\in\operatorname{dom}f\cap\operatorname{dom}g,\,(f(x)-a)(g(x)-a)<0\}da.$$

(iii) Suppose that μ is σ-finite, that T is a σ-subalgebra of Σ, that $E\in\Sigma$ and that $g:X\to[0,1]$ is T-measurable. Show that there is an $F\in\mathrm{T}$ such that $\mu(E\triangle F)\le\int|\chi E-g|d\mu$.

252 Notes and comments For a volume and a half now I have asked you to accept the idea of integrating partially-defined functions, insisting that sooner or later they would appear at the core of the subject. The moment has now come. If we wish to apply Fubini's and Tonelli's theorems in the most fundamental of all cases, with both factors equal to Lebesgue measure on the unit interval, it is surely natural to look at all functions which are integrable on the square for two-dimensional Lebesgue measure. Now two-dimensional Lebesgue measure is a complete measure, so, in particular, assigns zero measure to any set of the form $\{(x,b):x\in A\}$ or $\{(a,y):y\in A\}$, whether or not the set A is measured by one-dimensional measure. Accordingly, if f is a function of two variables which is integrable for two-dimensional Lebesgue measure, there is no reason why any particular section $x\mapsto f(x,b)$ or $y\mapsto f(a,y)$ should be measurable, let alone integrable. Consequently, even if f itself is defined everywhere, the outer integral of $\iint f(x,y)dxdy$ is likely to be applied to a function which is not defined for every y. Let me remark that the problem does not concern '∞'; the awkward functions are those with sections so irregular that they cannot be assigned an integral at all.

I have seen many approaches to this particular nettle, generally less whole-hearted than the one I have determined on for this treatise. Part of the difficulty is that Fubini's theorem really is at the centre of measure theory. Over large parts of the subject, it is possible to assert that a result is non-trivial if and only if it depends on Fubini's theorem. I am therefore unwilling to insert any local fix, saying that 'in this chapter, we shall integrate functions which are not defined everywhere'; before long, such a provision would have to be interpolated into the preambles to half the best theorems, or an explanation offered of why it wasn't necessary in their particular contexts. I suppose that one of the commonest responses is (like Halmos 50) to restrict attention to $\Sigma\widehat{\otimes}\mathrm{T}$-measurable functions, which eliminates measurability problems for the moment (252Xh, 252P); but unhappily (or rather, to my mind, happily) there are crucial applications in which the functions are not actually $\Sigma\widehat{\otimes}\mathrm{T}$-measurable, but belong to some wider class, and this restriction sooner or later leads to undignified contortions as we are forced to adapt limited results to unforeseen contexts. Besides, it leaves unsaid the really rather important information that if f is a measurable function of two variables then (under appropriate conditions) almost all its sections are measurable (252E).

In 252B and its corollaries there is a clumsy restriction: we assume that one of the measures is σ-finite and the other is either strictly localizable or complete and locally determined. The obvious question is, whether we need these hypotheses. From 252K we see that the hypothesis 'σ-finite' on the second factor can certainly not be abandoned, even when the first factor is a complete probability measure. The requirement 'μ is either strictly localizable or complete and locally determined' is in fact fractionally stronger than what is needed, as well as disagreeably elaborate. The 'right' hypothesis is that the completion of μ should be locally determined (see 252Ya). The point is that because the product of two measures is the same as the product of their c.l.d. versions (251T), no theorem which leads from the product measure to the factor measures can distinguish between a measure and its c.l.d. version; so that, in 252B, we must expect to need μ and its c.l.d. version to give rise to the same integrals. The proof of 252B would be better focused if the hypothesis was simplified to 'ν is σ-finite and μ is complete and locally determined'. But this would just transfer part of the argument into the proof of 252C.

We also have to work a little harder in 252B in order to cover functions and integrals taking the values $\pm\infty$. Fubini's theorem is so central to measure theory that I believe it is worth taking a bit of extra trouble to state the results in maximal generality. This is especially important because we frequently apply it in multiply repeated integrals, as in 252Xd, in which we have even less control than usual over the intermediate functions to be integrated.

I have expressed all the main results of this section in terms of the 'c.l.d.' product measure. In the case of σ-finite spaces, of course, which is where the theory works best, we could just as well use the 'primitive' product measure. Indeed, Fubini's theorem itself has a version in terms of the primitive product measure which is rather more elegant than 252B as stated (252Yc), and covers the great majority of applications. (Integrals with respect to the primitive and c.l.d. product measures are of course very closely related; see 252Yd.) But we do sometimes need to look at non-σ-finite spaces, and in these cases the asymmetric form in 252B is close to the best we can do. Using the primitive product measure does not help at all with the most substantial obstacle, the phenomenon in 252K (see 252Yk).

The pre-calculus concept of an integral as 'the area under a curve' is given expression in 252N: the integral of a non-negative function is the measure of its ordinate set. This is unsatisfactory as a definition of the integral, not just because of the requirement that the base space should be complete and locally determined (which can be dealt with by using the primitive product measure, as in 252Yl), but because the construction of the product measure involves integration (part (c) of the proof of 251E). The idea of 252N is to relate the measure of an ordinate set to the integral of the measures of its vertical sections. Curiously, if instead we integrate the measures of its *horizontal* sections, as in 252O, we get a more versatile result. (Indeed this one does not involve the concept of 'product measure', and could have appeared at any point after §123.) Note that the integral $\int_0^\infty \ldots dt$ here is applied to a monotonic function, so may be interpreted as an improper Riemann integral. If you think you know enough about the Riemann integral to make this a tempting alternative to the construction in §122, the tricky bit now becomes the proof that the integral is additive.

A different line of argument is to use integration over sections to define a product measure. The difficulty with this approach is that unless we take great care we may find ourselves with an asymmetric construction. My own view is that such an asymmetry is acceptable only when there is no alternative. But in Chapter 43 of Volume 4 I will describe a couple of examples.

Of the two examples I give here, 252K is supposed to show that when I call for σ-finite spaces they are really necessary, while 252L is supposed to show that joint measurability is essential in Tonelli's theorem and its corollaries. The factor spaces in 252K, Lebesgue measure and counting measure, are chosen to show that it is only the lack of σ-finiteness that can be the problem; they are otherwise as regular as one can reasonably ask. In 252L I have used the countable-cocountable measure on ω_1, which you may feel is fit only for counter-examples; and the question does arise, whether the same phenomenon occurs with Lebesgue measure. This leads into deep water, and I will return to it in Chapter 53 of Volume 5.

I ought perhaps to note explicitly that in Fubini's theorem, we really do need to have a function which is integrable for the product measure. I include 252Xf and 252Xg to remind you that even in the best-regulated circumstances, the repeated integrals $\iint f\,dxdy$, $\iint f\,dydx$ may fail to be equal if f is not integrable as a function of two variables.

There are many ways to calculate the volume β_r of an r-dimensional ball; the one I have used in 252Q follows a line that would have been natural to me before I ever heard of measure theory. In 252Xi I suggest another method. The idea of integration-by-substitution, used in part (b) of the argument for 252Q, is there supported by an ad hoc argument; I will present a different, more generally applicable, approach in Chapter 26. Elsewhere (252Xi, 252Yf, 252Yh, 252Yu) I find myself taking for granted substitutions of the form $t\mapsto at$, $t\mapsto a+t$, $t\mapsto t^2$; for a systematic justification, see §263. Of course an enormous number of other formulae of advanced calculus are also based on repeated integration of one kind or another, and I give a sample handful of such results (252Xb, 252Ye-252Yh, 252Yu).

Version of 18.4.08

253 Tensor products

The theorems of the last section show that the integrable functions on a product of two measure spaces can be effectively studied in terms of integration on each factor space separately. In this section I present a very striking relationship between the L^1 space of a product measure and the L^1 spaces of its factors, which actually determines the product L^1 up to isomorphism as Banach lattice. I start with a brief note on bilinear operators (253A) and a description of the canonical bilinear operator from $L^1(\mu)\times L^1(\nu)$ to $L^1(\mu\times\nu)$ (253B-253E). The main theorem of the section is 253F, showing that this canonical map is universal for continuous bilinear operators from $L^1(\mu)\times L^1(\nu)$ to Banach spaces; it also determines the ordering of $L^1(\mu\times\nu)$ (253G). I end with a description of a fundamental type of conditional expectation operator (253H) and notes on products of indefinite-integral measures (253I) and upper integrals of special kinds of function (253J, 253K).

253A Bilinear operators Before looking at any of the measure theory in this section, I introduce a concept from the theory of linear spaces.

(a) Let U, V and W be linear spaces over $\mathbb{R}$ (or, indeed, any field). A map $\phi:U\times V\to W$ is **bilinear** if it is linear in each variable separately, that is,

$$\phi(u_1+u_2,v)=\phi(u_1,v)+\phi(u_2,v),$$

$$\phi(u,v_1+v_2)=\phi(u,v_1)+\phi(u,v_2),$$

$$\phi(\alpha u,v)=\alpha\phi(u,v)=\phi(u,\alpha v)$$

for all u, u_1, $u_2\in U$, v, v_1, $v_2\in V$ and scalars α. Observe that such a ϕ gives rise to, and in turn can be defined by, a linear operator $T:U\to\mathsf{L}(V;W)$, writing $\mathsf{L}(V;W)$ for the space of linear operators from V to W, where

$$(Tu)(v)=\phi(u,v)$$

for all $u\in U$, $v\in V$. Hence, or otherwise, we can see, for instance, that $\phi(0,v)=\phi(u,0)=0$ whenever $u\in U$ and $v\in V$.

If W' is another linear space over the same field, and $S:W\to W'$ is a linear operator, then $S\phi:U\times V\to W'$ is bilinear.

(b) Now suppose that U, V and W are normed spaces, and $\phi : U \times V \to W$ a bilinear operator. Then we say that ϕ is **bounded** if $\sup\{\|\phi(u,v)\| : \|u\| \le 1,\ \|v\| \le 1\}$ is finite, and in this case we call this supremum the norm $\|\phi\|$ of ϕ. Note that $\|\phi(u,v)\| \le \|\phi\|\|u\|\|v\|$ for all $u \in U$, $v \in V$ (because

$$\|\phi(u,v)\| = \alpha\beta\|\phi(\alpha^{-1}u, \beta^{-1}v)\| \le \alpha\beta\|\phi\|$$

whenever $\alpha > \|u\|$, $\beta > \|v\|$).

If W' is another normed space and $S : W \to W'$ is a bounded linear operator, then $S\phi : U \times V \to W'$ is a bounded bilinear operator, and $\|S\phi\| \le \|S\|\|\phi\|$.

253B Definition The most important bilinear operators of this section are based on the following idea. Let f and g be real-valued functions. I will write $f \otimes g$ for the function $(x,y) \mapsto f(x)g(y) : \operatorname{dom} f \times \operatorname{dom} g \to \mathbb{R}$.

253C Proposition (a) Let X and Y be sets, and Σ, T σ-algebras of subsets of X, Y respectively. If f is a Σ-measurable real-valued function defined on a subset of X, and g is a T-measurable real-valued function defined on a subset of Y, then $f \otimes g$, as defined in 253B, is $\Sigma\widehat{\otimes}\mathrm{T}$-measurable.

(b) Let (X, Σ, μ) and (Y, T, ν) be measure spaces, and λ the c.l.d. product measure on $X \times Y$. If $f \in \mathcal{L}^0(\mu)$ and $g \in \mathcal{L}^0(\nu)$, then $f \otimes g \in \mathcal{L}^0(\lambda)$.

Remark Recall from 241A that $\mathcal{L}^0(\mu)$ is the space of μ-virtually measurable real-valued functions defined on μ-conegligible subsets of X.

proof (a) The point is that $f \otimes \chi Y$ is $\Sigma\widehat{\otimes}\mathrm{T}$-measurable, because for any $\alpha \in \mathbb{R}$ there is an $E \in \Sigma$ such that

$$\{x : f(x) \ge \alpha\} = E \cap \operatorname{dom} f,$$

so that

$$\{(x,y) : (f \otimes \chi Y)(x,y) \ge \alpha\} = (E \cap \operatorname{dom} f) \times Y = (E \times Y) \cap \operatorname{dom}(f \otimes \chi Y),$$

and of course $E \times Y \in \Sigma\widehat{\otimes}\mathrm{T}$. Similarly, $\chi X \otimes g$ is $\Sigma\widehat{\otimes}\mathrm{T}$-measurable and $f \otimes g = (f \otimes \chi Y) \times (\chi X \otimes g)$ is $\Sigma\widehat{\otimes}\mathrm{T}$-measurable.

(b) Let $E \in \Sigma$, $F \in \mathrm{T}$ be conegligible subsets of X, Y respectively such that $E \subseteq \operatorname{dom} f$, $F \subseteq \operatorname{dom} g$, $f{\restriction}E$ is Σ-measurable and $g{\restriction}F$ is T-measurable. Write Λ for the domain of λ. Then $\Sigma\widehat{\otimes}\mathrm{T} \subseteq \Lambda$ (251Ia). Also $E \times F$ is λ-conegligible, because

$$\begin{aligned}\lambda((X \times Y) \setminus (E \times F)) &\le \lambda((X \setminus E) \times Y) + \lambda(X \times (Y \setminus F)) \\ &= \mu(X \setminus E) \cdot \nu Y + \mu X \cdot \nu(Y \setminus F) = 0\end{aligned}$$

(also from 251Ia). So $\operatorname{dom}(f \otimes g) \supseteq E \times F$ is conegligible. Also, by (a), $(f \otimes g){\restriction}(E \times F) = (f{\restriction}E) \otimes (g{\restriction}F)$ is $\Sigma\widehat{\otimes}\mathrm{T}$-measurable, therefore Λ-measurable, and $f \otimes g$ is virtually measurable. Thus $f \otimes g \in \mathcal{L}^0(\lambda)$, as claimed.

253D Now we can apply the ideas of 253B-253C to integrable functions.

Proposition Let (X, Σ, μ) and (Y, T, ν) be measure spaces, and write λ for the c.l.d. product measure on $X \times Y$. If $f \in \mathcal{L}^1(\mu)$ and $g \in \mathcal{L}^1(\nu)$, then $f \otimes g \in \mathcal{L}^1(\lambda)$ and $\int f \otimes g \, d\lambda = \int f \, d\mu \int g \, d\nu$.

Remark I follow §242 in writing $\mathcal{L}^1(\mu)$ for the space of μ-integrable real-valued functions.

proof (a) Consider first the case $f = \chi E$, $g = \chi F$ where $E \in \Sigma$, $F \in \mathrm{T}$ have finite measure; then $f \otimes g = \chi(E \times F)$ is λ-integrable with integral

$$\lambda(E \times F) = \mu E \cdot \nu F = \int f \, d\mu \cdot \int g \, d\nu,$$

by 251Ia.

(b) It follows at once that $f \otimes g$ is λ-simple, with $\int f \otimes g \, d\lambda = \int f \, d\mu \int g \, d\nu$, whenever f is a μ-simple function and g is a ν-simple function.

(c) If f and g are non-negative integrable functions, there are non-decreasing sequences $\langle f_n \rangle_{n \in \mathbb{N}}$, $\langle g_n \rangle_{n \in \mathbb{N}}$ of non-negative simple functions converging almost everywhere to f, g respectively; now note that if $E \subseteq X$,

$F\subseteq Y$ are conegligible, $E\times F$ is conegligible in $X\times Y$, as remarked in the proof of 253C, so the non-decreasing sequence $\langle f_n\times g_n\rangle_{n\in\mathbb{N}}$ of λ-simple functions converges almost everywhere to $f\otimes g$, and

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

by B.Levi's theorem.

(d) Finally, for general f and g, we can express them as the differences f^+-f^-, g^+-g^- of non-negative integrable functions, and see that

$$\int f\otimes g\,d\lambda=\int f^+\otimes g^+-f^+\otimes g^--f^-\otimes g^++f^-\otimes g^-d\lambda=\int f\,d\mu\int g\,d\nu.$$

253E The canonical map $L^1\times L^1\to L^1$ I continue the argument from 253D. Because $E\times F$ is conegligible in $X\times Y$ whenever E and F are conegligible subsets of X and Y, $f_1\otimes g_1=f\otimes g$ λ-a.e. whenever $f=f_1$ μ-a.e. and $g=g_1$ ν-a.e. We may therefore define $u\otimes v\in L^1(\lambda)$, for $u\in L^1(\mu)$ and $v\in L^1(\nu)$, by saying that $u\otimes v=(f\otimes g)^\bullet$ whenever $u=f^\bullet$ and $v=g^\bullet$.

Now if f, f_1, $f_2\in\mathcal{L}^1(\mu)$, g, g_1, $g_2\in\mathcal{L}^1(\nu)$ and $a\in\mathbb{R}$,

$$(f_1+f_2)\otimes g=(f_1\otimes g)+(f_2\otimes g),$$

$$f\otimes(g_1+g_2)=(f\otimes g_1)+(f\otimes g_2),$$

$$(af)\otimes g=a(f\otimes g)=f\otimes(ag).$$

It follows at once that the map $(u,v)\mapsto u\otimes v$ is bilinear.

Moreover, if $f\in\mathcal{L}^1(\mu)$ and $g\in\mathcal{L}^1(\nu)$, $|f|\otimes|g|=|f\otimes g|$, so $\int|f\otimes g|d\lambda=\int|f|d\mu\int|g|d\nu$. Accordingly

$$\|u\otimes v\|_1=\|u\|_1\|v\|_1$$

for all $u\in L^1(\mu)$, $v\in L^1(\nu)$. In particular, the bilinear operator $\otimes$ is bounded, with norm 1 (except in the trivial case in which one of $L^1(\mu)$, $L^1(\nu)$ is 0-dimensional).

253F We are now ready for the main theorem of this section.

Theorem Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and let λ be the c.l.d. product measure on $X\times Y$. Let W be any Banach space and $\phi:L^1(\mu)\times L^1(\nu)\to W$ a bounded bilinear operator. Then there is a unique bounded linear operator $T:L^1(\lambda)\to W$ such that $T(u\otimes v)=\phi(u,v)$ for all $u\in L^1(\mu)$ and $v\in L^1(\nu)$, and $\|T\|=\|\phi\|$.

proof (a) The centre of the argument is the following fact: if $E_0,\dots,E_n$ are measurable sets of finite measure in X, $F_0,\dots,F_n$ are measurable sets of finite measure in Y, $a_0,\dots,a_n\in\mathbb{R}$ and $\sum_{i=0}^n a_i\chi(E_i\times F_i)=0$ λ-a.e., then $\sum_{i=0}^n a_i\phi(\chi E_i^\bullet,\chi F_i^\bullet)=0$ in W. **P** We can find a disjoint family $\langle G_j\rangle_{j\le m}$ of measurable sets of finite measure in X such that each E_i is expressible as a union of some subfamily of the G_j; so that χE_i is expressible in the form $\sum_{j=0}^m b_{ij}\chi G_j$ (see 122Ca). Similarly, we can find a disjoint family $\langle H_k\rangle_{k\le l}$ of measurable sets of finite measure in Y such that each χF_i is expressible as $\sum_{k=0}^l c_{ik}\chi H_k$. Now

$$\sum_{j=0}^m\sum_{k=0}^l\bigl(\sum_{i=0}^n a_ib_{ij}c_{ik}\bigr)\chi(G_j\times H_k)=\sum_{i=0}^n a_i\chi(E_i\times F_i)=0\ \lambda\text{-a.e.}$$

Because the $G_j\times H_k$ are disjoint, and $\lambda(G_j\times H_k)=\mu G_j\cdot\nu H_k$ for all j, k, it follows that for every $j\le m$, $k\le l$ we have either $\mu G_j=0$ or $\nu H_k=0$ or $\sum_{i=0}^n a_ib_{ij}c_{ik}=0$. In any of these three cases, $\sum_{i=0}^n a_ib_{ij}c_{ik}\phi(\chi G_j^\bullet,\chi H_k^\bullet)=0$ in W. But this means that

$$0=\sum_{j=0}^m\sum_{k=0}^l\bigl(\sum_{i=0}^n a_ib_{ij}c_{ik}\bigr)\phi(\chi G_j^\bullet,\chi H_k^\bullet)=\sum_{i=0}^n a_i\phi(\chi E_i^\bullet,\chi F_i^\bullet),$$

as claimed. **Q**

(b) It follows that if $E_0,\dots,E_n$, $E_0',\dots,E_m'$ are measurable sets of finite measure in X, $F_0,\dots,F_n$, $F_0',\dots,F_m'$ are measurable sets of finite measure in Y, $a_0,\dots,a_n,a_0',\dots,a_m'\in\mathbb{R}$ and $\sum_{i=0}^n a_i\chi(E_i\times F_i)=\sum_{i=0}^m a_i'\chi(E_i'\times F_i')$ λ-a.e., then

$$\sum_{i=0}^n a_i\phi(\chi E_i^\bullet,\chi F_i^\bullet)=\sum_{i=0}^m a_i'\phi(\chi E_i'^\bullet,\chi F_i'^\bullet)$$

in W. Let M be the linear subspace of $L^1(\lambda)$ generated by

$$\{\chi(E\times F)^\bullet : E\in\Sigma,\ \mu E<\infty,\ F\in\mathrm{T},\ \nu F<\infty\};$$

then we have a unique map $T_0 : M \to W$ such that

$$T_0(\sum_{i=0}^n a_i\chi(E_i\times F_i)^\bullet) = \sum_{i=0}^n a_i\phi(\chi E_i^\bullet, \chi F_i^\bullet)$$

whenever $E_0,\dots,E_n$ are measurable sets of finite measure in X, $F_0,\dots,F_n$ are measurable sets of finite measure in Y and $a_0,\dots,a_n\in\mathbb{R}$. Of course T_0 is linear.

(c) Some of the same calculations show that $\|T_0u\|\le\|\phi\|\|u\|_1$ for every $u\in M$. **P** If $u\in M$, then, by the arguments of (a), we can express u as $\sum_{j=0}^m\sum_{k=0}^l a_{jk}\chi(G_j\times H_k)^\bullet$, where $\langle G_j\rangle_{j\le m}$ and $\langle H_k\rangle_{k\le l}$ are disjoint families of sets of finite measure. Now

$$\begin{aligned}\|T_0u\| &= \|\sum_{j=0}^m\sum_{k=0}^l a_{jk}\phi(\chi G_j^\bullet,\chi H_k^\bullet)\| \le \sum_{j=0}^m\sum_{k=0}^l |a_{jk}|\|\phi(\chi G_j^\bullet,\chi H_k^\bullet)\|\\ &\le \sum_{j=0}^m\sum_{k=0}^l |a_{jk}|\|\phi\|\|\chi G_j^\bullet\|_1\|\chi H_k^\bullet\|_1 = \|\phi\|\sum_{j=0}^m\sum_{k=0}^l |a_{jk}|\mu G_j\cdot\nu H_k\\ &= \|\phi\|\sum_{j=0}^m\sum_{k=0}^l |a_{jk}|\lambda(G_j\times H_k) = \|\phi\|\|u\|_1,\end{aligned}$$

as claimed. **Q**

(d) The next point is to observe that M is dense in $L^1(\lambda)$ for $\|\,\|_1$. **P** Repeating the ideas above once again, we observe that if $E_0,\dots,E_n$ are sets of finite measure in X and $F_0,\dots,F_n$ are sets of finite measure in Y, then $\chi(\bigcup_{i\le n}E_i\times F_i)^\bullet\in M$; this is because, expressing each E_i as a union of G_j, where the G_j are disjoint, we have

$$\bigcup_{i\le n}E_i\times F_i = \bigcup_{j\le m}G_j\times F'_j,$$

where $F'_j=\bigcup\{F_i : G_j\subseteq E_i\}$ for each j; now $\langle G_j\times F'_j\rangle_{j\le m}$ is disjoint, so

$$\chi(\bigcup_{j\le m}G_j\times F_j)^\bullet = \sum_{j=0}^m\chi(G_j\times F'_j)^\bullet\in M.$$

So 251Ie tells us that whenever $\lambda H<\infty$ and $\epsilon>0$ there is a G such that $\lambda(H\triangle G)\le\epsilon$ and $\chi G^\bullet\in M$; now

$$\|\chi H^\bullet - \chi G^\bullet\|_1 = \lambda(G\triangle H)\le\epsilon,$$

so $\chi H^\bullet$ is approximated arbitrarily closely by members of M, and belongs to the closure $\overline{M}$ of M in $L^1(\lambda)$. Because M is a linear subspace of $L^1(\lambda)$, so is $\overline{M}$ (2A4Cb); accordingly $\overline{M}$ contains the equivalence classes of all λ-simple functions; but these are dense in $L^1(\lambda)$ (242Mb), so $\overline{M}=L^1(\lambda)$, as claimed. **Q**

(e) Because W is a Banach space, it follows that there is a bounded linear operator $T : L^1(\lambda)\to W$ extending T_0, with $\|T\|=\|T_0\|\le\|\phi\|$ (2A4I). Now $T(u\otimes v)=\phi(u,v)$ for all $u\in L^1(\mu)$, $v\in L^1(\nu)$. **P** If $u=\chi E^\bullet$ and $v=\chi F^\bullet$, where E, F are measurable sets of finite measure, then

$$T(u\otimes v) = T(\chi(E\times F)^\bullet) = T_0(\chi(E\times F)^\bullet) = \phi(\chi E^\bullet,\chi F^\bullet) = \phi(u,v).$$

Because ϕ and $\otimes$ are bilinear and T is linear,

$$T(f^\bullet\otimes g^\bullet) = \phi(f^\bullet, g^\bullet)$$

whenever f and g are simple functions. Now whenever $u\in L^1(\mu)$, $v\in L^1(\nu)$ and $\epsilon>0$, there are simple functions f, g such that $\|u-f^\bullet\|_1\le\epsilon$, $\|v-g^\bullet\|_1\le\epsilon$ (242Mb again); so that

$$\begin{aligned}\|\phi(u,v)-\phi(f^\bullet,g^\bullet)\| &\le \|\phi(u-f^\bullet, v-g^\bullet)\| + \|\phi(u,g^\bullet-v)\| + \|\phi(f^\bullet-u,v)\|\\ &\le \|\phi\|(\epsilon^2+\epsilon\|u\|_1+\epsilon\|v\|_1).\end{aligned}$$

Similarly

$$\|u\otimes v - f^\bullet\otimes g^\bullet\|_1 \le \epsilon(\epsilon+\|u\|_1+\|v\|_1),$$

so

$$\|T(u\otimes v)-T(f^\bullet\otimes g^\bullet)\|\le\epsilon\|T\|(\epsilon+\|u\|_1+\|v\|_1);$$

because $T(f^\bullet\otimes g^\bullet)=\phi(f^\bullet,g^\bullet)$,

$$\|T(u\otimes v)-\phi(u,v)\|\le\epsilon(\|T\|+\|\phi\|)(\epsilon+\|u\|_1+\|v\|_1).$$

As ϵ is arbitrary, $T(u\otimes v)=\phi(u,v)$, as required. **Q**

(f) The argument of (e) ensured that $\|T\|\le\|\phi\|$. Because $\|u\otimes v\|_1\le\|u\|_1\|v\|_1$ for all $u\in L^1(\mu)$ and $v\in L^1(\nu)$, $\|\phi(u,v)\|\le\|T\|\|u\|_1\|v\|_1$ for all u, v, and $\|\phi\|\le\|T\|$; so $\|T\|=\|\phi\|$.

(g) Thus T has the required properties. To see that it is unique, we have only to observe that any bounded linear operator $S:L^1(\lambda)\to W$ such that $S(u\otimes v)=\phi(u,v)$ for all $u\in L^1(\mu)$, $v\in L^1(\nu)$ must agree with T on objects of the form $\chi(E\times F)^\bullet$ where E and F are of finite measure, and therefore on every member of M; because M is dense and both S and T are continuous, they agree everywhere in $L^1(\lambda)$.

253G The order structure of L^1 In 253F I have treated the L^1 spaces exclusively as normed linear spaces. In general, however, the order structure of an L^1 space (see 242C) is as important as its norm. The map $\otimes:L^1(\mu)\times L^1(\nu)\to L^1(\lambda)$ respects the order structures of the three spaces in the following strong sense.

Proposition Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and λ the c.l.d. product measure on $X\times Y$. Then

(a) $u\otimes v\ge 0$ in $L^1(\lambda)$ whenever $u\ge 0$ in $L^1(\mu)$ and $v\ge 0$ in $L^1(\nu)$.

(b) The positive cone $\{w:w\ge 0\}$ of $L^1(\lambda)$ is precisely the closed convex hull C of $\{u\otimes v:u\ge 0,\ v\ge 0\}$ in $L^1(\lambda)$.

*(c) Let W be any Banach lattice, and $T:L^1(\lambda)\to W$ a bounded linear operator. Then the following are equiveridical:

(i) $Tw\ge 0$ in W whenever $w\ge 0$ in $L^1(\lambda)$;

(ii) $T(u\otimes v)\ge 0$ in W whenever $u\ge 0$ in $L^1(\mu)$ and $v\ge 0$ in $L^1(\nu)$.

proof (a) If u, $v\ge 0$ then they are expressible as $f^\bullet$, $g^\bullet$ where $f\in\mathcal{L}^1(\mu)$, $g\in\mathcal{L}^1(\nu)$, $f\ge 0$ and $g\ge 0$. Now $f\otimes g\ge 0$ so $u\otimes v=(f\otimes g)^\bullet\ge 0$.

(b)(i) Write $L^1(\lambda)^+$ for $\{w:w\in L^1(\lambda),\ w\ge 0\}$. Then $L^1(\lambda)^+$ is a closed convex set in $L^1(\lambda)$ (242De); by (a), it contains $u\otimes v$ whenever $u\in L^1(\mu)^+$ and $v\in L^1(\nu)^+$, so it must include C.

(ii)(α) Of course $0=0\otimes 0\in C$. (β) If $u\in M$, as defined in the proof of 253F, and $u>0$, then u is expressible as $\sum_{j\le m,k\le l}a_{jk}\chi(G_j\times H_k)^\bullet$, where $G_0,\dots,G_m$ and $H_0,\dots,H_l$ are disjoint sequences of sets of finite measure, as in (a) of the proof of 253F. Now a_{jk} can be negative only if $\chi(G_j\times H_k)^\bullet=0$, so replacing every a_{jk} by $\max(0,a_{jk})$ if necessary, we can suppose that $a_{jk}\ge 0$ for all j, k. Not all the a_{jk} can be zero, so $a=\sum_{j\le m,k\le l}a_{jk}>0$, and

$$u=\sum_{j\le m,k\le l}\frac{a_{jk}}{a}\cdot a\chi(G_j\times H_k)^\bullet=\sum_{j\le m,k\le l}\frac{a_{jk}}{a}\cdot(a\chi G_j^\bullet)\otimes\chi H_k^\bullet\in C.$$

(γ) If $w\in L^1(\lambda)^+$ and $\epsilon>0$, express w as $h^\bullet$ where $h\ge 0$ in $\mathcal{L}^1(\lambda)$. There is a simple function $h_1\ge 0$ such that $h_1\le_{\text{a.e.}}h$ and $\int h\le\int h_1+\epsilon$. Express h_1 as $\sum_{i=0}^n a_i\chi H_i$ where $\lambda H_i<\infty$ and $a_i\ge 0$ for each i, and for each $i\le n$ choose sets $G_{i0},\dots,G_{im_i}\in\Sigma$, $F_{i0},\dots,F_{im_i}\in\mathrm{T}$, all of finite measure, such that $G_{i0},\dots,G_{im_i}$ are disjoint and $\lambda(H_i\triangle\bigcup_{j\le m_i}G_{ij}\times F_{ij})\le\epsilon/(n+1)(a_i+1)$, as in (d) of the proof of 253F. Set

$$w_0=\sum_{i=0}^n a_i\sum_{j=0}^{m_i}\chi(G_{ij}\times F_{ij})^\bullet.$$

Then $w_0\in C$ because $w_0\in M$ and $w_0\ge 0$. Also

$$\begin{aligned}\|w-w_0\|_1&\le\|w-h_1^\bullet\|_1+\|h_1^\bullet-w_0\|_1\\&\le\int(h-h_1)d\lambda+\sum_{i=0}^n a_i\int|\chi H_i-\sum_{j=0}^{m_i}\chi(G_{ij}\times F_{ij})|d\lambda\\&\le\epsilon+\sum_{i=0}^n a_i\lambda(H\triangle\bigcup_{j\le m_i}G_{ij}\times F_{ij})\le 2\epsilon.\end{aligned}$$

As ϵ is arbitrary and C is closed, $w \in C$. As w is arbitrary, $L^1(\lambda)^+ \subseteq C$ and $C = L^1(\lambda)^+$.

(c) Part (a) tells us that (i)⇒(ii). For the reverse implication, we need a fragment from the theory of Banach lattices: $W^+ = \{w : w \in W,\, w \ge 0\}$ is a closed set in W. **P** If w, $w' \in W$, then

$$w = (w - w') + w' \le |w - w'| + w' \le |w - w'| + |w'|,$$

$$-w = (w' - w) - w' \le |w - w'| - w' \le |w - w'| + |w'|,$$

$$|w| \le |w - w'| + |w'|, \quad |w| - |w'| \le |w - w'|,$$

because $|w| = w \vee (-w)$ and the order of W is translation-invariant (241Ec). Similarly, $|w'| - |w| \le |w - w'|$ and $||w| - |w'|| \le |w - w'|$, so $\||w| - |w'|\| \le \|w - w'\|$, by the definition of Banach lattice (242G). Setting $\phi(w) = |w| - w$, we see that $\|\phi(w) - \phi(w')\| \le 2\|w - w'\|$ for all w, $w' \in W$, so that ϕ is continuous.

Now, because the order is invariant under multiplication by positive scalars,

$$w \ge 0 \iff 2w \ge 0 \iff w \ge -w \iff w = |w| \iff \phi(w) = 0,$$

so $W^+ = \{w : \phi(w) = 0\}$ is closed. **Q**

Now suppose that (ii) is true, and set $C_1 = \{w : w \in L^1(\lambda),\, Tw \ge 0\}$. Then C_1 contains $u \otimes v$ whenever u, $v \ge 0$; but also it is convex, because T is linear, and closed, because T is continuous and $C_1 = T^{-1}[W^+]$. By (b), C_1 includes $\{w : w \in L^1(\lambda),\, w \ge 0\}$, as required by (i).

253H Conditional expectations The ideas of this section and the preceding one provide us with some of the most important examples of conditional expectations.

Theorem Let (X, Σ, μ) and (Y, T, ν) be complete probability spaces, with c.l.d. product $(X \times Y, \Lambda, \lambda)$. Set $\Lambda_1 = \{E \times Y : E \in \Sigma\}$. Then Λ_1 is a σ-subalgebra of Λ. Given a λ-integrable real-valued function f, set

$$g(x, y) = \int f(x, z)\nu(dz)$$

whenever $x \in X$, $y \in Y$ and the integral is defined in $\mathbb{R}$. Then g is a conditional expectation of f on Λ_1.

proof We know that $\Lambda_1 \subseteq \Lambda$, by 251Ia, and Λ_1 is a σ-algebra of sets because Σ is. Fubini's theorem (252B, 252C) tells us that $f_1(x) = \int f(x, z)\nu(dz)$ is defined for almost every x, and therefore that $g = f_1 \otimes \chi Y$ is defined almost everywhere in $X \times Y$. f_1 is μ-virtually measurable; because μ is complete, f_1 is Σ-measurable, so g is Λ_1-measurable (since $\{(x, y) : g(x, y) \le \alpha\} = \{x : f_1(x) \le \alpha\} \times Y$ for every $\alpha \in \mathbb{R}$). Finally, if $W \in \Lambda_1$, then $W = E \times Y$ for some $E \in \Sigma$, so

$$\int_W g\, d\lambda = \int (f_1 \otimes \chi Y) \times (\chi E \otimes \chi Y) d\lambda = \int f_1 \times \chi E\, d\mu \int \chi Y\, d\nu$$

(by 253D)

$$= \iint \chi E(x) f(x, y)\nu(dy)\mu(dx) = \int f \times \chi(E \times Y) d\lambda$$

(by Fubini's theorem)

$$= \int_W f\, d\lambda.$$

So g is a conditional expectation of f.

253I This is a convenient moment to set out a useful result on products of indefinite-integral measures.

Proposition Let (X, Σ, μ) and (Y, T, ν) be measure spaces, and $f \in \mathcal{L}^0(\mu)$, $g \in \mathcal{L}^0(\nu)$ non-negative functions. Let μ', ν' be the corresponding indefinite-integral measures (see §234). Let λ be the c.l.d. product of μ and ν, and λ' the indefinite-integral measure defined from λ and $f \otimes g \in \mathcal{L}^0(\lambda)$ (253Cb). Then λ' is the c.l.d. product of μ' and ν'.

proof Write θ for the c.l.d. product of μ' and ν'.

(a) If we replace μ by its completion, we do not change μ' (234Ke); at the same time, we do not change λ, by 251T. The same applies to ν. So it will be enough to prove the result on the assumption that μ and ν are complete; in which case f and g are measurable and have measurable domains.

Set $F = \{x : x \in \operatorname{dom} f,\, f(x) > 0\}$ and $G = \{y : y \in \operatorname{dom} g,\, g(y) > 0\}$, so that $F \times G = \{w : w \in \operatorname{dom}(f \otimes g),\, (f \otimes g)(w) > 0\}$. Then F is μ'-conegligible and G is ν'-conegligible, so $F \times G$ is θ-conegligible as well as λ'-conegligible. Because both θ and λ' are complete (251Ic, 234I), it will be enough to show that the subspace measures $\theta_{F\times G}$, $\lambda'_{F\times G}$ on $F \times G$ are equal. But note that $\theta_{F\times G}$ can be identified with the product of μ'_F and ν'_G, where μ'_F and ν'_G are the subspace measures on F, G respectively (251Q(ii-α)). At the same time, μ'_F is the indefinite-integral measure defined from the subspace measure μ_F on F and the function $f{\restriction}F$, ν'_G is the indefinite-integral measure defined from the subspace measure ν_G on G and $g{\restriction}G$, and $\lambda'_{F\times G}$ is defined from the subspace measure $\lambda_{F\times G}$ and $(f{\restriction}F) \otimes (g{\restriction}G)$. Finally, by 251Q again, $\lambda_{F\times G}$ is the product of μ_F and ν_G.

What all this means is that it will be enough to deal with the case in which $F = X$ and $G = Y$, that is, f and g are everywhere defined and strictly positive; which is what I will suppose from now on.

(b) In this case $\operatorname{dom}\mu' = \Sigma$ and $\operatorname{dom}\nu' = \mathrm{T}$ (234La). Similarly, $\operatorname{dom}\lambda' = \Lambda$ is just the domain of λ. Set

$$F_n = \{x : x \in X,\, 2^{-n} \le f(x) \le 2^n\}, \quad G_n = \{y : y \in Y,\, 2^{-n} \le g(y) \le 2^n\}$$

for $n \in \mathbb{N}$.

(c) Set

$$\mathcal{A} = \{W : W \in \operatorname{dom}\theta \cap \operatorname{dom}\lambda',\, \theta(W) = \lambda'(W)\}.$$

If $\mu'E$ and $\nu'H$ are defined and finite, then $f \times \chi E$ and $g \times \chi H$ are integrable, so

$$\lambda'(E \times H) = \int (f \otimes g) \times \chi(E \times H) d\lambda = \int (f \times \chi E) \otimes (g \times \chi H) d\lambda$$
$$= \int f \times \chi E\, d\mu \cdot \int g \times \chi H\, d\nu = \theta(E \times H)$$

by 253D and 251Ia, that is, $E \times H \in \mathcal{A}$. If we now look at $\mathcal{A}_{EH} = \{W : W \subseteq X \times Y,\, W \cap (E \times H) \in \mathcal{A}\}$, then we see that

$\mathcal{A}_{EH}$ contains $E' \times H'$ for every $E' \in \Sigma$, $H' \in \mathrm{T}$,

if $\langle W_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in $\mathcal{A}_{EH}$ then $\bigcup_{n\in\mathbb{N}} W_n \in \mathcal{A}_{EH}$,

if W, $W' \in \mathcal{A}_{EH}$ and $W \subseteq W'$ then $W' \setminus W \in \mathcal{A}_{EH}$.

Thus $\mathcal{A}_{EH}$ is a Dynkin class of subsets of $X \times Y$, and by the Monotone Class Theorem (136B) includes the σ-algebra generated by $\{E' \times H' : E' \in \Sigma,\, H' \in \mathrm{T}\}$, which is $\Sigma\widehat{\otimes}\mathrm{T}$.

(d) Now suppose that $W \in \Lambda$. In this case $W \in \operatorname{dom}\theta$ and $\theta W \le \lambda' W$. **P** Take $n \in \mathbb{N}$, and $E \in \Sigma$, $H \in \mathrm{T}$ such that $\mu'E$ and $\nu'H$ are both finite. Set $E' = E \cap F_n$, $H' = H \cap G_n$ and $W' = W \cap (E' \times H')$. Then $W' \in \Lambda$, while $\mu E' \le 2^n \mu' E$ and $\nu H' \le 2^n \nu' H$ are finite. By 251Ib there is a $V \in \Sigma\widehat{\otimes}\mathrm{T}$ such that $V \subseteq W'$ and $\lambda V = \lambda W'$. Similarly, there is a $V' \in \Sigma\widehat{\otimes}\mathrm{T}$ such that $V' \subseteq (E' \times H') \setminus W'$ and $\lambda V' = \lambda((E' \times H') \setminus W')$. This means that $\lambda((E' \times H') \setminus (V \cup V')) = 0$, so $\lambda'((E' \times H') \setminus (V \cup V')) = 0$. But $(E' \times H') \setminus (V \cup V') \in \mathcal{A}$, by (c), so $\theta((E' \times H') \setminus (V \cup V')) = 0$ and $W' \in \operatorname{dom}\theta$, while

$$\theta W' = \theta V = \lambda' V \le \lambda' W.$$

Since E and H are arbitrary, $W \cap (F_n \times G_n) \in \operatorname{dom}\theta$ (251H) and $\theta(W \cap (F_n \times G_n)) \le \lambda' W$. Since $\langle F_n\rangle_{n\in\mathbb{N}}$, $\langle G_n\rangle_{n\in\mathbb{N}}$ are non-decreasing sequences with unions X, Y respectively,

$$\theta W = \sup_{n\in\mathbb{N}} \theta(W \cap (F_n \times G_n)) \le \lambda' W. \text{ } \mathbf{Q}$$

(e) In the same way, $\lambda' W$ is defined and less than or equal to θW for every $W \in \operatorname{dom}\theta$. **P** The arguments are very similar, but a refinement seems to be necessary at the last stage. Take $n \in \mathbb{N}$, and $E \in \Sigma$, $H \in \mathrm{T}$ such that μE and νH are both finite. Set $E' = E \cap F_n$, $H' = H \cap G_n$ and $W' = W \cap (E' \times H')$. Then $W' \in \operatorname{dom}\theta$, while $\mu' E' \le 2^n \mu E$ and $\nu' H' \le 2^n \nu H$ are finite. This time, there are V, $V' \in \Sigma\widehat{\otimes}\mathrm{T}$ such that $V \subseteq W'$, $V' \subseteq (E' \times H') \setminus W'$, $\theta V = \theta W'$ and $\theta V' = \theta((E' \times H') \setminus W')$. Accordingly

$$\lambda' V + \lambda' V' = \theta V + \theta V' = \theta(E' \times H') = \lambda'(E' \times H'),$$

so that $\lambda' W'$ is defined and equal to $\theta W'$.

What this means is that $W \cap (F_n \times G_n) \cap (E \times H) \in \mathcal{A}$ whenever μE and νH are finite. So $W \cap (F_n \times G_n) \in \Lambda$, by 251H; as n is arbitrary, $W \in \Lambda$ and $\lambda' W$ is defined.

? Suppose, if possible, that $\lambda' W > \theta W$. Then there is some $n \in \mathbb{N}$ such that $\lambda'(W \cap (F_n \times G_n)) > \theta W$. Because λ is semi-finite, 213B tells us that there is some λ-simple function h such that $h \le (f \otimes g) \times \chi(W \cap (F_n \times G_n))$ and $\int h\,d\lambda > \theta W$; setting $V = \{(x, y) : h(x, y) > 0\}$, we see that $V \subseteq W \cap (F_n \times G_n)$, λV is defined and finite and $\lambda' V > \theta W$. Now there must be sets $E \in \Sigma$, $H \in \mathrm{T}$ such that μE and νF are both finite and $\lambda(V \setminus (E \times H)) < 4^{-n}(\lambda' V - \theta W)$. But in this case $V \in \Lambda \subseteq \operatorname{dom}\theta$ (by (d)), so we can apply the argument just above to V and conclude that $V \cap (E \times H) = V \cap (F_n \times G_n) \cap (E \times H)$ belongs to $\mathcal{A}$. And now

$$\begin{aligned}\lambda' V &= \lambda'(V \cap (E \times H)) + \lambda'(V \setminus (E \times H)) \\ &\le \theta(V \cap (E \times H)) + 4^n \lambda(V \setminus (E \times H)) < \theta V + \lambda' V - \theta W \le \lambda' V,\end{aligned}$$

which is absurd. **X**

So $\lambda' W$ is defined and not greater than θW. **Q**

(f) Putting this together with (d), we see that $\lambda' = \theta$, as claimed.

Remark If μ' and ν' are totally finite, so that they are 'truly continuous' with respect to μ and ν in the sense of 232Ab, then f and g are integrable, so $f \otimes g$ is λ-integrable, and $\theta = \lambda'$ is truly continuous with respect to λ.

The proof above can be simplified using fragments of the general theory of complete locally determined spaces, which will be given in §412 in Volume 4.

***253J Upper integrals** The idea of 253D can be repeated in terms of upper integrals, as follows.

Proposition Let (X, Σ, μ) and (Y, T, ν) be σ-finite measure spaces, with c.l.d. product measure λ. Then for any functions f and g, defined on conegligible subsets of X and Y respectively, and taking values in $[0, \infty]$,

$$\overline{\int} f \otimes g\,d\lambda = \overline{\int} f d\mu \cdot \overline{\int} g\,d\nu.$$

Remark Here $(f \otimes g)(x, y) = f(x)g(y)$ for all $x \in \operatorname{dom} f$ and $y \in \operatorname{dom} g$, taking $0 \cdot \infty = 0$, as in §135.

proof (a) I show first that $\overline{\int} f \otimes g \le \overline{\int} f \overline{\int} g$. **P** If $\overline{\int} f = 0$, then $f = 0$ a.e., so $f \otimes g = 0$ a.e. and the result is immediate. The same argument applies if $\overline{\int} g = 0$. If both $\overline{\int} f$ and $\overline{\int} g$ are non-zero, and either is infinite, the result is trivial. So let us suppose that both are finite. In this case there are integrable f_0, g_0 such that $f \le_{\text{a.e.}} f_0$, $g \le_{\text{a.e.}} g_0$, $\overline{\int} f = \int f_0$ and $\overline{\int} g = \int g_0$ (133Ja/135Ha). So $f \otimes g \le_{\text{a.e.}} f_0 \otimes g_0$, and

$$\overline{\int} f \otimes g \le \int f_0 \otimes g_0 = \int f_0 \int g_0 = \overline{\int} f \overline{\int} g,$$

by 253D. **Q**

(b) For the reverse inequality, we need consider only the case in which $\overline{\int} f \otimes g$ is finite, so that there is a λ-integrable function h such that $f \otimes g \le_{\text{a.e.}} h$ and $\overline{\int} f \otimes g = \int h$. Set

$$f_0(x) = \int h(x, y)\nu(dy)$$

whenever this is defined in $\mathbb{R}$, which is almost everywhere, by Fubini's theorem (252B-252C). Then $f_0(x) \ge f(x)\overline{\int} g\,d\nu$ for every $x \in \operatorname{dom} f_0 \cap \operatorname{dom} f$, which is a conegligible set in X; so

$$\overline{\int} f \otimes g = \int h\,d\lambda = \int f_0 d\mu \ge \overline{\int} f \overline{\int} g,$$

as required.

***253K** A similar argument applies to upper integrals of sums, as follows.

Proposition Let (X, Σ, μ) and (Y, T, ν) be probability spaces, with c.l.d. product measure λ. Then for any real-valued functions f, g defined on conegligible subsets of X, Y respectively,

$$\overline{\int} f(x) + g(y)\,\lambda(d(x,y)) = \overline{\int} f(x)\mu(dx) + \overline{\int} g(y)\nu(dy),$$

at least when the right-hand side is defined in $[-\infty,\infty]$.

proof Set $h(x,y) = f(x) + g(y)$ for $x \in \operatorname{dom} f$ and $y \in \operatorname{dom} g$, so that $\operatorname{dom} h$ is λ-conegligible.

(a) As in 253J, I start by showing that $\overline{\int} h \le \overline{\int} f + \overline{\int} g$. **P** If either $\overline{\int} f$ or $\overline{\int} g$ is ∞, this is trivial. Otherwise, take integrable functions f_0, g_0 such that $f \le_{\text{a.e.}} f_0$ and $g \le_{\text{a.e.}} g_0$. Set $h_0 = (f_0 \otimes \chi Y) + (\chi X \otimes g_0)$; then $h \le h_0$ λ-a.e., so

$$\overline{\int} h\,d\lambda \le \int h_0 d\lambda = \int f_0 d\mu + \int g_0 d\nu.$$

As f_0, g_0 are arbitrary, $\overline{\int} h \le \overline{\int} f + \overline{\int} g$. **Q**

(b) For the reverse inequality, suppose that $h \le h_0$ for λ-almost every (x,y), where h_0 is λ-integrable. Set $f_0(x) = \int h_0(x,y)\nu(dy)$ whenever this is defined in $\mathbb{R}$. Then $f_0(x) \ge f(x) + \overline{\int} g\,d\nu$ whenever $x \in \operatorname{dom} f \cap \operatorname{dom} f_0$, so

$$\int h_0\,d\lambda = \int f_0 d\mu \ge \overline{\int} f d\mu + \overline{\int} g\,d\nu.$$

As h_0 is arbitrary, $\overline{\int} h \ge \overline{\int} f + \overline{\int} g$, as required.

253L Complex spaces As usual, the ideas of 253F and 253H apply essentially unchanged to complex L^1 spaces. Writing $L^1_{\mathbb{C}}(\mu)$, etc., for the complex L^1 spaces involved, we have the following results. Throughout, let (X,Σ,μ) and (Y,T,ν) be measure spaces, and λ the c.l.d. product measure on $X \times Y$.

(a) If $f \in \mathcal{L}^0_{\mathbb{C}}(\mu)$ and $g \in \mathcal{L}^0_{\mathbb{C}}(\nu)$ then $f \otimes g$, defined by the formula $(f \otimes g)(x,y) = f(x)g(y)$ for $x \in \operatorname{dom} f$ and $y \in \operatorname{dom} g$, belongs to $\mathcal{L}^0_{\mathbb{C}}(\lambda)$.

(b) If $f \in \mathcal{L}^1_{\mathbb{C}}(\mu)$ and $g \in \mathcal{L}^1_{\mathbb{C}}(\nu)$ then $f \otimes g \in \mathcal{L}^1_{\mathbb{C}}(\lambda)$ and $\int f \otimes g\,d\lambda = \int f d\mu \int g\,d\nu$.

(c) We have a bilinear operator $(u,v) \mapsto u \otimes v : L^1_{\mathbb{C}}(\mu) \times L^1_{\mathbb{C}}(\nu) \to L^1_{\mathbb{C}}(\lambda)$ defined by writing $f^\bullet \otimes g^\bullet = (f \otimes g)^\bullet$ for all $f \in \mathcal{L}^1_{\mathbb{C}}(\mu)$, $g \in \mathcal{L}^1_{\mathbb{C}}(\nu)$.

(d) If W is any complex Banach space and $\phi : L^1_{\mathbb{C}}(\mu) \times L^1_{\mathbb{C}}(\nu) \to W$ is any bounded bilinear operator, then there is a unique bounded linear operator $T : L^1_{\mathbb{C}}(\lambda) \to W$ such that $T(u \otimes v) = \phi(u,v)$ for every $u \in L^1_{\mathbb{C}}(\mu)$ and $v \in L^1_{\mathbb{C}}(\nu)$, and $\|T\| = \|\phi\|$.

(e) If μ and ν are complete probability measures, and $\Lambda_1 = \{E \times Y : E \in \Sigma\}$, then for any $f \in \mathcal{L}^1_{\mathbb{C}}(\lambda)$ we have a conditional expectation g of f on Λ_1 given by setting $g(x,y) = \int f(x,z)\nu(dz)$ whenever this is defined.

253X Basic exercises >(a) Let U, V and W be linear spaces. Show that the set of bilinear operators from $U \times V$ to W has a natural linear structure agreeing with those of $\mathsf{L}(U;\mathsf{L}(V;W))$ and $\mathsf{L}(V;\mathsf{L}(U;W))$, writing $\mathsf{L}(U;W)$ for the linear space of linear operators from U to W.

>(b) Let U, V and W be normed spaces. (i) Show that for a bilinear operator $\phi : U \times V \to W$ the following are equiveridical: (α) ϕ is bounded in the sense of 253Ab; (β) ϕ is continuous; (γ) ϕ is continuous at some point of $U \times V$. (ii) Show that the space of bounded bilinear operators from $U \times V$ to W is a linear subspace of the space of all bilinear operators from $U \times V$ to W, and that the functional $\|\ \|$ defined in 253Ab is a norm, agreeing with the norms of $\mathsf{B}(U;\mathsf{B}(V;W))$ and $\mathsf{B}(V;\mathsf{B}(U;W))$, writing $\mathsf{B}(U;W)$ for the normed space of bounded linear operators from U to W.

(c) Let $(X_1,\Sigma_1,\mu_1),\dots,(X_n,\Sigma_n,\mu_n)$ be measure spaces, and λ the c.l.d. product measure on $X_1 \times \dots \times X_n$, as described in 251W. Let W be a Banach space, and suppose that $\phi : L^1(\mu_1) \times \dots \times L^1(\mu_n) \to W$ is **multilinear** (that is, linear in each variable separately) and **bounded** (that is, $\|\phi\| = \sup\{\phi(u_1,\dots,u_n) : \|u_i\|_1 \le 1\ \forall\ i \le n\} < \infty$). Show that there is a unique bounded linear operator $T : L^1(\lambda) \to W$ such that $T\otimes = \phi$, where $\otimes : L^1(\mu_1) \times \dots \times L^1(\mu_n) \to L^1(\lambda)$ is a canonical multilinear operator (to be defined).

(d) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and λ the c.l.d. product measure on $X\times Y$. Show that if $A\subseteq L^1(\mu)$ and $B\subseteq L^1(\nu)$ are both uniformly integrable, then $\{u\otimes v:u\in A,\,v\in B\}$ is uniformly integrable in $L^1(\lambda)$.

>(e) Let (X,Σ,μ) and (Y,T,ν) be measure spaces and λ the c.l.d. product measure on $X\times Y$. Show that

(i) we have a bilinear operator $(u,v)\mapsto u\otimes v:L^0(\mu)\times L^0(\nu)\to L^0(\lambda)$ given by setting $f^\bullet\otimes g^\bullet=(f\otimes g)^\bullet$ for all $f\in\mathcal{L}^0(\mu)$ and $g\in\mathcal{L}^0(\nu)$;

(ii) if $1\le p\le\infty$ then $u\otimes v\in L^p(\lambda)$ and $\|u\otimes v\|_p=\|u\|_p\|v\|_p$ for all $u\in L^p(\mu)$ and $v\in L^p(\nu)$;

(iii) if $u,\,u'\in L^2(\mu)$ and $v,\,v'\in L^2(\nu)$ then the inner product $(u\otimes v|u'\otimes v')$, taken in $L^2(\lambda)$, is just $(u|u')(v|v')$;

(iv) the map $(u,v)\mapsto u\otimes v:L^0(\mu)\times L^0(\nu)\to L^0(\lambda)$ is continuous if $L^0(\mu)$, $L^0(\nu)$ and $L^0(\lambda)$ are all given their topologies of convergence in measure.

(f) In 253Xe, assume that μ and ν are semi-finite. Show that if $u_0,\dots,u_n$ are linearly independent members of $L^0(\mu)$ and $v_0,\dots,v_n\in L^0(\nu)$ are not all 0, then $\sum_{i=0}^n u_i\otimes v_i\ne 0$ in $L^0(\lambda)$. (*Hint*: start by finding sets $E\in\Sigma$, $F\in\mathrm{T}$ of finite measure such that $u_0\times\chi E^\bullet,\dots,u_n\times\chi E^\bullet$ are linearly independent and $v_0\times\chi F^\bullet,\dots,v_n\times\chi F^\bullet$ are not all 0.)

(g) In 253Xe, assume that μ and ν are semi-finite. If U, V are linear subspaces of $L^0(\mu)$ and $L^0(\nu)$ respectively, write $U\otimes V$ for the linear subspace of $L^0(\lambda)$ generated by $\{u\otimes v:u\in U,\,v\in V\}$. Show that if W is any linear space and $\phi:U\times V\to W$ is a bilinear operator, there is a unique linear operator $T:U\otimes V\to W$ such that $T(u\otimes v)=\phi(u,v)$ for all $u\in U$, $v\in V$. (*Hint*: start by showing that if $u_0,\dots,u_n\in U$ and $v_0,\dots,v_n\in V$ are such that $\sum_{i=0}^n u_i\otimes v_i=0$, then $\sum_{i=0}^n\phi(u_i,v_i)=0$ – do this by expressing the u_i as linear combinations of some linearly independent family and applying 253Xf.)

>(h) Let (X,Σ,μ) and (Y,T,ν) be complete probability spaces, with c.l.d. product measure λ. Suppose that $p\in[1,\infty]$ and that $f\in\mathcal{L}^p(\lambda)$. Set $g(x)=\int f(x,y)\nu(dy)$ whenever this is defined. Show that $g\in\mathcal{L}^p(\mu)$ and that $\|g\|_p\le\|f\|_p$. (*Hint*: 253H, 244M.)

(i) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, with c.l.d. product measure λ, and $p\in[1,\infty[$. Show that $\{w:w\in L^p(\lambda),\,w\ge 0\}$ is the closed convex hull in $L^p(\lambda)$ of $\{u\otimes v:u\in L^p(\mu),\,v\in L^p(\nu),\,u\ge 0,\,v\ge 0\}$ (see 253Xe(ii) above).

253Y Further exercises (a) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and λ_0 the primitive product measure on $X\times Y$. Show that if $f\in\mathcal{L}^0(\mu)$ and $g\in\mathcal{L}^0(\nu)$, then $f\otimes g\in\mathcal{L}^0(\lambda_0)$.

(b) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and λ_0 the primitive product measure on $X\times Y$. Show that if $f\in\mathcal{L}^1(\mu)$ and $g\in\mathcal{L}^1(\nu)$, then $f\otimes g\in\mathcal{L}^1(\lambda_0)$ and $\int f\otimes g\,d\lambda_0=\int f\,d\mu\int g\,d\nu$.

(c) Let (X,Σ,μ) and (Y,T,ν) be measure spaces, and λ_0, λ the primitive and c.l.d. product measures on $X\times Y$. Show that the embedding $\mathcal{L}^1(\lambda_0)\subsetneqq\mathcal{L}^1(\lambda)$ induces a Banach lattice isomorphism between $L^1(\lambda_0)$ and $L^1(\lambda)$.

(d) Let (X,Σ,μ), (Y,T,ν) be strictly localizable measure spaces, with c.l.d. product measure λ. Show that $L^\infty(\lambda)$ can be identified with $L^1(\lambda)^*$. Show that under this identification $\{w:w\in L^\infty(\lambda),\,w\ge 0\}$ is the weak*-closed convex hull of $\{u\otimes v:u\in L^\infty(\mu),\,v\in L^\infty(\nu),\,u\ge 0,\,v\ge 0\}$.

(e) Find a version of 253J valid when one of μ, ν is not σ-finite.

(f) Let (X,Σ,μ) be any measure space and V any Banach space. Write $\mathcal{L}^1_V=\mathcal{L}^1_V(\mu)$ for the set of functions f such that (α) $\operatorname{dom}f$ is a conegligible subset of X (β) f takes values in V (γ) there is a conegligible set $D\subseteq\operatorname{dom}f$ such that $f[D]$ is separable and $D\cap f^{-1}[G]\in\Sigma$ for every open set $G\subseteq V$ (δ) the integral $\int\|f(x)\|\mu(dx)$ is finite. (These are the **Bochner integrable** functions from X to V.) For f, $g\in\mathcal{L}^1_V$ write $f\sim g$ if $f=g$ μ-a.e.; let L^1_V be the set of equivalence classes in $\mathcal{L}^1_V$ under $\sim$. Show that

(i) $f+g$, $cf\in\mathcal{L}^1_V$ for all f, $g\in\mathcal{L}^1_V$, $c\in\mathbb{R}$;

(ii) L^1_V has a natural linear space structure, defined by writing $f^\bullet + g^\bullet = (f+g)^\bullet$, $cf^\bullet = (cf)^\bullet$ for f, $g \in \mathcal{L}^1_V$ and $c \in \mathbb{R}$;

(iii) L^1_V has a norm $\| \, \|$, defined by writing $\|f^\bullet\| = \int \|f(x)\|\mu(dx)$ for $f \in \mathcal{L}^1_V$;

(iv) L^1_V is a Banach space under this norm;

(v) there is a natural map $\otimes : \mathcal{L}^1 \times V \to \mathcal{L}^1_V$ defined by writing $(f \otimes v)(x) = f(x)v$ when $f \in \mathcal{L}^1 = \mathcal{L}^1_{\mathbb{R}}(\mu)$, $v \in V$ and $x \in \operatorname{dom} f$;

(vi) there is a canonical bilinear operator $\otimes : L^1 \times V \to L^1_V$ defined by writing $f^\bullet \otimes v = (f \otimes v)^\bullet$ for $f \in \mathcal{L}^1$ and $v \in V$;

(vii) whenever W is a Banach space and $\phi : L^1 \times V \to W$ is a bounded bilinear operator, there is a unique bounded linear operator $T : L^1_V \to W$ such that $T(u \otimes v) = \phi(u,v)$ for all $u \in L^1$ and $v \in V$, and $\|T\| = \|\phi\|$. (When $W = V$ and $\phi(u,v) = (\int u)v$ for $u \in L^1$ and $v \in V$, $Tf^\bullet$ is called the **Bochner integral** of f.)

(g) Let (X, Σ, μ) and (Y, T, ν) be measure spaces, and λ_0 the primitive product measure on $X \times Y$. If f is a λ_0-integrable function, write $f_x(y) = f(x,y)$ whenever this is defined. Show that we have a map $x \mapsto f_x^\bullet$ from a conegligible subset D_0 of X to $L^1(\nu)$. Show that this map is a Bochner integrable function, as defined in 253Yf, and that its Bochner integral is $\int f d\lambda_0$.

(h) Let (X, Σ, μ) and (Y, T, ν) be measure spaces, and suppose that ϕ is a function from X to a separable subset of $L^1(\nu)$ which is measurable in the sense that $\phi^{-1}[G] \in \Sigma$ for every open $G \subseteq L^1(\nu)$. Show that there is a Λ-measurable function f from $X \times Y$ to $\mathbb{R}$, where Λ is the domain of the c.l.d. product measure on $X \times Y$, such that $\phi(x) = f_x^\bullet$ for every $x \in X$, writing $f_x(y) = f(x,y)$ for $x \in X$, $y \in Y$.

(i) Let (X, Σ, μ) and (Y, T, ν) be measure spaces, and λ the c.l.d. product measure on $X \times Y$. Show that 253Yg provides a canonical identification between $L^1(\lambda)$ and $L^1_{L^1(\nu)}(\mu)$.

(j) Let (X, Σ, μ) and (Y, T, ν) be complete locally determined measure spaces, with c.l.d. product measure λ. (i) Suppose that $K \in \mathcal{L}^2(\lambda)$, $f \in \mathcal{L}^2(\mu)$. Show that $h(y) = \int K(x,y)f(x)dx$ is defined for almost all $y \in Y$ and that $h \in \mathcal{L}^2(\nu)$. (*Hint*: to see that h is defined a.e., consider $\int_{E \times F} K(x,y)f(x)d(x,y)$ for μE, $\nu F < \infty$; to see that $h \in \mathcal{L}^2$ consider $\int h \times g$ where $g \in \mathcal{L}^2(\nu)$.) (ii) Show that the map $f \mapsto h$ corresponds to a bounded linear operator $T_K : L^2(\mu) \to L^2(\nu)$. (iii) Show that the map $K \mapsto T_K$ corresponds to a bounded linear operator, of norm at most 1, from $L^2(\lambda)$ to $\mathsf{B}(L^2(\mu); L^2(\nu))$.

(k) Suppose that p, $q \in [1, \infty]$ and that $\frac{1}{p} + \frac{1}{q} = 1$, interpreting $\frac{1}{\infty}$ as 0 as usual. Let (X, Σ, μ), (Y, T, ν) be complete locally determined measure spaces with c.l.d. product measure λ. Show that the ideas of 253Yj can be used to define a bounded linear operator, of norm at most 1, from $L^p(\lambda)$ to $\mathsf{B}(L^q(\mu); L^p(\nu))$.

(l) In 253Xc, suppose that W is a Banach lattice. Show that the following are equiveridical: (i) $Tu \geq 0$ whenever $u \in L^1(\lambda)$; (ii) $\phi(u_1, \dots, u_n) \geq 0$ whenever $u_i \geq 0$ in $L^1(\mu_i)$ for each $i \leq n$.

253 Notes and comments Throughout the main arguments of this section, I have written the results in terms of the c.l.d. product measure; of course the isomorphism noted in 253Yc means that they could just as well have been expressed in terms of the primitive product measure. The more restricted notion of integrability with respect to the primitive product measure is indeed the one appropriate for the ideas of 253Yg.

Theorem 253F is a ‘universal mapping theorem’; it asserts that every bounded bilinear operator on $L^1(\mu) \times L^1(\nu)$ factors through $\otimes : L^1(\mu) \times L^1(\nu) \to L^1(\lambda)$, at least if the range space is a Banach space. It is easy to see that this property defines the pair $(L^1(\lambda), \otimes)$ up to Banach space isomorphism, in the following sense: if V is a Banach space, and $\psi : L^1(\mu) \times L^1(\nu) \to V$ is a bounded bilinear operator such that for every bounded bilinear operator ϕ from $L^1(\mu) \times L^1(\nu)$ to any Banach space W there is a unique bounded linear operator $T : V \to W$ such that $T\psi = \phi$ and $\|T\| = \|\phi\|$, then there is an isometric Banach space isomorphism $S : L^1(\lambda) \to V$ such that $S\otimes = \psi$. There is of course a general theory of bilinear operators between Banach spaces; in the language of this theory, $L^1(\lambda)$ is, or is isomorphic to, the ‘projective tensor product’ of $L^1(\mu)$ and $L^1(\nu)$. For an introduction to this subject, see Defant & Floret 93, §I.3, or Semadeni 71, §20. I should perhaps emphasise, for the sake of those who have not encountered tensor products before, that this theorem is special to L^1 spaces. While some of the same ideas can be applied to other function spaces (see 253Xe-253Xg), there is no other class to which 253F applies.

There is also a theory of tensor products of Banach lattices, for which I do not think we are quite ready (it needs general ideas about ordered linear spaces for which I mean to wait until Chapter 35 in the next volume). However 253G shows that the ordering, and therefore the Banach lattice structure, of $L^1(\lambda)$ is determined by the ordering of $L^1(\mu)$ and $L^1(\nu)$ and the map $\otimes : L^1(\mu) \times L^1(\nu) \to L^1(\lambda)$.

The conditional expectation operators described in 253H are of very great importance, largely because in this special context we have a realization of the conditional expectation operator as a function P_0 from $\mathcal{L}^1(\lambda)$ to $\mathcal{L}^1(\lambda\restriction\Lambda_1)$, not just as a function from $L^1(\lambda)$ to $L^1(\lambda\restriction\Lambda_1)$, as in 242J. As described here, $P_0(f+f')$ need not be equal, in the strict sense, to $P_0 f + P_0 f'$; it can have a larger domain. In applications, however, one might be willing to restrict attention to the linear space $\mathcal{U}$ of bounded $\Sigma\widehat{\otimes}\mathrm{T}$-measurable functions defined everywhere on $X \times Y$, so that P_0 becomes an operator from $\mathcal{U}$ to itself (see 252P).

Version of 23.2.16

254 Infinite products

I come now to the second basic idea of this chapter: the description of a product measure on the product of a (possibly large) family of probability spaces. The section begins with a construction on similar lines to that of §251 (254A-254F) and its defining property in terms of inverse-measure-preserving functions (254G). I discuss the usual measure on $\{0,1\}^I$ (254J-254K), subspace measures (254L) and various properties of subproducts (254M-254T), including a study of the associated conditional expectation operators (254R-254T).

254A Definitions (a) Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces. Set $X=\prod_{i\in I}X_i$, the family of functions x with domain I such that $x(i)\in X_i$ for every $i\in I$. In this context, I will say that a **measurable cylinder** is a subset of X expressible in the form

$$C=\prod_{i\in I}C_i,$$

where $C_i\in\Sigma_i$ for every $i\in I$ and $\{i:C_i\ne X_i\}$ is finite. Note that for a non-empty $C\subseteq X$ this expression is unique. **P** Suppose that $C=\prod_{i\in I}C_i=\prod_{i\in I}C'_i$. For each $i\in I$ set

$$D_i=\{x(i):x\in C\}.$$

Of course $D_i\subseteq C_i$. Because $C\ne\emptyset$, we can fix on some $z\in C$. If $i\in I$ and $t\in C_i$, consider $x\in X$ defined by setting

$$x(i)=t,\quad x(j)=z(j)\text{ for }j\ne i;$$

then $x\in C$ so $t=x(i)\in D_i$. Thus $D_i=C_i$ for $i\in I$. Similarly, $D_i=C'_i$. **Q**

(b) We can therefore define a functional $\theta_0:\mathcal{C}\to[0,1]$, where $\mathcal{C}$ is the set of measurable cylinders, by setting

$$\theta_0C=\prod_{i\in I}\mu_iC_i$$

whenever $C_i\in\Sigma_i$ for every $i\in I$ and $\{i:C_i\ne X_i\}$ is finite, noting that only finitely many terms in the product can differ from 1, so that it can safely be treated as a finite product. If $C=\emptyset$, one of the C_i must be empty, so θ_0C is surely 0, even though the expression of C as $\prod_{i\in I}C_i$ is no longer unique.

(c) Now define $\theta:\mathcal{P}X\to[0,1]$ by setting

$$\theta A=\inf\{\sum_{n=0}^\infty\theta_0C_n : C_n\in\mathcal{C}\text{ for every }n\in\mathbb{N},\ A\subseteq\bigcup_{n\in\mathbb{N}}C_n\}.$$

254B Lemma The functional θ defined in 254Ac is always an outer measure on X.

proof Use exactly the same arguments as those in 251B above.

254C Definition Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be any indexed family of probability spaces, and X the Cartesian product $\prod_{i\in I}X_i$. The **product measure** on X is the measure defined by Carathéodory's method (113C) from the outer measure θ defined in 254A.

254D Remarks (a) In 254Ab, I asserted that if $C\in\mathcal{C}$ and no C_i is empty, then nor is $C=\prod_{i\in I}C_i$. This is the 'Axiom of Choice': the product of any family $\langle C_i\rangle_{i\in I}$ of non-empty sets is non-empty, that is, there is a 'choice function' x with domain I picking out a distinguished member $x(i)$ of each C_i. In this volume I have not attempted to be scrupulous in indicating uses of the axiom of choice. In fact the use here is not an absolutely vital one; I mean, the theory of infinite products, even uncountable products, of probability spaces does not change character completely in the absence of the full axiom of choice (provided, that is, that we allow ourselves to use the countable axiom of choice). The point is that all we really need, in the present context, is that $X=\prod_{i\in I}X_i$ should be non-empty; and in many contexts we can prove this, for the particular cases of interest, without using the axiom of choice, by actually exhibiting a member of X. The simplest case in which this is difficult is when the X_i are uncontrolled Borel subsets of $[0,1]$; and even then, if they are presented with coherent descriptions, we may, with appropriate labour, be able to construct a member of X. But clearly such a process is liable to slow us down a good deal, and for the moment I think there is no great virtue in taking so much trouble.

(b) I have given this section the title 'infinite products', but it is useful to be able to apply the ideas to finite I; I should mention in particular the cases $\#(I)\le 2$.

(i) If $I=\emptyset$, X consists of the unique function with domain I, the empty function. If we identify a function with its graph, then X is actually $\{\emptyset\}$; in any case, X is to be a singleton set, with $\lambda X=1$.

(ii) If I is a singleton $\{i\}$, then we can identify X with X_i; $\mathcal{C}$ becomes identified with Σ_i and θ_0 with μ_i, so that θ can be identified with μ_i^* and the 'product measure' becomes the measure on X_i defined from μ_i^*, that is, the completion of μ_i (see 213Xa(iv)).

(iii) If I is a doubleton $\{i,j\}$, then we can identify X with $X_i\times X_j$; in this case the definitions of 254A and 254C match exactly with those of 251A and 251C, so that λ here can be identified with the primitive product measure as defined in 251C. Because μ_i and μ_j are both totally finite, this agrees with the c.l.d. product measure of 251F.

(c) In Volume 4, when considering products of probability spaces endowed with certain kinds of topology, I will introduce some alternative product measures. In such contexts I may speak of the product measure here as the 'ordinary' product measure.

254E Definition Let $\langle X_i\rangle_{i\in I}$ be any family of sets, and $X=\prod_{i\in I}X_i$. If Σ_i is a σ-subalgebra of subsets of X_i for each $i\in I$, I write $\widehat{\bigotimes}_{i\in I}\Sigma_i$ for the σ-algebra of subsets of X generated by

$$\{\{x:x\in X,\,x(i)\in E\}:i\in I,\,E\in\Sigma_i\}.$$

(Compare 251D.)

254F Theorem Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces, and let λ be the product measure on $X=\prod_{i\in I}X_i$ defined as in 254C; let Λ be its domain.

(a) $\lambda X=1$.

(b) If $E_i\in\Sigma_i$ for every $i\in I$, and $\{i:E_i\ne X_i\}$ is countable, then $\prod_{i\in I}E_i\in\Lambda$, and $\lambda(\prod_{i\in I}E_i)=\prod_{i\in I}\mu_iE_i$. In particular, $\lambda C=\theta_0C$ for every measurable cylinder C, as defined in 254A, and if $j\in I$ then $x\mapsto x(j):X\to X_j$ is inverse-measure-preserving.

(c) $\widehat{\bigotimes}_{i\in I}\Sigma_i\subseteq\Lambda$.

(d) λ is complete.

(e) For every $W\in\Lambda$ and $\epsilon>0$ there is a finite family $C_0,\dots,C_n$ of measurable cylinders such that $\lambda(W\triangle\bigcup_{k\le n}C_k)\le\epsilon$.

(f) For every $W\in\Lambda$ there are W_1, $W_2\in\widehat{\bigotimes}_{i\in I}\Sigma_i$ such that $W_1\subseteq W\subseteq W_2$ and $\lambda(W_2\setminus W_1)=0$.

Remark Perhaps I should pause to interpret the product $\prod_{i\in I}\mu_i E_i$. Because all the $\mu_i E_i$ belong to $[0,1]$, this is simply $\inf_{J\subseteq I, J \text{ is finite}}\prod_{i\in J}\mu_i E_i$, taking the empty product to be 1.

proof Throughout this proof, define $\mathcal{C}$, θ_0 and θ as in 254A. I will write out an argument which applies to finite I as well as infinite I, but you may reasonably prefer to assume that I is infinite on first reading.

(a) Of course $\lambda X = \theta X$, so I have to show that $\theta X = 1$. Because X, $\emptyset \in \mathcal{C}$ and $\theta_0 X = \prod_{i\in I}\mu_i X_i = 1$ and $\theta_0\emptyset = 0$,

$$\theta X \le \theta_0 X + \theta_0\emptyset + \ldots = 1.$$

I therefore have to show that $\theta X \ge 1$. **?** Suppose, if possible, otherwise.

(i) There is a sequence $\langle C_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{C}$, covering X, such that $\sum_{n=0}^{\infty}\theta_0 C_n < 1$. For each $n \in \mathbb{N}$, express C_n as $\{x : x(i) \in E_{ni}\ \forall\ i \in I\}$, where every $E_{ni} \in \Sigma_i$ and $J_n = \{i : E_{ni} \ne X_i\}$ is finite. No J_n can be empty, because $\theta_0 C_n < 1 = \theta_0 X$; set $J = \bigcup_{n\in\mathbb{N}} J_n$. Then J is a countable non-empty subset of I. Set $K = \mathbb{N}$ if J is infinite, $\{k : 0 \le k < \#(J)\}$ if J is finite; let $k \mapsto i_k : K \to J$ be a bijection.

For each $k \in K$, set $L_k = \{i_j : j < k\} \subseteq J$, and set $\alpha_{nk} = \prod_{i\in I\setminus L_k}\mu_i E_{ni}$ for $n \in \mathbb{N}$, $k \in K$. If J is finite, then we can identify $L_{\#(J)}$ with J, and set $\alpha_{n,\#(J)} = 1$ for every n. We have $\alpha_{n0} = \theta_0 C_n$ for each n, so $\sum_{n=0}^{\infty}\alpha_{n0} < 1$. For $n \in \mathbb{N}$, $k \in K$ and $t \in X_{i_k}$ set

$$\begin{aligned} f_{nk}(t) &= \alpha_{n,k+1} \text{ if } t \in E_{n,i_k},\\ &= 0 \text{ otherwise.}\end{aligned}$$

Then

$$\int f_{nk}d\mu_{i_k} = \alpha_{n,k+1}\mu_{i_k}E_{n,i_k} = \alpha_{nk}.$$

(ii) Choose $t_k \in X_{i_k}$ inductively, for $k \in K$, as follows. The inductive hypothesis will be that $\sum_{n\in M_k}\alpha_{nk} < 1$, where $M_k = \{n : n \in \mathbb{N},\ t_j \in E_{n,i_j}\ \forall\ j < k\}$; of course $M_0 = \mathbb{N}$, so the induction starts. Given that

$$1 > \sum_{n\in M_k}\alpha_{nk} = \sum_{n\in M_k}\int f_{nk}d\mu_{i_k} = \int(\sum_{n\in M_k} f_{nk})d\mu_{i_k}$$

(by B.Levi's theorem), there must be a $t_k \in X_{i_k}$ such that $\sum_{n\in M_k} f_{nk}(t_k) < 1$. Now for such a choice of t_k, $\alpha_{n,k+1} = f_{nk}(t_k)$ for every $n \in M_{k+1}$, so that $\sum_{n\in M_{k+1}}\alpha_{n,k+1} < 1$, and the induction continues, unless J is finite and $k + 1 = \#(J)$. In this last case we must just have $M_{\#(J)} = \emptyset$, because $\alpha_{n,\#(J)} = 1$ for every n.

(iii) If J is infinite, we obtain a full sequence $\langle t_k\rangle_{k\in\mathbb{N}}$; if J is finite, we obtain just a finite sequence $\langle t_k\rangle_{k<\#(J)}$. In either case, there is an $x \in X$ such that $x(i_k) = t_k$ for each $k \in K$. Now there must be some $m \in \mathbb{N}$ such that $x \in C_m$. Because $J_m = \{i : E_{mi} \ne X_i\}$ is finite, there is a $k \in \mathbb{N}$ such that $J_m \subseteq L_k$ (allowing $k = \#(J)$ if J is finite). Now $m \in M_k$, so in fact we cannot have $k = \#(J)$, and $\alpha_{mk} = 1$, so $\sum_{n\in M_k}\alpha_{nk} \ge 1$, contrary to the inductive hypothesis. **X**

This contradiction shows that $\theta X = 1$.

(b)(i) I take the particular case first. Suppose that $j \in I$ and $E \in \Sigma_j$, and let $C \in \mathcal{C}$; set $W = \{x : x \in X, x(j) \in E\}$; then $C \cap W$ and $C \setminus W$ both belong to $\mathcal{C}$, and $\theta_0 C = \theta_0(C \cap W) + \theta_0(C \setminus W)$. **P** If $C = \prod_{i\in I} C_i$, where $C_i \in \Sigma_i$ for each i, then $C \cap W = \prod_{i\in I} C_i'$, where $C_i' = C_i$ if $i \ne j$, and $C_j' = C_j \cap E$; similarly, $C \setminus W = \prod_{i\in I} C_i''$, where $C_i'' = C_i$ if $i \ne j$, and $C_j'' = C_j \setminus E$. So both belong to $\mathcal{C}$, and

$$\theta_0(C \cap W) + \theta_0(C \setminus W) = (\mu_j(C_j \cap E) + \mu_j(C_j \setminus E))\prod_{i\ne j}\mu C_i = \prod_{i\in I}\mu C_i = \theta_0 C. \ \mathbf{Q}$$

(ii) Now suppose that $A \subseteq X$ is any set, and $\epsilon > 0$. Then there is a sequence $\langle C_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{C}$ such that $A \subseteq \bigcup_{n\in\mathbb{N}} C_n$ and $\sum_{n=0}^{\infty}\theta_0 C_n \le \theta A + \epsilon$. In this case

$$A \cap W \subseteq \bigcup_{n\in\mathbb{N}} C_n \cap W, \quad A \setminus W \subseteq \bigcup_{n\in\mathbb{N}} C_n \setminus W,$$

so

$$\theta(A \cap W) \le \sum_{n=0}^{\infty}\theta_0(C_n \cap W), \quad \theta(A \setminus W) \le \sum_{n=0}^{\infty}\theta_0(C_n \setminus W),$$

and

$$\theta(A\cap W)+\theta(A\setminus W)\le\sum_{n=0}^\infty\theta_0(C_n\cap W)+\theta_0(C_n\setminus W)=\sum_{n=0}^\infty\theta_0C_n\le\theta A+\epsilon.$$

As ϵ is arbitrary, $\theta(A\cap W)+\theta(A\setminus W)\le\theta A$; as A is arbitrary, $W\in\Lambda$.

(iii) I show next that if $J\subseteq I$ is finite and $C_i\in\Sigma_i$ for each $i\in J$, and $C=\{x:x\in X,\,x(i)\in C_i\ \forall\ i\in J\}$, then $C\in\Lambda$ and $\lambda C=\prod_{i\in J}\mu_iC_i$. **P** Induce on $\#(J)$. If $\#(J)=0$, that is, $J=\emptyset$, then $C=X$ and this is part (a). For the inductive step to $\#(J)=n+1$, take any $j\in J$ and set $J'=J\setminus\{j\}$,

$$C'=\{x:x\in X,\,x(i)\in C_i\ \forall\ i\in J'\},$$

$$C''=C'\setminus C=\{x:x\in C',\,x(j)\in X_j\setminus C_j\}.$$

Then C, C', C'' all belong to $\mathcal{C}$, and $\theta_0C'=\prod_{i\in J'}\mu_iC_i=\alpha$ say, $\theta_0C=\alpha\mu_jC_j$, $\theta_0C''=\alpha(1-\mu_jC_j)$. Moreover, by the inductive hypothesis, $C'\in\Lambda$ and $\alpha=\lambda C'=\theta C'$. So $C=C'\cap\{x:x(j)\in C_j\}\in\Lambda$ by (ii), and $C''=C'\setminus C\in\Lambda$.

We surely have $\lambda C=\theta C\le\theta_0C$, $\lambda C''\le\theta_0C''$; but also

$$\alpha=\lambda C'=\lambda C+\lambda C''\le\theta_0C+\theta_0C''=\alpha,$$

so in fact

$$\lambda C=\theta_0C=\alpha\mu_jC_j=\prod_{i\in J}\mu C_i,$$

and the induction proceeds. **Q**

(iv) Now let us return to the general case of a set W of the form $\prod_{i\in I}E_i$ where $E_i\in\Sigma_i$ for each i, and $K=\{i:E_i\ne X_i\}$ is countable. If K is finite then $W=\{x:x(i)\in E_i\ \forall\ i\in K\}$ so $W\in\Lambda$ and

$$\lambda W=\prod_{i\in K}\mu_iE_i=\prod_{i\in I}\mu_iE_i.$$

Otherwise, let $\langle i_n\rangle_{n\in\mathbb{N}}$ be an enumeration of K. For each $n\in\mathbb{N}$ set $W_n=\{x:x\in X,\,x(i_k)\in E_{i_k}\ \forall\ k\le n\}$; then we know that $W_n\in\Lambda$ and that $\lambda W_n=\prod_{k=0}^n\mu_{i_k}E_{i_k}$. But $\langle W_n\rangle_{n\in\mathbb{N}}$ is a non-increasing sequence with intersection W, so $W\in\Lambda$ and

$$\lambda W=\lim_{n\to\infty}\lambda W_n=\prod_{i\in K}\mu_iE_i=\prod_{i\in I}\mu_iE_i.$$

(c) is an immediate consequence of (b) and the definition of $\widehat{\bigotimes}_{i\in I}\Sigma_i$.

(d) Because λ is constructed by Carathéodory's method it must be complete.

(e) Let $\langle C_n\rangle_{n\in\mathbb{N}}$ be a sequence in $\mathcal{C}$ such that $W\subseteq\bigcup_{n\in\mathbb{N}}C_n$ and $\sum_{n=0}^\infty\theta_0C_n\le\theta W+\frac12\epsilon$. Set $V=\bigcup_{n\in\mathbb{N}}C_n$; by (b), $V\in\Lambda$. Let $n\in\mathbb{N}$ be such that $\sum_{i=n+1}^\infty\theta_0C_i\le\frac12\epsilon$, and consider $W'=\bigcup_{k\le n}C_k$. Since $V\setminus W'\subseteq\bigcup_{i>n}C_i$,

$$\begin{aligned}\lambda(W\triangle W')&\le\lambda(V\setminus W)+\lambda(V\setminus W')=\lambda V-\lambda W+\lambda(V\setminus W')=\theta V-\theta W+\theta(V\setminus W')\\&\le\sum_{i=0}^\infty\theta_0C_i-\theta W+\sum_{i=n+1}\theta_0C_i\le\frac12\epsilon+\frac12\epsilon=\epsilon.\end{aligned}$$

(f)(i) If $W\in\Lambda$ and $\epsilon>0$ there is a $V\in\widehat{\bigotimes}_{i\in I}\Sigma_i$ such that $W\subseteq V$ and $\lambda V\le\lambda W+\epsilon$. **P** Let $\langle C_n\rangle_{n\in\mathbb{N}}$ be a sequence in $\mathcal{C}$ such that $W\subseteq\bigcup_{n\in\mathbb{N}}C_n$ and $\sum_{n=0}^\infty\theta_0C_n\le\theta W+\epsilon$. Then $C_n\in\widehat{\bigotimes}_{i\in I}\Sigma_i$ for each n, so $V=\bigcup_{n\in\mathbb{N}}C_n\in\widehat{\bigotimes}_{i\in I}\Sigma_i$. Now $W\subseteq V$, and

$$\lambda V=\theta V\le\sum_{n=0}^\infty\theta_0C_n\le\theta W+\epsilon=\lambda W+\epsilon.\ \mathbf{Q}$$

(ii) Now, given $W\in\Lambda$, let $\langle V_n\rangle_{n\in\mathbb{N}}$ be a sequence of sets in $\widehat{\bigotimes}_{i\in I}\Sigma_i$ such that $W\subseteq V_n$ and $\lambda V_n\le\lambda W+2^{-n}$ for each n; then $W_2=\bigcap_{n\in\mathbb{N}}V_n$ belongs to $\widehat{\bigotimes}_{i\in I}\Sigma_i$ and $W\subseteq W_2$ and $\lambda W_2=\lambda W$. Similarly, there is a $W_2'\in\widehat{\bigotimes}_{i\in I}\Sigma_i$ such that $X\setminus W\subseteq W_2'$ and $\lambda W_2'=\lambda(X\setminus W)$, so we may take $W_1=X\setminus W_2'$ to complete the proof.

254G The following is a fundamental, indeed defining, property of product measures. (Compare 251L.)

Theorem Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces with product (X,Λ,λ). Let (Y,T,ν) be a complete probability space and $\phi:Y\to X$ a function. Suppose that $\nu^*\phi^{-1}[C]\le\lambda C$ for every measurable cylinder $C\subseteq X$. Then ϕ is inverse-measure-preserving. In particular, ϕ is inverse-measure-preserving iff $\phi^{-1}[C]\in\mathrm{T}$ and $\nu\phi^{-1}[C]=\lambda C$ for every measurable cylinder $C\subseteq X$.

Remark By ν^* I mean the usual outer measure defined from ν as in §132.

proof (a) First note that, writing θ for the outer measure of 254A, $\nu^*\phi^{-1}[A]\le\theta A$ for every $A\subseteq X$. **P** Given $\epsilon>0$, there is a sequence $\langle C_n\rangle_{n\in\mathbb{N}}$ of measurable cylinders such that $A\subseteq\bigcup_{n\in\mathbb{N}}C_n$ and $\sum_{n=0}^\infty\theta_0C_n\le\theta A+\epsilon$, where θ_0 is the functional of 254A. But we know that $\theta_0C=\lambda C$ for every measurable cylinder C (254Fb), so

$$\nu^*\phi^{-1}[A]\le\nu^*(\bigcup_{n\in\mathbb{N}}\phi^{-1}[C_n])\le\sum_{n=0}^\infty\nu^*\phi^{-1}[C_n]\le\sum_{n=0}^\infty\lambda C_n\le\theta A+\epsilon.$$

As ϵ is arbitrary, $\nu^*\phi^{-1}[A]\le\theta A$. **Q**

(b) Now take any $W\in\Lambda$. Then there are F, $F'\in\mathrm{T}$ such that

$$\phi^{-1}[W]\subseteq F,\quad \phi^{-1}[X\setminus W]\subseteq F',$$

$$\nu F=\nu^*\phi^{-1}[W]\le\theta W=\lambda W,\quad \nu F'\le\lambda[X\setminus W].$$

We have

$$F\cup F'\supseteq\phi^{-1}[W]\cup\phi^{-1}[X\setminus W]=Y,$$

so

$$\nu(F\cap F')=\nu F+\nu F'-\nu(F\cup F')\le\lambda W+\lambda(X\setminus W)-1=0.$$

Now

$$F\setminus\phi^{-1}[W]\subseteq F\cap\phi^{-1}[X\setminus W]\subseteq F\cap F'$$

is ν-negligible. Because ν is complete, $F\setminus\phi^{-1}[W]\in\mathrm{T}$ and $\phi^{-1}[W]=F\setminus(F\setminus\phi^{-1}[W])$ belongs to T. Moreover,

$$1=\nu F+\nu F'\le\lambda W+\lambda(X\setminus W)=1,$$

so we must have $\nu F=\lambda W$; but this means that $\nu\phi^{-1}[W]=\nu W$. As W is arbitrary, ϕ is inverse-measure-preserving.

254H Corollary Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ and $\langle(Y_i,\mathrm{T}_i,\nu_i)\rangle_{i\in I}$ be two families of probability spaces, with products (X,Λ,λ) and (Y,Λ',λ'). Suppose that for each $i\in I$ we are given an inverse-measure-preserving function $\phi_i:X_i\to Y_i$. Set $\phi(x)=\langle\phi_i(x(i))\rangle_{i\in I}$ for $x\in X$. Then $\phi:X\to Y$ is inverse-measure-preserving.

proof If $C=\prod_{i\in I}C_i$ is a measurable cylinder in Y, then $\phi^{-1}[C]=\prod_{i\in I}\phi_i^{-1}[C_i]$ is a measurable cylinder in X, and

$$\lambda\phi^{-1}[C]=\prod_{i\in I}\mu_i\phi_i^{-1}[C_i]=\prod_{i\in I}\nu_iC_i=\lambda'C.$$

Since λ is a complete probability measure, 254G tells us that ϕ is inverse-measure-preserving.

254I Corresponding to 251T we have the following.

Proposition Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces, λ the product measure on $X=\prod_{i\in I}X_i$, and Λ its domain. Then λ is also the product of the completions $\hat\mu_i$ of the μ_i (212C).

proof Write $\hat\lambda$ for the product of the $\hat\mu_i$, and $\hat\Lambda$ for its domain. (i) The identity map from X_i to itself is inverse-measure-preserving if regarded as a map from $(X_i,\hat\mu_i)$ to (X_i,μ_i), so the identity map on X is inverse-measure-preserving if regarded as a map from $(X,\hat\lambda)$ to (X,λ), by 254H; that is, $\Lambda\subseteq\hat\Lambda$ and $\lambda=\hat\lambda\restriction\Lambda$. (ii) If C is a measurable cylinder for $\langle\hat\mu_i\rangle_{i\in I}$, that is, $C=\prod_{i\in I}C_i$ where $C_i\in\hat\Sigma_i$ for every i and $\{i:C_i\ne X_i\}$ is finite, then for each $i\in I$ we can find a $C_i'\in\Sigma_i$ such that $C_i\subseteq C_i'$ and $\mu_iC_i'=\hat\mu_iC_i$; setting $C'=\prod_{i\in I}C_i'$, we get

$$\lambda^*C\le\lambda C'=\prod_{i\in I}\mu_iC_i'=\prod_{i\in I}\hat\mu_iC_i=\hat\lambda C.$$

By 254G, λW must be defined and equal to $\hat{\lambda}W$ whenever $W \in \hat{\Lambda}$. Putting this together with (i), we see that $\lambda = \hat{\lambda}$.

254J The product measure on $\{0,1\}^I$ **(a)** Perhaps the most important of all examples of infinite product measures is the case in which each factor X_i is just $\{0,1\}$ and each μ_i is the 'fair-coin' probability measure, setting

$$\mu_i\{0\} = \mu_i\{1\} = \frac{1}{2}.$$

In this case, the product $X = \{0,1\}^I$ has a family $\langle E_i \rangle_{i \in I}$ of measurable sets such that, writing λ for the product measure on X,

$$\lambda(\bigcap_{i \in J} E_i) = 2^{-\#(J)} \text{ if } J \subseteq I \text{ is finite.}$$

(Just take $E_i = \{x : x(i) = 1\}$ for each i.) I will call this λ the **usual measure** on $\{0,1\}^I$. Observe that if I is finite then $\lambda\{x\} = 2^{-\#(I)}$ for each $x \in X$ (using 254Fb). On the other hand, if I is infinite, then $\lambda\{x\} = 0$ for every $x \in X$ (because, again using 254Fb, $\lambda^*\{x\} \le 2^{-n}$ for every n).

(b) There is a natural bijection between $\{0,1\}^I$ and $\mathcal{P}I$, matching $x \in \{0,1\}^I$ with $\{i : i \in I,\, x(i) = 1\}$. So we get a standard measure $\tilde{\lambda}$ on $\mathcal{P}I$, which I will call the **usual measure on** $\mathcal{P}I$. Note that for any finite $b \subseteq I$ and any $c \subseteq b$ we have

$$\tilde{\lambda}\{a : a \cap b = c\} = \lambda\{x : x(i) = 1 \text{ for } i \in c,\, x(i) = 0 \text{ for } i \in b \setminus c\} = 2^{-\#(b)}.$$

(c) Of course we can apply 254G to these measures; if (Y, T, ν) is a complete probability space, a function $\phi : Y \to \{0,1\}^I$ is inverse-measure-preserving iff

$$\nu\{y : y \in Y,\, \phi(y) \restriction J = z\} = 2^{-\#(J)}$$

whenever $J \subseteq I$ is finite and $z \in \{0,1\}^J$; this is because the measurable cylinders in $\{0,1\}^I$ are precisely the sets of the form $\{x : x \restriction J = z\}$ where $J \subseteq I$ is finite.

(d) Define addition on X by setting $(x + y)(i) = x(i) +_2 y(i)$ for every $i \in I$, x, $y \in X$, where $0 +_2 0 = 1 +_2 1 = 0$, $0 +_2 1 = 1 +_2 0 = 1$. If $y \in X$, the map $x \mapsto x + y : X \to X$ is inverse-measure-preserving. **P** If $J \subseteq I$ is finite and $z \in \{0,1\}^J$, set $z' = \langle z(j) +_2 y(j) \rangle_{j \in J}$; then

$$\lambda\{x : (x + y) \restriction J = z\} = \lambda\{x : x \restriction J = z'\} = 2^{-\#(J)}.$$

As J is arbitrary, (c) tells us that $x \mapsto x + y$ is inverse-measure-preserving. **Q** Now since

$$(x + y) + y = x + (y + y) = x + 0 = x$$

for every x, the map $x \mapsto x + y : X \to X$ is bijective and equal to its inverse, so it is actually a measure space automorphism of (X, λ).

***(e)** Just because all the factors (X_i, μ_i) are the same, we have another class of automorphisms of (X, λ), corresponding to permutations of I. If $\pi : I \to I$ is any permutation, then we have a corresponding function $x \mapsto x\pi : X \to X$. If $J \subseteq I$ is finite and $z \in \{0,1\}^J$, set $J' = \pi[J]$ and $z' = z\pi^{-1} \in \{0,1\}^{J'}$; then

$$\lambda\{x : (x\pi) \restriction J = z\} = \lambda\{x : x \restriction J' = z'\} = 2^{-\#(J')} = 2^{-\#(J)}.$$

So $x \mapsto x\pi$ is inverse-measure-preserving. This time, its inverse is $x \mapsto x\pi^{-1}$, which is again inverse-measure-preserving; so $x \mapsto x\pi$ is a measure space automorphism.

254K In the case of countably infinite I, we have a very important relationship between the usual product measure of $\{0,1\}^I$ and Lebesgue measure on $[0,1]$.

Proposition Let λ be the usual measure on $X = \{0,1\}^{\mathbb{N}}$, and let μ be Lebesgue measure on $[0,1]$; write Λ for the domain of λ and Σ for the domain of μ.

(i) For $x \in X$ set $\phi(x) = \sum_{i=0}^{\infty} 2^{-i-1} x(i)$. Then

$\phi^{-1}[E] \in \Lambda$ and $\lambda\phi^{-1}[E] = \mu E$ for every $E \in \Sigma$;

$\phi[F] \in \Sigma$ and $\mu\phi[F] = \lambda F$ for every $F \in \Lambda$.

(ii) There is a bijection $\tilde{\phi} : X \to [0,1]$ which is equal to ϕ at all but countably many points, and any such bijection is an isomorphism between (X, Λ, λ) and $([0,1], \Sigma, \mu)$.

proof (a) The first point to observe is that ϕ itself is nearly a bijection. Setting

$$H = \{x : x \in X, \exists\, m \in \mathbb{N},\, x(i) = x(m)\ \forall\, i \ge m\},$$

$$H' = \{2^{-n}k : n \in \mathbb{N},\, k \le 2^n\},$$

then H and H' are countable and $\phi\restriction X \setminus H$ is a bijection between $X \setminus H$ and $[0,1] \setminus H'$. (For $t \in [0,1] \setminus H'$, $\phi^{-1}(t)$ is the binary expansion of t.) Because H and H' are countably infinite, there is a bijection between them; combining this with $\phi\restriction X \setminus H$, we have a bijection between X and $[0,1]$ equal to ϕ except at countably many points. For the rest of this proof, let $\tilde{\phi}$ be any such bijection. Let M be the countable set $\{x : x \in X, \phi(x) \ne \tilde{\phi}(x)\}$, and N the countable set $\phi[M] \cup \tilde{\phi}[M]$; then $\phi[A] \triangle \tilde{\phi}[A] \subseteq N$ for every $A \subseteq X$.

(b) To see that $\lambda\tilde{\phi}^{-1}[E]$ exists and is equal to μE for every $E \in \Sigma$, I consider successively more complex sets E.

(α) If $E = \{t\}$ then $\lambda\tilde{\phi}^{-1}[E] = \lambda\{\tilde{\phi}^{-1}(t)\}$ exists and is zero.

(β) If E is of the form $[2^{-n}k, 2^{-n}(k+1)[$, where $n \in \mathbb{N}$ and $0 \le k < 2^n$, then $\phi^{-1}[E]$ differs by at most two points from a set of the form $\{x : x(i) = z(i)\ \forall\, i < n\}$, so $\tilde{\phi}^{-1}[E]$ differs from this by a countable set, and

$$\lambda\tilde{\phi}^{-1}[E] = 2^{-n} = \mu E.$$

(γ) If E is of the form $[2^{-n}k, 2^{-n}l[$, where $n \in \mathbb{N}$ and $0 \le k < l \le 2^n$, then

$$E = \bigcup_{k \le i < l} [2^{-n}i, 2^{-n}(i+1)[,$$

so

$$\lambda\tilde{\phi}^{-1}[E] = 2^{-n}(l-k) = \mu E.$$

(δ) If E is of the form $[t, u[$, where $0 \le t < u \le 1$, then for each $n \in \mathbb{N}$ set $k_n = \lfloor 2^n t \rfloor$, the integer part of $2^n t$, $l_n = \lfloor 2^n u \rfloor$ and $E_n = [2^{-n}(k_n+1), 2^{-n}l_n[$; then $\langle E_n \rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence and $\bigcup_{n\in\mathbb{N}} E_n$ is $]t, u[$. So (using (α))

$$\lambda\tilde{\phi}^{-1}[E] = \lambda\tilde{\phi}^{-1}[\bigcup_{n\in\mathbb{N}} E_n] = \lim_{n\to\infty} \lambda\tilde{\phi}^{-1}[E_n]$$
$$= \lim_{n\to\infty} \mu E_n = \mu E.$$

(ϵ) If $E \in \Sigma$, then for any $\epsilon > 0$ there is a sequence $\langle I_n \rangle_{n\in\mathbb{N}}$ of half-open subintervals of $[0,1[$ such that $E \setminus \{1\} \subseteq \bigcup_{n\in\mathbb{N}} I_n$ and $\sum_{n=0}^{\infty} \mu I_n \le \mu E + \epsilon$; now $\tilde{\phi}^{-1}[E] \subseteq \{\tilde{\phi}^{-1}(1)\} \cup \bigcup_{n\in\mathbb{N}} \phi^{-1}[I_n]$, so

$$\lambda^*\tilde{\phi}^{-1}[E] \le \lambda(\bigcup_{n\in\mathbb{N}} \tilde{\phi}^{-1}[I_n]) \le \sum_{n=0}^{\infty} \lambda\tilde{\phi}^{-1}[I_n] = \sum_{n=0}^{\infty} \mu I_n \le \mu E + \epsilon.$$

As ϵ is arbitrary, $\lambda^*\tilde{\phi}^{-1}[E] \le \mu E$, and there is a $V \in \Lambda$ such that $\tilde{\phi}^{-1}[E] \subseteq V$ and $\lambda V \le \mu E$.

(ζ) Similarly, there is a $V' \in \Lambda$ such that $V' \supseteq \tilde{\phi}^{-1}[[0,1] \setminus E]$ and $\lambda V' \le \mu([0,1] \setminus E)$. Now $V \cup V' = X$, so

$$\lambda(V \cap V') = \lambda V + \lambda V' - \lambda(V \cup V') \le \mu E + (1 - \mu E) - 1 = 0$$

and

$$\tilde{\phi}^{-1}[E] = (X \setminus V') \cup (V \cap V' \cap \tilde{\phi}^{-1}[E])$$

belongs to Λ, with

$$\lambda\tilde{\phi}^{-1}[E] \le \lambda V \le \mu E;$$

at the same time,

$$1 - \lambda\tilde{\phi}^{-1}[E] \le \lambda V' \le 1 - \mu E$$

so $\lambda\tilde{\phi}^{-1}[E] = \mu E$.

(c) Now suppose that $C \subseteq X$ is a measurable cylinder of the special form $\{x : x(0) = \epsilon_0, \dots, x(n) = \epsilon_n\}$ for some $\epsilon_0, \dots, \epsilon_n \in \{0, 1\}$. Then $\phi[C] = [t, t + 2^{-n-1}]$ where $t = \sum_{i=0}^{n} 2^{-i-1}\epsilon_i$, so that $\mu\phi[C] = \lambda C$. Since $\tilde{\phi}[C]\triangle\phi[C] \subseteq N$ is countable, $\mu\tilde{\phi}[C] = \lambda C$.

If $C \subseteq X$ is any measurable cylinder, then it is of the form $\{x : x{\restriction}J = z\}$ for some finite $J \subseteq \mathbb{N}$; taking n so large that $J \subseteq \{0, \dots, n\}$, C is expressible as a disjoint union of $2^{n+1-\#(J)}$ sets of the form just considered, being just those in which $\epsilon_i = z(i)$ for $i \in J$. Summing their measures, we again get $\mu\tilde{\phi}[C] = \lambda C$. Now 254G tells us that $\tilde{\phi}^{-1} : [0, 1] \to X$ is inverse-measure-preserving, that is, $\tilde{\phi}[W]$ is Lebesgue measurable, with measure λW, for every $W \in \Lambda$.

Putting this together with (b), $\tilde{\phi}$ must be an isomorphism between (X, Λ, λ) and $([0, 1], \Sigma, \mu)$, as claimed in (ii) of the proposition.

(d) As for (i), if $E \in \Sigma$ then $\phi^{-1}[E]\triangle\tilde{\phi}^{-1}[E] \subseteq M$ is countable, so $\lambda\phi^{-1}[E] = \lambda\tilde{\phi}^{-1}[E] = \mu E$. While if $W \in \Lambda$, $\phi[F]\triangle\tilde{\phi}[W] \subseteq N$ is countable, so $\mu\phi[W] = \mu\tilde{\phi}[W] = \lambda W$.

254L Subspaces Just as in 251Q, we can consider the product of subspace measures. There is a simplification in the form of the result because in the present context we are restricted to probability measures.

Theorem Let $\langle(X_i, \Sigma_i, \mu_i)\rangle_{i\in I}$ be a family of probability spaces, and (X, Λ, λ) their product.

(a) For each $i \in I$, let $A_i \subseteq X_i$ be a set of full outer measure, and write $\tilde{\mu}_i$ for the subspace measure on A_i (214B). Let $\tilde{\lambda}$ be the product measure on $A = \prod_{i\in I} A_i$. Then $\tilde{\lambda}$ is the subspace measure on A induced by λ.

(b) $\lambda^*(\prod_{i\in I} A_i) = \prod_{i\in I} \mu_i^* A_i$ whenever $A_i \subseteq X_i$ for every i.

proof (a) Write λ_A for the subspace measure on A defined from λ, and Λ_A for its domain; write $\tilde{\Lambda}$ for the domain of $\tilde{\lambda}$.

(i) Let $\phi : A \to X$ be the identity map. If $C \subseteq X$ is a measurable cylinder, say $C = \prod_{i\in I} C_i$ where $C_i \in \Sigma_i$ for each i, then $\phi^{-1}[C] = \prod_{i\in I}(C_i \cap A_i)$ is a measurable cylinder in A, and

$$\tilde{\lambda}\phi^{-1}[C] = \prod_{i\in I} \tilde{\mu}_i(C_i \cap A_i) \le \prod_{i\in I} \mu_i C_i = \mu C.$$

By 254G, ϕ is inverse-measure-preserving, that is, $\tilde{\lambda}(A \cap W) = \lambda W$ for every $W \in \Lambda$. But this means that $\tilde{\lambda}V$ is defined and equal to $\lambda_A V = \lambda^* V$ for every $V \in \Lambda_A$, since for any such V there is a $W \in \Lambda$ such that $V = A \cap W$ and $\lambda W = \lambda_A V$. In particular, $\lambda_A A = 1$.

(ii) Now regard ϕ as a function from the measure space $(A, \Lambda_A, \lambda_A)$ to $(A, \tilde{\Lambda}, \tilde{\lambda})$. If D is a measurable cylinder in A, we can express it as $\prod_{i\in I} D_i$ where every D_i belongs to the domain of $\tilde{\mu}_i$ and $D_i = A_i$ for all but finitely many i. Now for each i we can find $C_i \in \Sigma_i$ such that $D_i = C_i \cap A_i$ and $\mu C_i = \tilde{\mu}_i D_i$, and we can suppose that $C_i = X_i$ whenever $D_i = A_i$. In this case $C = \prod_{i\in I} C_i \in \Lambda$ and

$$\lambda C = \prod_{i\in I} \mu_i C_i = \prod_{i\in I} \tilde{\mu}_i D_i = \tilde{\lambda} D.$$

Accordingly

$$\lambda_A\phi^{-1}[D] = \lambda_A(A \cap C) \le \lambda C = \tilde{\lambda} D.$$

By 254G again, ϕ is inverse-measure-preserving in this manifestation, that is, $\lambda_A V$ is defined and equal to $\tilde{\lambda}V$ for every $V \in \tilde{\Lambda}$. Putting this together with (i), we have $\lambda_A = \tilde{\lambda}$, as claimed.

(b) For each $i \in I$, choose a set $E_i \in \Sigma_i$ such that $A_i \subseteq E_i$ and $\mu_i E_i = \mu_i^* A_i$; do this in such a way that $E_i = X_i$ whenever $\mu_i^* A_i = 1$. Set $B_i = A_i \cup (X_i \setminus E_i)$, so that $\mu_i^* B_i = 1$ for each i (if $F \in \Sigma_i$ and $F \supseteq B_i$ then $F \cap E_i \supseteq A_i$, so

$$\mu_i F = \mu_i(F \cap E_i) + \mu_i(F \setminus E_i) = \mu_i E_i + \mu_i(X_i \setminus E_i) = 1.)$$

By (a), we can identify the subspace measure λ_B on $B = \prod_{i\in I} B_i$ with the product of the subspace measures $\tilde{\mu}_i$ on B_i. In particular, $\lambda^* B = \lambda_B B = 1$. Now $A_i = B_i \cap E_i$ so (writing $A = \prod_{i\in I} A_i$), $A = B \cap \prod_{i\in I} E_i$.

If $\prod_{i\in I}\mu_i^*A_i=0$, then for every $\epsilon>0$ there is a finite $J\subseteq I$ such that $\prod_{i\in J}\mu_i^*A_i\le\epsilon$; consequently (using 254Fb)

$$\lambda^*A\le\lambda\{x:x(i)\in E_i \text{ for every } i\in J\}=\textstyle\prod_{i\in J}\mu_iE_i\le\epsilon.$$

As ϵ is arbitrary, $\lambda^*A=0$. If $\prod_{i\in I}\mu_i^*A_i>0$, then for every $n\in\mathbb{N}$ the set $\{i:\mu^*A_i\le 1-2^{-n}\}$ must be finite, so

$$J=\{i:\mu^*A_i<1\}=\{i:E_i\ne X_i\}$$

is countable. By 254Fb again, applied to $\langle E_i\cap B_i\rangle_{i\in I}$ in the product $\prod_{i\in I}B_i$,

$$\begin{aligned}\lambda^*(\prod_{i\in I}A_i)&=\lambda_B(\prod_{i\in I}A_i)=\lambda_B\{x:x\in B,\,x(i)\in E_i\cap B_i \text{ for every } i\in J\}\\&=\prod_{i\in J}\tilde\mu_i(E_i\cap B_i)=\prod_{i\in I}\mu_i^*A_i,\end{aligned}$$

as required.

254M I now turn to the basic results which make it possible to use these product measures effectively. First, I offer a vocabulary for dealing with subproducts. Let $\langle X_i\rangle_{i\in I}$ be a family of sets, with product X.

(a) For $J\subseteq I$, write X_J for $\prod_{i\in J}X_i$. We have a canonical bijection $x\mapsto(x\restriction J,x\restriction I\setminus J):X\to X_I\times X_{I\setminus J}$. Associated with this we have the map $x\mapsto\pi_J(x)=x\restriction J:X\to X_J$. Now I will say that a set $W\subseteq X$ is **determined by coordinates in** J if there is a $V\subseteq X_J$ such that $W=\pi_J^{-1}[V]$; that is, W corresponds to $V\times X_{I\setminus J}\subseteq X_J\times X_{I\setminus J}$.

It is easy to see that

$$\begin{aligned}&W \text{ is determined by coordinates in } J\\&\iff x'\in W \text{ whenever } x\in W,\ x'\in X \text{ and } x'\restriction J=x\restriction J\\&\iff W=\pi_J^{-1}[\pi_J[W]].\end{aligned}$$

It follows that if W is determined by coordinates in J, and $J\subseteq K\subseteq I$, W is also determined by coordinates in K. The family $\mathcal{W}_J$ of subsets of X determined by coordinates in J is closed under complementation and arbitrary unions and intersections. **P** If $W\in\mathcal{W}_J$, then

$$X\setminus W=X\setminus\pi_J^{-1}[\pi_J[W]]=\pi_J^{-1}[X_J\setminus\pi_J[W]]\in\mathcal{W}_J.$$

If $\mathcal{V}\subseteq\mathcal{W}_J$, then

$$\textstyle\bigcup\mathcal{V}=\bigcup_{V\in\mathcal{V}}\pi_J^{-1}[\pi_J[V]]=\pi_J^{-1}[\bigcup_{V\in\mathcal{V}}\pi_J[V]]\in\mathcal{W}_J.\ \mathbf{Q}$$

(b) It follows that

$$\mathcal{W}=\bigcup\{\mathcal{W}_J:J\subseteq I \text{ is countable}\},$$

the family of subsets of X determined by coordinates in some countable set, is a σ-algebra of subsets of X. **P** (i) X and $\emptyset$ are determined by coordinates in $\emptyset$ (recall that $X_\emptyset$ is a singleton, and that $X=\pi_\emptyset^{-1}[X_\emptyset]$, $\emptyset=\pi_\emptyset^{-1}[\emptyset]$). (ii) If $W\in\mathcal{W}$, there is a countable $J\subseteq I$ such that $W\in\mathcal{W}_J$; now

$$X\setminus W=\pi_J^{-1}[X_J\setminus\pi_J[W]]\in\mathcal{W}_J\subseteq\mathcal{W}.$$

(iii) If $\langle W_n\rangle_{n\in\mathbb{N}}$ is a sequence in $\mathcal{W}$, then for each $n\in\mathbb{N}$ there is a countable $J_n\subseteq I$ such that $W\in\mathcal{W}_{J_n}$. Now $J=\bigcup_{n\in\mathbb{N}}J_n$ is a countable subset of I, and every W_n belongs to $\mathcal{W}_J$, so

$$\textstyle\bigcup_{n\in\mathbb{N}}W_n\in\mathcal{W}_J\subseteq\mathcal{W}.\ \mathbf{Q}$$

(c) If $i\in I$ and $E\subseteq X_i$ then $\{x:x\in X,\ x(i)\in E\}$ is determined by the single coordinate i, so surely belongs to $\mathcal{W}$; accordingly $\mathcal{W}$ must include $\widehat{\bigotimes}_{i\in I}\mathcal{P}X_i$. *A fortiori*, if Σ_i is a σ-algebra of subsets of X_i for each i, $\mathcal{W}\supseteq\widehat{\bigotimes}_{i\in I}\Sigma_i$; that is, every member of $\widehat{\bigotimes}_{i\in I}\Sigma_i$ is determined by coordinates in some countable set.

254N Theorem Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces and $\langle K_j\rangle_{j\in J}$ a partition of I. For each $j\in J$ let λ_j be the product measure on $Z_j=\prod_{i\in K_j}X_i$, and write λ for the product measure on $X=\prod_{i\in I}X_i$. Then the natural bijection

$$x\mapsto\phi(x)=\langle x\restriction K_j\rangle_{j\in J}:X\to\prod_{j\in J}Z_j$$

identifies λ with the product of the family $\langle\lambda_j\rangle_{j\in J}$.

In particular, if $K\subseteq I$ is any set, then λ can be identified with the c.l.d. product of the product measures on $\prod_{i\in K}X_i$ and $\prod_{i\in I\setminus K}X_i$.

proof (Compare 251N.) Write $Z=\prod_{j\in J}Z_j$ and $\tilde\lambda$ for the product measure on Z; let Λ, $\tilde\Lambda$ be the domains of λ and $\tilde\lambda$.

(a) Let $C\subseteq Z$ be a measurable cylinder. Then $\lambda^*\phi^{-1}[C]\le\tilde\lambda C$. **P** Express C as $\prod_{j\in J}C_j$ where $C_j\subseteq Z_j$ belongs to the domain Λ_j of λ_j for each j. Set $L=\{j:C_j\ne Z_j\}$, so that L is finite. Let $\epsilon>0$. For each $j\in L$ let $\langle C_{jn}\rangle_{n\in\mathbb{N}}$ be a sequence of measurable cylinders in $Z_j=\prod_{i\in K_j}X_i$ such that $C_j\subseteq\bigcup_{n\in\mathbb{N}}C_{jn}$ and $\sum_{n=0}^\infty\lambda_jC_{jn}\le\lambda C_j+\epsilon$. Express each C_{jn} as $\prod_{i\in K_j}C_{jni}$ where $C_{jni}\in\Sigma_i$ for $i\in K_j$ (and $\{i:C_{jni}\ne X_i\}$ is finite).

For $f\in\mathbb{N}^L$, set

$$D_f=\{x:x\in X,\,x(i)\in C_{j,f(j),i}\text{ whenever }j\in L,\,i\in K_j\}.$$

Because $\bigcup_{j\in L}\{i:C_{j,f(j),i}\ne X_i\}$ is finite, D_f is a measurable cylinder in X, and

$$\lambda D_f=\prod_{j\in L}\prod_{i\in K_j}\mu_iC_{j,f(j),i}=\prod_{j\in L}\lambda_jC_{j,f(j)}.$$

Also

$$\bigcup\{D_f:f\in\mathbb{N}^L\}\supseteq\phi^{-1}[C]$$

because if $\phi(x)\in C$ then $\phi(x)(j)\in C_j$ for each $j\in L$, so there must be an $f\in\mathbb{N}^L$ such that $\phi(x)(j)\in C_{j,f(j)}$ for every $j\in L$. But (because $\mathbb{N}^L$ is countable) this means that

$$\lambda^*\phi^{-1}[C]\le\sum_{f\in\mathbb{N}^L}\lambda D_f=\sum_{f\in\mathbb{N}^L}\prod_{j\in L}\lambda_jC_{j,f(j)}$$
$$=\prod_{j\in L}\sum_{n=0}^\infty\lambda_jC_{jn}\le\prod_{j\in L}(\lambda_jC_j+\epsilon).$$

As ϵ is arbitrary,

$$\lambda^*\phi^{-1}[C]\le\prod_{j\in L}\lambda_jC_j=\tilde\lambda C.\ \mathbf{Q}$$

By 254G, it follows that $\lambda\phi^{-1}[W]$ is defined, and equal to $\tilde\lambda W$, whenever $W\in\tilde\Lambda$.

(b) Next, $\tilde\lambda\phi[D]=\lambda D$ for every measurable cylinder $D\subseteq X$. **P** This is easy. Express D as $\prod_{i\in I}D_i$ where $D_i\in\Sigma_i$ for every $i\in I$ and $\{i:D_i\ne\Sigma_i\}$ is finite. Then $\phi[D]=\prod_{j\in J}\tilde D_j$, where $\tilde D_j=\prod_{i\in K_j}D_i$ is a measurable cylinder for each $j\in J$. Because $\{j:\tilde D_j\ne Z_j\}$ must also be finite (in fact, it cannot have more members than the finite set $\{i:D_i\ne X_i\}$), $\prod_{j\in J}\tilde D_j$ is itself a measurable cylinder in Z, and

$$\tilde\lambda\phi[D]=\prod_{j\in J}\lambda_j\tilde D_j=\prod_{j\in J}\prod_{i\in K_j}\mu D_i=\lambda D.\ \mathbf{Q}$$

Applying 254G to $\phi^{-1}:Z\to X$, it follows that $\tilde\lambda\phi[W]$ is defined, and equal to λW, for every $W\in\Lambda$. But together with (a) this means that for any $W\subseteq X$,

if $W\in\Lambda$ then $\phi[W]\in\tilde\Lambda$ and $\tilde\lambda\phi[W]=\lambda W$,

if $\phi[W]\in\tilde\Lambda$ then $W\in\Lambda$ and $\lambda W=\tilde\lambda\phi[W]$.

And of course this is just what is meant by saying that ϕ is an isomorphism between (X,Λ,λ) and $(Z,\tilde\Lambda,\tilde\lambda)$.

254O Proposition Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces. For each $J\subseteq I$ let λ_J be the product probability measure on $X_J=\prod_{i\in J}X_i$, and Λ_J its domain; write $X=X_I$, $\lambda=\lambda_I$ and $\Lambda=\Lambda_I$. For $x\in X$ and $J\subseteq I$ set $\pi_J(x)=x\restriction J\in X_J$.

(a) For every $J \subseteq I$, λ_J is the image measure $\lambda\pi_J^{-1}$ (234D); in particular, $\pi_J : X \to X_J$ is inverse-measure-preserving for λ and λ_J.

(b) If $J \subseteq I$ and $W \in \Lambda$ is determined by coordinates in J (254M), then $\lambda_J\pi_J[W]$ is defined and equal to λW. Consequently there are W_1, W_2 belonging to the σ-algebra of subsets of X generated by

$$\{\{x : x(i) \in E\} : i \in J,\, E \in \Sigma_i\}$$

such that $W_1 \subseteq W \subseteq W_2$ and $\lambda(W_2 \setminus W_1) = 0$.

(c) For every $W \in \Lambda$, we can find a countable set J and W_1, $W_2 \in \Lambda$, both determined by coordinates in J, such that $W_1 \subseteq W \subseteq W_2$ and $\lambda(W_2 \setminus W_1) = 0$.

(d) For every $W \in \Lambda$, there is a countable set $J \subseteq I$ such that $\pi_J[W] \in \Lambda_J$ and $\lambda_J\pi_J[W] = \lambda W$; so that $W' = \pi_J^{-1}[\pi_J[W]]$ belongs to Λ, and $\lambda(W' \setminus W) = 0$.

proof (a)(i) By 254N, we can identify λ with the product of λ_J and $\lambda_{I\setminus J}$ on $X_J \times X_{I\setminus J}$. Now $\pi_J^{-1}[E] \subseteq X$ corresponds to $E \times X_{I\setminus J} \subseteq X_J \times X_{I\setminus J}$, so

$$\lambda(\pi^{-1}[E]) = \lambda_J E \cdot \lambda_{I\setminus J} X_{I\setminus J} = \lambda_J E,$$

by 251E or 251Ia, whenever $E \in \Lambda_J$. This shows that π_J is inverse-measure-preserving.

(ii) To see that λ_J is actually the image measure, suppose that $E \subseteq X_J$ is such that $\pi_J^{-1}[E] \in \Lambda$. Identifying $\pi_J^{-1}[E]$ with $E \times X_{I\setminus J}$, as before, we are supposing that $E \times X_{I\setminus J}$ is measured by the product measure on $X_J \times X_{I\setminus J}$. But this means that for $\lambda_{I\setminus J}$-almost every $z \in X_{I\setminus J}$, $E_z = \{y : (y,z) \in E \times X_{I\setminus J}\}$ belongs to Λ_J (252D(ii), because λ_J is complete). Since $E_z = E$ for every z, E itself belongs to Λ_J, as claimed.

(b) If $W \in \Lambda$ is determined by coordinates in J, set $H = \pi_J[W]$; then $\pi_J^{-1}[H] = W$, so $H \in \Lambda_J$ by (a) just above. By 254Ff, there are H_1, $H_2 \in \widehat{\bigotimes}_{i\in J}\Sigma_i$ such that $H_1 \subseteq H \subseteq H_2$ and $\lambda_J(H_2 \setminus H_1) = 0$.

Let T_J be the σ-algebra of subsets of X generated by sets of the form $\{x : x(i) \in E\}$ where $i \in J$ and $E \in \Sigma_J$. Consider $\mathrm{T}'_J = \{G : G \subseteq X_J,\, \pi_J^{-1}[G] \in \mathrm{T}_J\}$. This is a σ-algebra of subsets of X_J, and it contains $\{y : y \in X_J,\, y(i) \in E\}$ whenever $i \in J$, $E \in \Sigma_J$ (because

$$\pi_J^{-1}[\{y : y \in X_J,\, y(i) \in E\}] = \{x : x \in X,\, x(i) \in E\}$$

whenever $i \in J$, $E \subseteq X_i$). So T'_J must include $\widehat{\bigotimes}_{i\in J}\Sigma_i$. In particular, H_1 and H_2 both belong to T'_J, that is, $W_k = \pi_J^{-1}[H_k]$ belongs to T_J for both k. Of course $W_1 \subseteq W \subseteq W_2$, because $H_1 \subseteq H \subseteq H_2$, and

$$\lambda(W_2 \setminus W_1) = \lambda_J(H_2 \setminus H_1) = 0,$$

as required.

(c) Now take any $W \in \Lambda$. By 254Ff, there are W_1 and $W_2 \in \widehat{\bigotimes}_{i\in I}\Sigma_i$ such that $W_1 \subseteq W \subseteq W_2$ and $\lambda(W_2 \setminus W_1) = 0$. By 254Mc, there are countable sets J_1, $J_2 \subseteq I$ such that, for each k, W_k is determined by coordinates in J_k. Setting $J = J_1 \cup J_2$, J is a countable subset of I and both W_1 and W_2 are determined by coordinates in J.

(d) Continuing the argument from (c), $\pi_J[W_1]$, $\pi_J[W_2] \in \Lambda_J$, by (b), and $\lambda_J(\pi_J[W_2] \setminus \pi_J[W_1]) = 0$. Since $\pi_J[W_1] \subseteq \pi_J[W] \subseteq \pi_J[W_2]$, it follows that $\pi_J[W] \in \Lambda_J$, with $\lambda_J\pi_J[W] = \lambda_J\pi_J[W_2]$; so that, setting $W' = \pi_J^{-1}[\pi_J[W]]$, $W' \in \Lambda$, and

$$\lambda W' = \lambda_J\pi_J[W] = \lambda_J\pi_J[W_2] = \lambda\pi_J^{-1}[\pi_J[W_2]] = \lambda W_2 = \lambda W.$$

254P Proposition Let $\langle(X_i, \Sigma_i, \mu_i)\rangle_{i\in I}$ be a family of probability spaces, and for each $J \subseteq I$ let λ_J be the product probability measure on $X_J = \prod_{i\in J} X_i$, and Λ_J its domain; write $X = X_I$, $\Lambda = \Lambda_I$ and $\lambda = \lambda_I$. For $x \in X$ and $J \subseteq I$ set $\pi_J(x) = x\restriction J \in X_J$.

(a) If $J \subseteq I$ and g is a real-valued function defined on a subset of X_J, then g is Λ_J-measurable iff $g\pi_J$ is Λ-measurable.

(b) Whenever f is a Λ-measurable real-valued function defined on a λ-conegligible subset of X, we can find a countable set $J \subseteq I$ and a Λ_J-measurable function g defined on a λ_J-conegligible subset of X_J such that f extends $g\pi_J$.

proof (a)(i) If g is Λ_J-measurable and $a \in \mathbb{R}$, there is an $H \in \Lambda_J$ such that $\{y : y \in \operatorname{dom} g,\ g(y) \geq a\} = H \cap \operatorname{dom} g$. Now $\pi_J^{-1}[H] \in \Lambda$, by 254Oa, and $\{x : x \in \operatorname{dom} g\pi_J,\ g\pi_J(x) \geq a\} = \pi_J^{-1}[H] \cap \operatorname{dom} g\pi_J$. So $g\pi_J$ is Λ-measurable.

(ii) If $g\pi_J$ is Λ-measurable and $a \in \mathbb{R}$, then there is a $W \in \Lambda$ such that $\{x : g\pi_J(x) \geq a\} = W \cap \operatorname{dom} g\pi_J$. As in the proof of 254Oa, we may identify λ with the product of λ_J and $\lambda_{I\setminus J}$, and 252D(ii) tells us that, if we identify W with the corresponding subset of $X_J \times X_{I\setminus J}$, there is at least one $z \in X_{I\setminus J}$ such that $W_z = \{y : y \in X_J,\ (y,z) \in W\}$ belongs to Λ_J. But since (on this convention) $g\pi_J(y,z) = g(y)$ for every $y \in X_J$, we see that $\{y : y \in \operatorname{dom} g,\ g(y) \geq a\} = W_z \cap \operatorname{dom} g$. As a is arbitrary, g is Λ_J-measurable.

(b) For rational numbers q, set $W_q = \{x : x \in \operatorname{dom} f,\ f(x) \geq q\}$. By 254Oc we can find for each q a countable set $J_q \subseteq I$ and sets W'_q, W''_q, both determined by coordinates in J_q, such that $W'_q \subseteq W_q \subseteq W''_q$ and $\lambda(W''_q \setminus W'_q) = 0$. Set $J = \bigcup_{q\in\mathbb{Q}} J_q$, $V = X \setminus \bigcup_{q\in\mathbb{Q}}(W''_q \setminus W'_q)$; then J is a countable subset of I and V is a conegligible subset of X; moreover, V is determined by coordinates in J because all the W'_q, W''_q are.

For every $q \in \mathbb{Q}$, $W_q \cap V = W'_q \cap V$, because $V \cap (W_q \setminus W'_q) \subseteq V \cap (W''_q \setminus W'_q) = \emptyset$; so $W_q \cap V$ is determined by coordinates in J. Consequently $V \cap \operatorname{dom} f = \bigcup_{q\in\mathbb{Q}} V \cap W_q$ also is determined by coordinates in J. Also

$$\{x : x \in V \cap \operatorname{dom} f,\ f(x) \geq a\} = \bigcap_{q\leq a} V \cap W_q$$

is determined by coordinates in J. What this means is that if x, $x' \in V$ and $\pi_J x = \pi_J x'$, then $x \in \operatorname{dom} f$ iff $x' \in \operatorname{dom} f$ and in this case $f(x) = f(x')$. Setting $H = \pi_J[V \cap \operatorname{dom} f]$, we have $\pi_J^{-1}[H] = V \cap \operatorname{dom} f$ a conegligible subset of X, so (because $\lambda_J = \lambda\pi_J^{-1}$) H is conegligible in X_J. Also, for $y \in H$, $f(x) = f(x')$ whenever $\pi_J x = \pi_J x' = y$, so there is a function $g : H \to \mathbb{R}$ defined by saying that $g\pi_J(x) = f(x)$ whenever $x \in V \cap \operatorname{dom} f$. Thus g is defined almost everywhere in X_J and f extends $g\pi_J$. Finally, for any $a \in \mathbb{R}$,

$$\pi_J^{-1}[\{y : g(y) \geq a\}] = \{x : x \in V \cap \operatorname{dom} f,\ f(x) \geq a\} \in \Lambda;$$

by 254Oa, $\{y : g(y) \geq a\} \in \Lambda_J$; as a is arbitrary, g is measurable.

254Q Proposition Let $\langle(X_i, \Sigma_i, \mu_i)\rangle_{i\in I}$ be a family of probability spaces, and for each $J \subseteq I$ let λ_J be the product probability measure on $X_J = \prod_{i\in J} X_i$; write $X = X_I$, $\lambda = \lambda_I$. For $x \in X$, $J \subseteq I$ set $\pi_J(x) = x\restriction J \in X_J$.

(a) Let $\mathcal{S}$ be the linear subspace of $\mathbb{R}^X$ spanned by $\{\chi C : C \subseteq X$ is a measurable cylinder$\}$. Then for every λ-integrable real-valued function f and every $\epsilon > 0$ there is a $g \in \mathcal{S}$ such that $\int |f - g| d\lambda \leq \epsilon$.

(b) Whenever $J \subseteq I$ and g is a real-valued function defined on a subset of X_J, then $\int g\, d\lambda_J = \int g\pi_J d\lambda$ if either integral is defined in $[-\infty, \infty]$.

(c) Whenever f is a λ-integrable real-valued function, we can find a countable set $J \subseteq X$ and a λ_J-integrable function g such that f extends $g\pi_J$.

proof (a)(i) Write $\overline{\mathcal{S}}$ for the set of functions f satisfying the assertion, that is, such that for every $\epsilon > 0$ there is a $g \in \mathcal{S}$ such that $\int |f - g| \leq \epsilon$. Then $f_1 + f_2$ and $cf_1 \in \overline{\mathcal{S}}$ whenever f_1, $f_2 \in \overline{\mathcal{S}}$. **P** Given $\epsilon > 0$ there are g_1, $g_2 \in \mathcal{S}$ such that $\int |f_1 - g_1| \leq \frac{\epsilon}{2+|c|}$, $\int |f_2 - g_2| \leq \frac{\epsilon}{2}$; now $g_1 + g_2$, $cg_1 \in \mathcal{S}$ and $\int |(f_1 + f_2) - (g_1 + g_2)| \leq \epsilon$, $\int |cf_1 - cg_1| \leq \epsilon$. **Q** Also, of course, $f \in \overline{\mathcal{S}}$ whenever $f_0 \in \overline{\mathcal{S}}$ and $f =_{\text{a.e.}} f_0$.

(ii) Write $\mathcal{W}$ for $\{W : W \subseteq X,\ \chi W \in \overline{\mathcal{S}}\}$, and $\mathcal{C}$ for the family of measurable cylinders in X. Then it is plain from the definition in 254A that $C \cap C' \in \mathcal{C}$ for all C, $C' \in \mathcal{C}$, and of course $C \in \mathcal{W}$ for every $C \in \mathcal{C}$, because $\chi C \in \mathcal{S}$. Next, $W \setminus V \in \mathcal{W}$ whenever W, $V \in \mathcal{W}$ and $V \subseteq W$, because then $\chi(W \setminus V) = \chi W - \chi V$. Thirdly, $\bigcup_{n\in\mathbb{N}} W_n \in \mathcal{W}$ for every non-decreasing sequence $\langle W_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{W}$. **P** Set $W = \bigcup_{n\in\mathbb{N}} W_n$. Given $\epsilon > 0$, there is an $n \in \mathbb{N}$ such that $\lambda(W \setminus W_n) \leq \frac{\epsilon}{2}$. Now there is a $g \in \mathcal{S}$ such that $\int |\chi W_n - g| \leq \frac{\epsilon}{2}$, so that $\int |\chi W - g| \leq \epsilon$. **Q** Thus $\mathcal{W}$ is a Dynkin class of subsets of X.

By the Monotone Class Theorem (136B), $\mathcal{W}$ must include the σ-algebra of subsets of X generated by $\mathcal{C}$, which is $\widehat{\bigotimes}_{i\in I}\Sigma_i$. But this means that $\mathcal{W}$ contains every measurable subset of X, since by 254Ff any measurable set differs by a negligible set from some member of $\widehat{\bigotimes}_{i\in I}\Sigma_i$.

(iii) Thus $\overline{\mathcal{S}}$ contains the indicator function of any measurable subset of X. Because it is closed under addition and scalar multiplication, it contains all simple functions. But this means that it must contain all integrable functions. **P** If f is a real-valued function which is integrable over X, and $\epsilon > 0$, there is a simple function $h : X \to \mathbb{R}$ such that $\int |f - h| \le \frac{\epsilon}{2}$ (242M), and now there is a $g \in \mathcal{S}$ such that $\int |h - g| \le \frac{\epsilon}{2}$, so that $\int |f - g| \le \epsilon$. **Q**

This proves part (a) of the proposition.

(b) Put 254Oa and 235J together.

(c) By 254Pb, there are a countable $J \subseteq I$ and a real-valued function g defined on a conegligible subset of X_J such that f extends $g\pi_J$. Now $\operatorname{dom}(g\pi_J) = \pi_J^{-1}[\operatorname{dom} g]$ is conegligible, so $f =_{\text{a.e.}} g\pi_J$ and $g\pi_J$ is λ-integrable. By (b), g is λ_J-integrable.

254R Conditional expectations again Putting the ideas of 253H together with the work above, we obtain some results which are important not only for their direct applications but for the light they throw on the structures here.

Theorem Let $\langle(X_i, \Sigma_i, \mu_i)\rangle_{i\in I}$ be a family of probability spaces with product (X, Λ, λ). For $J \subseteq I$ let $\Lambda_J \subseteq \Lambda$ be the σ-subalgebra of sets determined by coordinates in J (254Ma). Then we may regard $L^0(\lambda\restriction\Lambda_J)$ as a subspace of $L^0(\lambda)$ (242Jh). Let $P_J : L^1(\lambda) \to L^1(\lambda\restriction\Lambda_J) \subseteq L^1(\lambda)$ be the corresponding conditional expectation operator (242Jd). Then

(a) for any J, $K \subseteq I$, $P_{K\cap J} = P_K P_J$;

(b) for any $u \in L^1(\lambda)$, there is a countable set $J^* \subseteq I$ such that $P_J u = u$ iff $J \supseteq J^*$;

(c) for any $u \in L^0(\lambda)$, there is a unique smallest set $J^* \subseteq I$ such that $u \in L^0(\lambda\restriction\Lambda_{J^*})$, and this J^* is countable;

(d) for any $W \in \Lambda$ there is a unique smallest set $J^* \subseteq I$ such that $W\triangle W'$ is negligible for some $W' \in \Lambda_{J^*}$, and this J^* is countable;

(e) for any Λ-measurable real-valued function $f : X \to \mathbb{R}$ there is a unique smallest set $J^* \subseteq I$ such that f is equal almost everywhere to a Λ_J^*-measurable function, and this J^* is countable.

proof For $J \subseteq I$, write $X_J = \prod_{i\in J} X_i$, let λ_J be the product measure on X_J, and set $\pi_J(x) = x\restriction J$ for $x \in X$. Write L^0_J for $L^0(\lambda\restriction\Lambda_J)$, regarded as a subset of $L^0 = L^0_I$, and L^1_J for $L^1(\lambda\restriction\Lambda_J) = L^1(\lambda) \cap L^0_J$, as in 242Jb; thus L^1_J is the set of values of the projection P_J.

(a)(i) Let $C \subseteq X$ be a measurable cylinder, expressed as $\prod_{i\in I} C_i$ where $C_i \in \Sigma_i$ for every i and $L = \{i : C_i \ne X_i\}$ is finite. Set

$$C'_i = C_i \text{ for } i \in J,\ X_i \text{ for } i \in I \setminus J, \quad C' = \prod_{i\in I} C'_i, \quad \alpha = \prod_{i\in I\setminus J} \mu_i C_i.$$

Then $\alpha\chi C'$ is a conditional expectation of χC on Λ_J. **P** By 254N, we can identify λ with the product of λ_J and $\lambda_{I\setminus J}$. This identifies Λ_J with $\{E \times X_{I\setminus J} : E \in \operatorname{dom}\lambda_J\}$. By 253H we have a conditional expectation g of χC defined by setting

$$g(y, z) = \int \chi C(y, t)\lambda_{I\setminus J}(dt)$$

for $y \in X_J$, $z \in X_{I\setminus J}$. But C is identified with $C_J \times C_{I\setminus J}$, where $C_J = \prod_{i\in J} C_i$, so that $g(y, z) = 0$ if $y \notin C_J$ and otherwise is $\lambda_{I\setminus J} C_{I\setminus J} = \alpha$. Thus $g = \alpha\chi(C_J \times X_{I\setminus J})$. But the identification between $X_I \times X_{I\setminus J}$ and X matches $C_J \times X_{I\setminus J}$ with C', as described above. So g becomes identified with $\alpha\chi C'$ and $\alpha\chi C'$ is a conditional expectation of χC. **Q**

(ii) Next, setting

$$C''_i = C'_i \text{ for } i \in K,\ X_i \text{ for } i \in I \setminus K, \quad C'' = \prod_{i\in I} C''_i,$$

$$\beta = \prod_{i\in I\setminus K} \mu_i C'_i = \prod_{i\in I\setminus(J\cup K)} \mu_i C_i,$$

the same arguments show that $\beta\chi C''$ is a conditional expectation of $\chi C'$ on Λ_K. So we have

$$P_K P_J(\chi C)^\bullet = \beta\alpha(\chi C'')^\bullet.$$

But if we look at $\beta\alpha$, this is just $\prod_{i\in I\setminus(K\cap J)} \mu_i C_i$, while $C''_i = C_i$ if $i \in K \cap J$, X_i for other i. So $\beta\alpha\chi C''$ is a conditional expectation of χC on $\Lambda_{K\cap J}$, and

$$P_K P_J(\chi C)^\bullet = P_{K\cap J}(\chi C)^\bullet.$$

(iii) Thus we see that the operators P_KP_J, $P_{K\cap J}$ agree on elements of the form $\chi C^\bullet$ where C is a measurable cylinder. Because they are both linear, they agree on linear combinations of these, that is, $P_KP_Jv=P_{K\cap J}v$ whenever $v=g^\bullet$ for some g in the space $\mathcal{S}$ of 254Q. But if $u\in L^1(\lambda)$ and $\epsilon>0$, there is a λ-integrable function f such that $f^\bullet=u$ and there is a $g\in\mathcal{S}$ such that $\int|f-g|\le\epsilon$ (254Qa), so that $\|u-v\|_1\le\epsilon$, where $v=g^\bullet$. Since P_J, P_K and $P_{K\cap J}$ are all linear operators of norm 1,

$$\|P_KP_Ju-P_{K\cap J}u\|_1\le 2\|u-v\|_1+\|P_KP_Jv-P_{K\cap J}v\|_1\le 2\epsilon.$$

As ϵ is arbitrary, $P_KP_Ju=P_{K\cap J}u$; as u is arbitrary, $P_KP_J=P_{K\cap J}$.

(b) Take $u\in L^1(\lambda)$. Let $\mathcal{J}$ be the family of all subsets J of I such that $P_Ju=u$. By (a), $J\cap K\in\mathcal{J}$ for all J, $K\in\mathcal{J}$. Next, $\mathcal{J}$ contains a countable set J_0. **P** Let f be a λ-integrable function such that $f^\bullet=u$. By 254Qc, we can find a countable set $J_0\subseteq I$ and a λ_{J_0}-integrable function g such that $f=_{\text{a.e.}}g\pi_{J_0}$. Now $g\pi_{J_0}$ is Λ_{J_0}-measurable and $u=(g\pi_{J_0})^\bullet$ belongs to $L^1_{J_0}$, so $J_0\in\mathcal{J}$. **Q**

Write $J^*=\bigcap\mathcal{J}$, so that $J^*\subseteq J_0$ is countable. Then $J^*\in\mathcal{J}$. **P** Let $\epsilon>0$. As in the proof of (a) above, there is a $g\in\mathcal{S}$ such that $\|u-v\|_1\le\epsilon$, where $v=g^\bullet$. But because g is a finite linear combination of indicator functions of measurable cylinders, each determined by coordinates in some finite set, there is a finite $K\subseteq I$ such that g is Λ_K-measurable, so that $P_Kv=v$. Because K is finite, there must be $J_1,\dots,J_n\in\mathcal{J}$ such that $J^*\cap K=\bigcap_{1\le i\le n}J_i\cap K$; but as $\mathcal{J}$ is closed under finite intersections, $J=J_1\cap\ldots\cap J_n\in\mathcal{J}$, and $J^*\cap K=J\cap K$.

Now we have

$$P_{J^*}v=P_{J^*}P_Kv=P_{J^*\cap K}v=P_{J\cap K}v=P_JP_Kv=P_Jv,$$

using (a) twice. Because both P_J and P_{J^*} have norm 1,

$$\begin{aligned}\|P_{J^*}u-u\|_1&\le\|P_{J^*}u-P_{J^*}v\|_1+\|P_{J^*}v-P_Jv\|_1+\|P_Jv-P_Ju\|_1+\|P_Ju-u\|_1\\&\le\|u-v\|_1+0+\|u-v\|_1+0\le 2\epsilon.\end{aligned}$$

As ϵ is arbitrary, $P_{J^*}u=u$ and $J^*\in\mathcal{J}$. **Q**

Now, for any $J\subseteq I$,

$$\begin{aligned}P_Ju=u&\Longrightarrow J\in\mathcal{J}\Longrightarrow J\supseteq J^*\\&\Longrightarrow P_Ju=P_JP_{J^*}u=P_{J\cap J^*}u=P_{J^*}u=u.\end{aligned}$$

Thus J^* has the required properties.

(c) Set $e=(\chi X)^\bullet$, $u_n=(-ne)\vee(u\wedge ne)$ for each $n\in\mathbb{N}$. Then, for any $J\subseteq I$, $u\in L^0_J$ iff $u_n\in L^0_J$ for every n. **P** (α) If $u\in L^0_J$, then u is expressible as $f^\bullet$ for some Λ_J-measurable f; now $f_n=(-n\chi X)\vee(f\wedge n\chi X)$ is Λ_J-measurable, so $u_n=f_n^\bullet\in L^0_J$ for every n. (β) If $u_n\in L^0_J$ for each n, then for each n we can find a Λ_J-measurable function f_n such that $f_n^\bullet=u_n$. But there is also a Λ-measurable function f such that $u=f^\bullet$, and we must have $f_n=_{\text{a.e.}}(-n\chi X)\vee(f\wedge n\chi X)$ for each n, so that $f=_{\text{a.e.}}\lim_{n\to\infty}f_n$ and $u=(\lim_{n\to\infty}f_n)^\bullet$. Since $\lim_{n\to\infty}f_n$ is Λ_J-measurable and defined on a $\mu\restriction\Lambda_J$-conegligible set, $u\in L^0_J$. **Q**

As every u_n belongs to L^1, we know that

$$u_n\in L^0_J\iff u_n\in L^1_J\iff P_Ju_n=u_n.$$

By (b), there is for each n a countable J_n^* such that $P_Ju_n=u_n$ iff $J\supseteq J_n^*$. So we see that $u\in L^0_J$ iff $J\supseteq J_n^*$ for every n, that is, $J\supseteq\bigcup_{n\in\mathbb{N}}J_n^*$. Thus $J^*=\bigcup_{n\in\mathbb{N}}J_n^*$ has the property claimed.

(d) Applying (c) to $u=(\chi W)^\bullet$, we have a (countable) unique smallest J^* such that $u\in L^0_{J^*}$. But if $J\subseteq I$, then there is a $W'\in\Lambda_J$ such that $W'\triangle W$ is negligible iff $u\in L^0_J$. So this is the J^* we are looking for.

(e) Again apply (c), this time to $f^\bullet$.

254S Proposition Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces, with product (X,Λ,λ).

(a) If $A\subseteq X$ is determined by coordinates in $I\setminus\{j\}$ for every $j\in I$, then its outer measure λ^*A must be either 0 or 1.

(b) If $W \in \Lambda$ and $\lambda W > 0$, then for every $\epsilon > 0$ there are a $W' \in \Lambda$ and a finite set $J \subseteq I$ such that $\lambda W' \geq 1 - \epsilon$ and for every $x \in W'$ there is a $y \in W$ such that $x \restriction I \setminus J = y \restriction I \setminus J$.

proof For $J \subseteq I$ write X_J for $\prod_{i \in J} X_i$ and λ_J for the product measure on X_J.

(a) Let W be a measurable envelope of A. By 254Rd, there is a smallest $J \subseteq I$ for which there is a $W' \in \Lambda$, determined by coordinates in J, with $\lambda(W \triangle W') = 0$. Now $J = \emptyset$. **P** Take any $j \in I$. Then A is determined by coordinates in $I \setminus \{j\}$, that is, can be regarded as $X_j \times A'$ for some $A' \subseteq X_{I \setminus \{j\}}$. We can also think of λ as the product of $\lambda_{\{j\}}$ and $\lambda_{I \setminus \{j\}}$ (254N). Let $\Lambda_{I \setminus \{j\}}$ be the domain of $\lambda_{I \setminus \{j\}}$. By 251S,

$$\lambda^* A = \lambda^*_{\{j\}} X_j \cdot \lambda^*_{I \setminus \{j\}} A' = \lambda^*_{I \setminus \{j\}} A'.$$

Let $V \in \Lambda_{I \setminus \{j\}}$ be measurable envelope of A'. Then $W' = X_j \times V$ belongs to Λ, includes A and has measure $\lambda^* A$, so $\lambda(W \cap W') = \lambda W = \lambda W'$ and $W \triangle W'$ is negligible. At the same time, W' is determined by coordinates in $I \setminus \{j\}$. This means that J must be included in $I \setminus \{j\}$. As j is arbitrary, $J = \emptyset$. **Q**

But the only subsets of X which are determined by coordinates in $\emptyset$ are X and $\emptyset$. Since W differs from one of these by a negligible set, $\lambda^* A = \lambda W \in \{0, 1\}$, as claimed.

(b) Set $\eta = \frac{1}{2} \min(\epsilon, 1) \lambda W$. By 254Fe, there is a measurable set V, determined by coordinates in a finite subset J of I, such that $\lambda(W \triangle V) \leq \eta$. Note that

$$\lambda V \geq \lambda W - \eta \geq \frac{1}{2} \lambda W > 0,$$

so

$$\lambda(W \triangle V) \leq \frac{1}{2} \epsilon \lambda W \leq \epsilon \lambda V.$$

We may identify λ with the c.l.d. product of λ_J and $\lambda_{I \setminus J}$ (254N). Let $\tilde{W}$, $\tilde{V} \subseteq X_I \times X_{I \setminus J}$ be the sets corresponding to W, $V \subseteq X$. Then $\tilde{V}$ can be expressed as $U \times X_{I \setminus J}$ where $\lambda_J U = \lambda V > 0$. Set $U' = \{z : z \in X_{I \setminus J}, \lambda_J \tilde{W}^{-1}[\{z\}] = 0\}$. Then U' is measured by $\lambda_{I \setminus J}$ (252D(ii) again, because both λ_J and $\lambda_{I \setminus J}$ are complete), and

$$\lambda_J U \cdot \lambda_{I \setminus J} U' \leq \int \lambda_J(\tilde{W}^{-1}[\{z\}] \triangle U) \lambda_{I \setminus J}(dz)$$

(because if $z \in U'$ then $\lambda_J(\tilde{W}^{-1}[\{z\}] \triangle U) = \lambda_J U$)

$$= \int \lambda_J(\tilde{W} \triangle \tilde{V})^{-1}[\{z\}] \lambda_{I \setminus J}(dz)$$

$$= (\lambda_J \times \lambda_{I \setminus J})(\tilde{W} \triangle \tilde{V})$$

(252D once more)

$$= \lambda(W \triangle V) \leq \epsilon \lambda V = \epsilon \lambda_J U.$$

This means that $\lambda_{I \setminus J} U' \leq \epsilon$. Set $W' = \{x : x \in X, x \restriction I \setminus J \notin U'\}$; then $\lambda W' \geq 1 - \epsilon$. If $x \in W'$, then $z = x \restriction I \setminus J \notin U'$, so $\tilde{W}^{-1}[\{z\}]$ is not empty, that is, there is a $y \in W$ such that $y \restriction I \setminus J = z$. So this W' has the required properties.

254T Remarks It is important to understand that the results above apply to L^0 and L^1 and measurable-sets-up-to-a-negligible-set, not to sets and functions themselves. One idea does apply to sets and functions, whether measurable or not.

(a) Let $\langle X_i \rangle_{i \in I}$ be a family of sets with Cartesian product X. For each $J \subseteq I$ let $\mathcal{W}_J$ be the set of subsets of X determined by coordinates in J. Then $\mathcal{W}_J \cap \mathcal{W}_K = \mathcal{W}_{J \cap K}$ for all J, $K \subseteq I$. **P** Of course $\mathcal{W}_J \cap \mathcal{W}_K \supseteq \mathcal{W}_{J \cap K}$, because $\mathcal{W}_J \supseteq \mathcal{W}_{J'}$ whenever $J' \subseteq J$. On the other hand, suppose $W \in \mathcal{W}_J \cap \mathcal{W}_K$, $x \in W$, $y \in X$ and $x \restriction J \cap K = y \restriction J \cap K$. Set $z(i) = x(i)$ for $i \in J$, $y(i)$ for $i \in I \setminus J$. Then $z \restriction J = x \restriction J$ so $z \in W$. Also $y \restriction K = z \restriction K$ so $y \in W$. As x, y are arbitrary, $W \in \mathcal{W}_{J \cap K}$; as W is arbitrary, $\mathcal{W}_J \cap \mathcal{W}_K \subseteq \mathcal{W}_{J \cap K}$. **Q** Accordingly, for any $W \subseteq X$, $\mathcal{F} = \{J : W \in \mathcal{W}_J\}$ is a filter on I (unless $W = X$ or $W = \emptyset$, in which case $\mathcal{F} = \mathcal{P}X$). But $\mathcal{F}$ does not necessarily have a least element, as the following example shows.

(b) Set $X = \{0,1\}^{\mathbb{N}}$,

$$W = \{x : x \in X,\, \lim_{i\to\infty} x(i) = 0\}.$$

Then for every $n \in \mathbb{N}$ W is determined by coordinates in $J_n = \{i : i \ge n\}$. But W is not determined by coordinates in $\bigcap_{n\in\mathbb{N}} J_n = \emptyset$. Note that

$$W = \bigcup_{n\in\mathbb{N}} \bigcap_{i\ge n} \{x : x(i) = 0\}$$

is measured by the usual measure on X. But it is also negligible (since it is countable); in 254Rd we have $J^* = \emptyset$, $W' = \emptyset$.

***254U** I am now in a position to describe a counter-example answering a natural question arising out of §251.

Example There are a localizable measure space (X, Σ, μ) and a probability space (Y, T, ν) such that the c.l.d. product measure λ on $X \times Y$ is not localizable.

proof (a) Take (X, Σ, μ) to be the space of 216E, so that $X = \{0,1\}^I$, where $I = \mathcal{P}C$ for some set C with cardinal greater than $\mathfrak{c}$. For each $\gamma \in C$ write E_γ for $\{x : x \in X,\, x(\{\gamma\}) = 1\}$ (that is, $G_{\{\gamma\}}$ in the notation of 216Ec); then $E_\gamma \in \Sigma$ and $\mu E_\gamma = 1$; also every measurable set of non-zero measure meets some E_γ in a set of non-zero measure, while $E_\gamma \cap E_\delta$ is negligible for all distinct γ, δ (see 216Ee).

Let (Y, T, ν) be $\{0,1\}^C$ with the usual measure (254J). For $\gamma \in C$, let F_γ be $\{y : y \in Y,\, y(\gamma) = 1\}$, so that $\nu F_\gamma = \frac{1}{2}$. Let λ be the c.l.d. product measure on $X \times Y$, and Λ its domain.

(b) Consider the family $\mathcal{W} = \{E_\gamma \times F_\gamma : \gamma \in C\} \subseteq \Lambda$. **?** Suppose, if possible, that V were an essential supremum of $\mathcal{W}$ in Λ in the sense of 211G. For $\gamma \in C$ write $H_\gamma = \{x : V[\{x\}]\triangle F_\gamma$ is negligible$\}$. Because $F_\gamma \triangle F_\delta$ is non-negligible, $H_\gamma \cap H_\delta = \emptyset$ for all $\gamma \ne \delta$.

Now $E_\gamma \setminus H_\gamma$ is μ-negligible for every $\gamma \in C$. **P** $\lambda((E_\gamma \times F_\gamma) \setminus V) = 0$, so $F_\gamma \setminus V[\{x\}]$ is negligible for almost every $x \in E_\gamma$, by 252D. On the other hand, if we set $F'_\gamma = Y \setminus F_\gamma$, $W_\gamma = (X \times Y) \setminus (E_\gamma \times F'_\gamma)$, then we see that

$$(E_\gamma \times F'_\gamma) \cap (E_\gamma \times F_\gamma) = \emptyset, \quad E_\gamma \times F_\gamma \subseteq W_\gamma,$$

$$\lambda((E_\delta \times F_\delta) \setminus W_\gamma) = \lambda((E_\gamma \times F'_\gamma) \cap (E_\delta \times F_\delta)) \le \mu(E_\gamma \cap E_\delta) = 0$$

for every $\delta \ne \gamma$, so W_γ is an essential upper bound for $\mathcal{W}$ and $V \cap (E_\gamma \times F'_\gamma) = V \setminus W_\gamma$ must be λ-negligible. Accordingly $V[\{x\}] \setminus F_\gamma = V[\{x\}] \cap F'_\gamma$ is ν-negligible for μ-almost every $x \in E_\gamma$. But this means that $V[\{x\}]\triangle F_\gamma$ is ν-negligible for μ-almost every $x \in E_\gamma$, that is, $\nu(E_\gamma \setminus H_\gamma) = 0$. **Q**

Now consider the family $\langle E_\gamma \cap H_\gamma \rangle_{\gamma\in C}$. This is a disjoint family of sets of finite measure in X. If $E \in \Sigma$ has non-zero measure, there is a $\gamma \in C$ such that $\mu(E_\gamma \cap H_\gamma \cap E) = \nu(E_\gamma \cap E) > 0$. But this means that $\mathcal{E} = \{E_\gamma \cap H_\gamma : \gamma \in C\}$ satisfies the conditions of 213Oa, and μ must be strictly localizable; which it isn't. **X**

(c) Thus we have found a family $\mathcal{W} \subseteq \Lambda$ with no essential supremum in Λ, and λ is not localizable.

Remark If (X, Σ, μ) and (Y, T, ν) are any localizable measure spaces with a non-localizable c.l.d. product measure, then their c.l.d. versions are still localizable (213Hb) and still have a non-localizable product (251T), which cannot be strictly localizable; so that at least one of the factors is not strictly localizable (251O). Thus any example of the type here must involve a complete locally determined localizable space which is not strictly localizable, as in 216E.

***254V** Corresponding to 251U and 251Wo, we have the following result on countable powers of atomless probability spaces.

Proposition Let (X, Σ, μ) be an atomless probability space and I a countable set. Let λ be the product probability measure on X^I. Then $\{x : x \in X^I,\, x \text{ is injective}\}$ is λ-conegligible.

proof For any pair $\{i, j\}$ of distinct elements of X, the set $\{z : z \in X^{\{i,j\}},\, z(i) = z(j)\}$ is negligible for the product measure on $X^{\{i,j\}}$, by 251U. By 254Oa, $\{x : x \in X,\, x(i) = x(j)\}$ is λ-negligible. Because I is countable, there are only countably many such pairs $\{i, j\}$, so $\{x : x \in X,\, x(i) = x(j)$ for some distinct

$i, j \in I\}$ is negligible, and its complement is conegligible; but this complement is just the set of injective functions from I to X.

254X Basic exercises (a) Let $\langle(X_i, \Sigma_i, \mu_i)\rangle_{i\in I}$ be any family of probability spaces, with product (X, Λ, μ). Write $\mathcal{E}$ for the family of subsets of X expressible as the union of a finite disjoint family of measurable cylinders. (i) Show that if $C \subseteq X$ is a measurable cylinder then $X \setminus C \in \mathcal{E}$. (ii) Show that $W \cap V \in \mathcal{E}$ for all W, $V \in \mathcal{E}$. (iii) Show that $X \setminus W \in \mathcal{E}$ for every $W \in \mathcal{E}$. (iv) Show that $\mathcal{E}$ is an algebra of subsets of X. (v) Show that for any $W \in \Lambda$, $\epsilon > 0$ there is a $V \in \mathcal{E}$ such that $\lambda(W \triangle V) \le \epsilon^2$. (vi) Show that for any $W \in \Lambda$ and $\epsilon > 0$ there are disjoint measurable cylinders $C_0, \dots, C_n$ such that $\lambda(W \cap C_j) \ge (1-\epsilon)\lambda C_j$ for every j and $\lambda(W \setminus \bigcup_{j\le n} C_j) \le \epsilon$. (*Hint*: select the C_j from the measurable cylinders composing a set V as in (v).) (vii) Show that if f, g are λ-integrable functions and $\int_C f \le \int_C g$ for every measurable cylinder $C \subseteq X$, then $f \le_{\text{a.e.}} g$. (*Hint*: show that $\int_W f \le \int_W f$ for every $W \in \Lambda$.)

>(b) Let $\langle(X_i, \Sigma_i, \mu_i)$ be a family of probability spaces, with product (X, Λ, λ). Show that the outer measure λ^* defined by λ is exactly the outer measure θ described in 254A, that is, that θ is a regular outer measure.

(c) Let $\langle(X_i, \Sigma_i, \mu_i)$ be a family of probability spaces, with product (X, Λ, λ). Write λ_0 for the restriction of λ to $\widehat{\bigotimes}_{i\in I}\Sigma_i$, and $\mathcal{C}$ for the family of measurable cylinders in X. Suppose that (Y, T, ν) is a probability space and $\phi : Y \to X$ a function. (i) Show that ϕ is inverse-measure-preserving when regarded as a function from (Y, T, ν) to $(X, \widehat{\bigotimes}_{i\in I}\Sigma_i, \lambda_0)$ iff $\phi^{-1}[C]$ belongs to T and $\nu\phi^{-1}[C] = \lambda_0 C$ for every $C \in \mathcal{C}$. (ii) Show that λ_0 is the only measure on X with this property. (*Hint*: 136C.)

>(d) Let I be a set and (Y, T, ν) a complete probability space. Show that a function $\phi : Y \to \{0,1\}^I$ is inverse-measure-preserving for ν and the usual measure on $\{0,1\}^I$ iff $\nu\{y : \phi(y)(i) = 1 \text{ for every } i \in J\} = 2^{-\#(J)}$ for every finite $J \subseteq I$.

>(e) Let I be any set and λ the usual measure on $X = \{0,1\}^I$. Define addition on X as in 254Jd. Show that the map $(x, y) \mapsto x + y : X \times X \to X$ is inverse-measure-preserving, if $X \times X$ is given its product measure.

>(f) Let I be any set and λ the usual measure on $\mathcal{P}I$. (i) Show that the map $a \mapsto a \triangle b : \mathcal{P}I \to \mathcal{P}I$ is inverse-measure-preserving for any $b \subseteq I$; in particular, $a \mapsto I \setminus a$ is inverse-measure-preserving. (ii) Show that the map $(a, b) \mapsto a \triangle b : \mathcal{P}I \times \mathcal{P}I \to \mathcal{P}I$ is inverse-measure-preserving.

>(g) Show that for any $q \in [0,1]$ and any set I there is a measure λ on $\mathcal{P}I$ such that $\lambda\{a : J \subseteq a\} = q^{\#(J)}$ for every finite $J \subseteq I$.

>(h) Let (Y, T, ν) be a complete probability space, and write μ for Lebesgue measure on $[0,1]$. Suppose that $\phi : Y \to [0,1]$ is a function such that $\nu\phi^{-1}[I]$ exists and is equal to μI for every interval I of the form $[2^{-n}k, 2^{-n}(k+1)]$, where $n \in \mathbb{N}$ and $0 \le k < 2^n$. Show that ϕ is inverse-measure-preserving for ν and μ.

(i) Show that if $\tilde{\phi} : \{0,1\}^{\mathbb{N}} \to [0,1]$ is any bijection constructed by the method of 254K, then $\{\tilde{\phi}^{-1}[E] : E \subseteq [0,1] \text{ is a Borel set}\}$ is just the σ-algebra of subsets of $\{0,1\}^{\mathbb{N}}$ generated by the sets $\{x : x(i) = 1\}$ for $i \in \mathbb{N}$.

(j) Let $\langle X_i\rangle_{i\in I}$ be a family of sets, and for each $i \in I$ let Σ_i be a σ-algebra of subsets of X_i. Show that for every $E \in \widehat{\bigotimes}_{i\in I}\Sigma_i$ there is a countable set $J \subseteq I$ such that E is expressible as $\pi_J^{-1}[F]$ for some $F \in \widehat{\bigotimes}_{i\in J}X_j$, writing $\pi_J(x) = x{\restriction}J \in \prod_{i\in J} X_i$ for $x \in \prod_{i\in I} X_i$.

(k)(i) Let ν be the usual measure on $X = \{0,1\}^{\mathbb{N}}$. Show that for any $k \ge 1$, (X, ν) is isomorphic to (X^k, ν_k), where ν_k is the measure on X^k which is the product measure obtained by giving each factor X the measure ν. (ii) Writing $\mu_{[0,1]}$ for Lebesgue measure on $[0,1]$, etc., show that for any $k \ge 1$, $([0,1]^k, \mu_{[0,1]^k})$ is isomorphic to $([0,1], \mu_{[0,1]})$.

(l)(i) Writing $\mu_{[0,1]}$ for Lebesgue measure on $[0,1]$, etc., show that $([0,1],\mu_{[0,1]})$ is isomorphic to $([0,1[\,,\mu_{[0,1[})$. (ii) Show that for any $k\ge 1$, $([0,1[^k\,,\mu_{[0,1[^k})$ is isomorphic to $([0,1[\,,\mu_{[0,1[})$. (iii) Show that for any $k\ge 1$, $(\mathbb{R},\mu_{\mathbb{R}})$ is isomorphic to $(\mathbb{R}^k,\mu_{\mathbb{R}^k})$.

(m) Let μ be Lebesgue measure on $[0,1]$ and λ the product measure on $[0,1]^{\mathbb{N}}$. Show that $([0,1],\mu)$ and $([0,1]^{\mathbb{N}},\lambda)$ are isomorphic.

(n) Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of complete probability spaces and λ the product measure on $\prod_{i\in I}X_i$, with domain Λ. Suppose that $A_i\subseteq X_i$ for each $i\in I$. Show that $\prod_{i\in I}A_i\in\Lambda$ iff either (i) $\prod_{i\in I}\mu_i^*A_i=0$ or (ii) $A_i\in\Sigma_i$ for every i and $\{i:A_i\ne X_i\}$ is countable. (*Hint*: assemble ideas from 252Xc, 254F, 254L and 254N.)

(o) Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces with product (X,Λ,λ). (i) Show that, for any $A\subseteq X$,

$$\lambda^*A=\min\{\lambda_J^*\pi_J[A]:J\subseteq I\text{ is countable}\},$$

where for $J\subseteq I$ I write λ_J for the product probability measure on $X_J=\prod_{i\in J}X_i$ and $\pi_J:X\to X_J$ for the canonical map. (ii) Show that if J, $K\subseteq I$ are disjoint and A, $B\subseteq X$ are determined by coordinates in J, K respectively, then $\lambda^*(A\cap B)=\lambda^*A\cdot\lambda^*B$.

(p) Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces with product (X,Λ,λ). Let $\mathcal{S}$ be the linear span of the set of indicator functions of measurable cylinders in X, as in 254Q. Show that $\{f^\bullet:f\in\mathcal{S}\}$ is dense in $L^p(\mu)$ for every $p\in[1,\infty[$.

(q) Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces, and (X,Λ,λ) their product; for $J\subseteq I$ let Λ_J be the σ-algebra of members of Λ determined by coordinates in J and $P_J:L^1=L^1(\lambda)\to L^1_J=L^1(\lambda\restriction\Lambda_J)$ the corresponding conditional expectation. (i) Show that if $u\in L^1_J$ and $v\in L^1_{I\setminus J}$ then $u\times v\in L^1$ and $\int u\times v=\int u\cdot\int v$. (*Hint*: 253D.) (ii) Show that if $\mathcal{J}\subseteq\mathcal{P}I$ is non-empty, with $J^*=\bigcap\mathcal{J}$, then $L^1_{J^*}=\bigcap_{J\in\mathcal{J}}L^1_J$.

(r)(i) Let I be any set and λ the usual measure on $\mathcal{P}I$. Let $A\subseteq\mathcal{P}I$ be such that $a\triangle b\in A$ whenever $a\in A$ and $b\subseteq I$ is finite. Show that λ^*A must be either 0 or 1. (ii) Let λ be the usual measure on $\{0,1\}^{\mathbb{N}}$, and Λ its domain. Let $f:\{0,1\}^{\mathbb{N}}\to\mathbb{R}$ be a function such that, for x, $y\in\{0,1\}^{\mathbb{N}}$, $f(x)=f(y)\iff\{n:n\in\mathbb{N},\ x(n)\ne y(n)\}$ is finite. Show that f is not Λ-measurable. (*Hint*: for any $q\in\mathbb{Q}$, $\lambda^*\{x:f(x)\le q\}$ is either 0 or 1.)

(s) Let $\langle X_i\rangle_{i\in I}$ be any family of sets and $A\subseteq B\subseteq\prod_{i\in I}X_i$. Suppose that A is determined by coordinates in $J\subseteq I$ and that B is determined by coordinates in K. Show that there is a set C such that $A\subseteq C\subseteq B$ and C is determined by coordinates in $J\cap K$.

254Y Further exercises (a) Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces, and for $J\subseteq I$ let λ_J be the product measure on $X_J=\prod_{i\in J}X_i$; write $X=X_I$, $\lambda=\lambda_I$ and $\pi_J(x)=x\restriction J$ for $x\in X$ and $J\subseteq I$.

(i) Show that for $K\subseteq J\subseteq I$ we have a natural linear, order-preserving and norm-preserving map $T_{JK}:L^1(\lambda_K)\to L^1(\lambda_J)$ defined by writing $T_{JK}(f^\bullet)=(f\pi_{KJ})^\bullet$ for every λ_K-integrable function f, where $\pi_{KJ}(y)=y\restriction K$ for $y\in X_J$.

(ii) Write $\mathcal{K}$ for the set of finite subsets of I. Show that if W is any Banach space and $\langle T_K\rangle_{K\in\mathcal{K}}$ is a family such that (α) T_K is a bounded linear operator from $L^1(\lambda_K)$ to W for every $K\in\mathcal{K}$ (β) $T_K=T_JT_{JK}$ whenever $K\subseteq J\in\mathcal{K}$ (γ) $\sup_{K\in\mathcal{K}}\|T_K\|<\infty$, then there is a unique bounded linear operator $T:L^1(\lambda)\to W$ such that $T_K=TT_{IK}$ for every $K\in\mathcal{K}$.

(iii) Write $\mathcal{J}$ for the set of countable subsets of I. Show that $L^1(\lambda)=\bigcup_{J\in\mathcal{J}}T_{IJ}[L^1(\lambda_J)]$.

(b) Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be any family of measure spaces. Set $X=\prod_{i\in I}X_i$ and let $\mathcal{F}$ be a filter on the set $[I]^{<\omega}$ of finite subsets of I such that $\{J:i\in J\in[I]^{<\omega}\}\in\mathcal{F}$ for every $i\in I$. Show that there is a complete locally determined measure λ on X such that $\lambda(\prod_{i\in I}E_i)$ is defined and equal to $\lim_{J\to\mathcal{F}}\prod_{i\in J}\mu_iE_i$ whenever $E_i\in\Sigma_i$ for every $i\in I$ and $\lim_{J\to\mathcal{F}}\prod_{i\in J}\mu_iE_i$ is defined in $[0,\infty[$. (*Hint*: Baker 04.)

(c) Let $\langle(X_i,\Sigma_i,\mu_i)\rangle_{i\in I}$ be a family of probability spaces, and λ a complete measure on $X=\prod_{i\in I}X_i$. Suppose that for every complete probability space (Y,T,ν) and function $\phi:Y\to X$, ϕ is inverse-measure-preserving for ν and λ iff $\nu\phi^{-1}[C]$ is defined and equal to θ_0C for every measurable cylinder $C\subseteq X$, writing θ_0 for the functional of 254A. Show that λ is the product measure on X.

(d) Let I be a set, and λ the usual measure on $\{0,1\}^I$. Show that $L^1(\lambda)$ is separable, in its norm topology, iff I is countable.

(e) Let $f:[0,1]\to[0,1]^2$ be a function which is inverse-measure-preserving for Lebesgue planar measure on $[0,1]^2$ and Lebesgue linear measure on $[0,1]$, as in 134Yl; let f_1, f_2 be the coordinates of f. Define $g:[0,1]\to[0,1]^{\mathbb{N}}$ by setting $g(t)=\langle f_1f_2^n(t)\rangle_{n\in\mathbb{N}}$ for $0\le t\le 1$. Show that g is inverse-measure-preserving. (*Hint*: show that $g_n:[0,1]\to[0,1]^{n+1}$ is inverse-measure-preserving for every $n\ge 1$, where $g_n(t)=(f_1(t),f_1f_2(t),\dots,f_1f_2^{n-1}(t),f_2^n(t))$ for $t\in[0,1]$.)

(f) Let I be a set, and λ the usual measure on $\mathcal{P}I$. Show that if $\mathcal{F}$ is a non-principal ultrafilter on I then $\lambda^*\mathcal{F}=1$. (*Hint*: 254Xr, 254Xf.)

(g) Let (X,Σ,μ), (Y,T,ν) and λ be as in 254U. Set $A=\{x_\gamma:\gamma\in C\}$ as defined in 216E. Let μ_A be the subspace measure on A, and $\tilde\lambda$ the c.l.d. product measure of μ_A and ν on $A\times Y$. Show that $\tilde\lambda$ is a proper extension of the subspace measure $\lambda_{A\times Y}$. (*Hint*: consider $\tilde W=\{(x_\gamma,y):\gamma\in C,\ y\in F_\gamma\}$, in the notation of 254U.)

(h) Let (X,Σ,μ) be an atomless probability space, I a set with cardinal at most $\#(X)$, and A the set of injective functions from I to X. Show that A has full outer measure for the product measure on X^I.

254 Notes and comments While there are many reasons for studying infinite products of probability spaces, one stands pre-eminent, from the point of view of abstract measure theory: they provide constructions of essentially new kinds of measure space. I cannot describe the nature of this 'newness' effectively without venturing into the territory of Volume 3. But the function spaces of Chapter 24 do give at least a form of words we can use: these are the first *probability* spaces (X,Λ,λ) we have seen for which $L^1(\lambda)$ need not be separable for its norm topology (254Yd).

The formulae of 254A, like those of 251A, lead very naturally to measures; the point at which they become more than a curiosity is when we find that the product measure λ is a probability measure (254Fa), which must be regarded as the crucial argument of this section, just as 251E is the essential basis of §251. It is I think remarkable that it makes no difference to the result here whether I is finite, countably infinite or uncountable. If you write out the proof for the case $I=\mathbb{N}$, it will seem natural to expand the sets J_n until they are initial segments of I itself, thereby avoiding altogether the auxiliary set K; but this is a misleading simplification, because it hides an essential feature of the argument, which is that any sequence in $\mathcal{C}$ involves only countably many coordinates, so that as long as we are dealing with only one such sequence the uncountability of the whole set I is irrelevant. This general principle naturally permeates the whole of the section; in 254O I have tried to spell out the way in which many of the questions we are interested in can be expressed in terms of countable subproducts of the factor spaces X_i. See also the exercises 254Xj, 254Xn and 254Ya(iii).

There is a slightly paradoxical side to this principle: even the best-behaved subsets E_i of X_i may fail to have measurable products $\prod_{i\in I}E_i$ if $E_i\ne X_i$ for uncountably many i. For instance, $]0,1[^I$ is not a measurable subset of $[0,1]^I$ if I is uncountable (254Xn). It has full outer measure and its own product measure is just the subspace measure (254L), but any measurable subset must have measure zero. The point is that the empty set is the only member of $\widehat{\bigotimes}_{i\in I}\Sigma_i$, where Σ_i is the algebra of Lebesgue measurable subsets of $[0,1]$ for each i, which is included in $]0,1[^I$ (see 254Xj).

As in §251, I use a construction which automatically produces a complete measure on the product space. I am sure that this is the best choice for 'the' product measure. But there are occasions when its restriction to the σ-algebra generated by the measurable cylinders is worth looking at; see 254Xc.

Lemma 254G is a result of a type which will be commoner in Volume 3 than in the present volume. It describes the product measure in terms not of what it *is* but of what it *does*; specifically, in terms of a property of the associated family of inverse-measure-preserving functions. It is therefore a 'universal mapping theorem'. (Compare 253F.) Because this description is sufficient to determine the product measure completely (254Yc), it is not surprising that I use it repeatedly.

The 'usual measure' on $\{0,1\}^I$ (254J) is sometimes called 'coin-tossing measure' because it can be used to model the concept of tossing a coin arbitrarily many times indexed by the set I, taking an $x \in \{0,1\}^I$ to represent the outcome in which the coin is 'heads' for just those $i \in I$ for which $x(i) = 1$. The sets, or 'events', in the class $\mathcal{C}$ are those which can be specified by declaring the outcomes of finitely many tosses, and the probability of any particular sequence of n results is $1/2^n$, regardless of which tosses we look at or in which order. In Chapter 27 I will return to the use of product measures to represent probabilities involving independent events.

In 254K I come to the first case in this treatise of a non-trivial isomorphism between two measure spaces. If you have been brought up on a conventional diet of modern abstract pure mathematics based on algebra and topology, you may already have been struck by the absence of emphasis on any concept of 'homomorphism' or 'isomorphism'. Here indeed I start to speak of 'isomorphisms' between measure spaces without even troubling to define them; I hope it really is obvious that an isomorphism between measure spaces (X, Σ, μ) and (Y, T, ν) is a bijection $\phi : X \to Y$ such that $\mathrm{T} = \{F : F \subseteq Y,\ \phi^{-1}[F] \in \Sigma\}$ and $\nu F = \mu\phi^{-1}[F]$ for every $F \in \mathrm{T}$, so that Σ is necessarily $\{E : E \subseteq X,\ \phi[E] \in \mathrm{T}\}$ and $\mu E = \nu\phi[E]$ for every $E \in \Sigma$. Put like this, you may, if you worked through the exercises of Volume 1, be reminded of some constructions of σ-algebras in 111Xc-111Xd and of the 'image measures' in 234C-234D. The result in 254K (see also 134Yo) naturally leads to two distinct notions of 'homomorphism' between two measure spaces (X, Σ, μ) and (Y, T, ν):

(i) a function $\phi : X \to Y$ such that $\phi^{-1}[F] \in \Sigma$ and $\mu\phi^{-1}[F] = \nu F$ for every $F \in \mathrm{T}$,

(ii) a function $\phi : X \to Y$ such that $\phi[E] \in \mathrm{T}$ and $\nu\phi[E] = \mu E$ for every $E \in \Sigma$.

On either definition, we find that a bijection $\phi : X \to Y$ is an isomorphism iff ϕ and ϕ^{-1} are both homomorphisms. (Also, of course, the composition of homomorphisms will be a homomorphism.) My own view is that (i) is the more important, and in this treatise I study such functions at length, calling them 'inverse-measure-preserving'. But both have their uses. The function ϕ of 254K not only satisfies both definitions, but is also 'nearly' an isomorphism in several different ways, of which possibly the most important is that there are conegligible sets $X' \subseteq \{0,1\}^{\mathbb{N}}$, $Y' \subseteq [0,1]$ such that $\phi\restriction X'$ is an isomorphism between X' and Y' when both are given their subspace measures.

Having once established the isomorphism between $[0,1]$ and $\{0,1\}^{\mathbb{N}}$, we are led immediately to many more; see 254Xk-254Xm. In fact Lebesgue measure on $[0,1]$ is isomorphic to a large proportion of the probability spaces arising in applications. In Volumes 3 and 4 I will discuss these isomorphisms at length.

The general notion of 'subproduct' is associated with some of the deepest and most characteristic results in the theory of product measures. Because we are looking at products of arbitrary families of probability spaces, the definition must ignore any possible structure in the index set I of 254A-254C. But many applications, naturally enough, deal with index sets with favoured subsets or partitions, and the first essential step is the 'associative law' (254N; compare 251Xe-251Xf and 251Wh). This is, for instance, the tool by which we can apply Fubini's theorem within infinite products. The natural projection maps from $\prod_{i\in I} X_i$ to $\prod_{i\in J} X_i$, where $J \subseteq I$, are related in a way which has already been used as the basis of theorems in §235; the product measure on $\prod_{i\in J} X_i$ is precisely the image of the product measure on $\prod_{i\in I} X_i$ (254Oa). In 254O-254Q I explore the consequences of this fact and the fact already noted that all measurable sets in the product are 'essentially' determined by coordinates in some countable set.

In 254R I go more deeply into this notion of a set $W \subseteq \prod_{i\in I} X_i$ 'determined by coordinates in' a set $J \subseteq I$. In its primitive form this is a purely set-theoretic notion (254M, 254Ta). I think that even a three-element set I can give us surprises; I invite you to try to visualize subsets of $[0,1]^3$ which are determined by pairs of coordinates. But the interactions of this with measure-theoretic ideas, and in particular with a willingness to add or discard negligible sets, lead to much more, and in particular to the unique minimal sets of coordinates associated with measurable sets and functions (254R). Of course these results can be elegantly and effectively described in terms of L^1 and L^0 spaces, in which negligible sets are swept out of sight as the spaces are constructed. The basis of all this is the fact that the conditional expectation operators associated with subproducts multiply together in the simplest possible way (254Ra); but some further idea is needed to show that if $\mathcal{J}$ is a non-empty family of subsets of I, then $L^0_{\bigcap\mathcal{J}} = \bigcap_{J\in\mathcal{J}} L^0_J$ (see part (b) of the proof of 254R, and 254Xq(iii)).

254Sa is a version of the 'zero-one law' (272O below). 254Sb is a strong version of the principle that measurable sets in a product must be approximable by sets determined by a *finite* set of coordinates (254Fe, 254Qa, 254Xa). Evidently it is not a coincidence that the set W of 254Tb is negligible. In §272 I will revisit many of the ideas of 254R-254S and 254Xq, in particular, in the more general context of 'independent σ-algebras'.

Finally, 254U and 254Yg hardly belong to this section at all; they are unfinished business from §251. They are here because the construction of 254A-254C is the simplest way to produce an adequately complex probability space (Y, T, ν).

Version of 3.7.08

255 Convolutions of functions

I devote a section to a construction which is of great importance – and will in particular be very useful in Chapters 27 and 28 – and may also be regarded as a series of exercises on the work so far.

I find it difficult to know how much repetition to indulge in in this section, because the natural unified expression of the ideas is in the theory of topological groups, and I do not think we are yet ready for the general theory (I will come to it in Chapter 44 in Volume 4). The groups we need for this volume are

$\mathbb{R}$;

$\mathbb{R}^r$, for $r \ge 2$;

$S^1 = \{z : z \in \mathbb{C}, |z| = 1\}$, the 'circle group';

$\mathbb{Z}$, the group of integers.

All the ideas already appear in the theory of convolutions on $\mathbb{R}$, and I will therefore present this material in relatively detailed form, before sketching the forms appropriate to the groups $\mathbb{R}^r$ and S^1 (or $]-\pi, \pi]$); $\mathbb{Z}$ can I think be safely left to the exercises.

255A This being a book on measure theory, it is perhaps appropriate for me to emphasize, as the basis of the theory of convolutions, certain measure space isomorphisms.

Theorem Let μ be Lebesgue measure on $\mathbb{R}$ and μ_2 Lebesgue measure on $\mathbb{R}^2$; write Σ, Σ_2 for their domains.

(a) For any $a \in \mathbb{R}$, the map $x \mapsto a + x : \mathbb{R} \to \mathbb{R}$ is a measure space automorphism of $(\mathbb{R}, \Sigma, \mu)$.

(b) The map $x \mapsto -x : \mathbb{R} \to \mathbb{R}$ is a measure space automorphism of $(\mathbb{R}, \Sigma, \mu)$.

(c) For any $a \in \mathbb{R}$, the map $x \mapsto a - x : \mathbb{R} \to \mathbb{R}$ is a measure space automorphism of $(\mathbb{R}, \Sigma, \mu)$.

(d) The map $(x, y) \mapsto (x + y, y) : \mathbb{R}^2 \to \mathbb{R}^2$ is a measure space automorphism of $(\mathbb{R}^2, \Sigma_2, \mu_2)$.

(e) The map $(x, y) \mapsto (x - y, y) : \mathbb{R}^2 \to \mathbb{R}^2$ is a measure space automorphism of $(\mathbb{R}^2, \Sigma_2, \mu_2)$.

Remark I ought to remark that (b), (d) and (e) may be regarded as simple special cases of Theorem 263A in the next chapter. I nevertheless feel that it is worth writing out separate proofs here, partly because the general case of linear operators dealt with in 263A requires some extra machinery not needed here, but more because the result here has nothing to do with the *linear* structure of $\mathbb{R}$ and $\mathbb{R}^2$; it is exclusively dependent on the *group* structure of $\mathbb{R}$, together with the links between its topology and measure, and the arguments I give now are adaptable to the proper generalizations to abelian topological groups.

proof (a) This is just the translation-invariance of Lebesgue measure, dealt with in §134. There I showed that if $E \in \Sigma$ then $E + a \in \Sigma$ and $\mu(E + a) = \mu E$ (134Ab); that is, writing $\phi(x) = x + a$, $\mu(\phi[E])$ exists and is equal to μE for every $E \in \Sigma$. But of course we also have

$$\mu(\phi^{-1}[E]) = \mu(E + (-a)) = \mu E$$

for every $E \in \Sigma$, so ϕ is an automorphism.

(b) The point is that $\mu^*(A) = \mu^*(-A)$ for every $A \subseteq \mathbb{R}$. **P** (I follow the definitions of Volume 1.) If $\epsilon > 0$, there is a sequence $\langle I_n \rangle_{n \in \mathbb{N}}$ of half-open intervals covering A with $\sum_{n=0}^{\infty} \mu I_n \le \mu^* A + \epsilon$. Now $-A \subseteq \bigcup_{n \in \mathbb{N}} (-I_n)$. But if $I_n = [a_n, b_n[$ then $-I_n =]-b_n, a_n]$, so

$$\mu^*(-A) \le \sum_{n=0}^{\infty} \mu(-I_n) = \sum_{n=0}^{\infty} \max(0, -a_n - (-b_n)) = \sum_{n=0}^{\infty} \mu I_n \le \mu^* A + \epsilon.$$

As ϵ is arbitrary, $\mu^*(-A) \le \mu^* A$. Also of course $\mu^* A \le \mu^*(-(-A)) = \mu^* A$, so $\mu^*(-A) = \mu^* A$. **Q**

This means that, setting $\phi(x) = -x$ this time, ϕ is an automorphism of the structure $(\mathbb{R}, \mu^*)$. But since μ is defined from μ^* by the abstract procedure of Carathéodory's method, ϕ must also be an automorphism of the structure $(\mathbb{R}, \Sigma, \mu)$.

(c) Put (a) and (b) together; $x \mapsto a - x$ is the composition of the automorphisms $x \mapsto -x$ and $x \mapsto a + x$, and the composition of automorphisms is surely an automorphism.

(d)(i) Write T for the set $\{E : E \in \Sigma_2,\ \phi[E] \in \Sigma_2\}$, where this time $\phi(x, y) = (x + y, y)$ for $x,\ y \in \mathbb{R}$, so that $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ is a permutation. Then T is a σ-algebra, being the intersection of the σ-algebras Σ_2 and $\{E : \phi[E] \in \Sigma_2\} = \{\phi^{-1}[F] : F \in \Sigma_2\}$. Moreover, $\mu_2 E = \mu_2(\phi[E])$ for every $E \in \mathrm{T}$. **P** By 252D, we have

$$\mu_2 E = \int \mu\{x : (x, y) \in E\}\mu(dy).$$

But applying the same result to $\phi[E]$ we have

$$\mu_2\phi[E] = \int \mu\{x : (x, y) \in \phi[E]\}\mu(dy) = \int \mu\{x : (x - y, y) \in E\}\mu(dy)$$

$$= \int \mu(E^{-1}[\{y\}] + y)\mu(dy) = \int \mu E^{-1}[\{y\}]\mu(dy)$$

(because Lebesgue measure is translation-invariant)

$$= \mu_2 E. \ \mathbf{Q}$$

(ii) Now ϕ and ϕ^{-1} are clearly continuous, so that $\phi[G]$ is open, and therefore measurable, for every open G; consequently all open sets must belong to T. Because T is a σ-algebra, it contains all Borel sets. Now let E be any measurable set. Then there are Borel sets H_1, H_2 such that $H_1 \subseteq E \subseteq H_2$ and $\mu_2(H_2 \setminus H_1) = 0$ (134Fb). We have $\phi[H_1] \subseteq \phi[E] \subseteq \phi[H_2]$ and

$$\mu(\phi[H_2] \setminus \phi[H_1]) = \mu\phi[H_2 \setminus H_1] = \mu(H_2 \setminus H_1) = 0.$$

Thus $\phi[E] \setminus \phi[H_1]$ must be negligible, therefore measurable, and $\phi[E] = \phi[H_1] \cup (\phi[E] \setminus \phi[H_1])$ is measurable. This shows that $\phi[E]$ is measurable whenever E is.

(iii) Repeating the same arguments with $-y$ in the place of y, we see that $\phi^{-1}[E]$ is measurable, and $\mu_2\phi^{-1}[E] = \mu_2 E$, for every $E \in \Sigma_2$. So ϕ is an automorphism of the structure $(\mathbb{R}^2, \Sigma_2, \mu_2)$.

(e) Of course this is an immediate corollary either of the proof of (d) or of (d) itself as stated, since $(x, y) \mapsto (x - y, y)$ is just the inverse of $(x, y) \mapsto (x + y, y)$.

255B Corollary (a) If $a \in \mathbb{R}$, then for any complex-valued function f defined on a subset of $\mathbb{R}$

$$\int f(x)dx = \int f(a + x)dx = \int f(-x)dx = \int f(a - x)dx$$

in the sense that if one of the integrals exists so do the others, and they are then all equal.

(b) If f is a complex-valued function defined on a subset of $\mathbb{R}^2$, then

$$\int f(x + y, y)d(x, y) = \int f(x - y, y)d(x, y) = \int f(x, y)d(x, y)$$

in the sense that if one of the integrals exists and is finite so does the other, and they are then equal.

255C Remarks (a) I am not sure whether it ought to be 'obvious' that if (X, Σ, μ), (Y, T, ν) are measure spaces and $\phi : X \to Y$ is an isomorphism, then for any function f defined on a subset of Y

$$\int f(\phi(x))\mu(dx) = \int f(y)\nu(dy)$$

in the sense that if one is defined so is the other, and they are then equal. If it is obvious then the obviousness must be contingent on the nature of the definition of integration: integrability with respect to the measure μ is something which depends on the structure (X, Σ, μ) and on no other properties of X. If it is not obvious then it is an easy deduction from Theorem 235A above, applied in turn to ϕ and ϕ^{-1} and to the real and imaginary parts of f. In any case the isomorphisms of 255A are just those needed to prove 255B.

(b) Note that in 255Bb I write $\int f(x,y)d(x,y)$ to emphasize that I am considering the integral of f with respect to two-dimensional Lebesgue measure. The fact that

$$\int\Bigl(\int f(x,y)dx\Bigr)dy = \int\Bigl(\int f(x+y,y)dx\Bigr)dy = \int\Bigl(\int f(x-y,y)dx\Bigr)dy$$

is actually easier, being an immediate consequence of the equality $\int f(a+x)dx = \int f(x)dx$. But applications of this result often depend essentially on the fact that the functions $(x,y) \mapsto f(x+y,y)$, $(x,y) \mapsto f(x-y,y)$ are measurable as functions of two variables.

(c) I have moved directly to complex-valued functions because these are necessary for the applications in Chapter 28. If however they give you any discomfort, either technically or aesthetically, all the measure-theoretic ideas of this section are already to be found in the real case, and you may wish at first to read it as if only real numbers were involved.

255D A further corollary of 255A will be useful.

Corollary Let f be a complex-valued function defined on a subset of $\mathbb{R}$.

(a) If f is measurable, then the functions $(x,y) \mapsto f(x+y)$, $(x,y) \mapsto f(x-y)$ are measurable.

(b) If f is defined almost everywhere in $\mathbb{R}$, then the functions $(x,y) \mapsto f(x+y)$, $(x,y) \mapsto f(x-y)$ are defined almost everywhere in $\mathbb{R}^2$.

proof Writing $g_1(x,y) = f(x+y)$, $g_2(x,y) = f(x-y)$ whenever these are defined, we have

$$g_1(x,y) = (f \otimes \chi\mathbb{R})(\phi(x,y)), \quad g_2(x,y) = (f \otimes \chi\mathbb{R})(\phi^{-1}(x,y)),$$

writing $\phi(x,y) = (x+y,y)$ as in 255B(d-e), and $(f \otimes \chi\mathbb{R})(x,y) = f(x)$, following the notation of 253B. By 253C, $f \otimes \chi\mathbb{R}$ is measurable if f is, and defined almost everywhere if f is. Because ϕ is a measure space automorphism, $(f \otimes \chi\mathbb{R})\phi = g_1$ and $(f \otimes \chi\mathbb{R})\phi^{-1} = g_2$ are measurable, or defined almost everywhere, if f is.

255E The basic formula Let f and g be measurable complex-valued functions defined almost everywhere in $\mathbb{R}$. Write $f * g$ for the function defined by the formula

$$(f*g)(x) = \int f(x-y)g(y)dy$$

whenever the integral exists (with respect to Lebesgue measure, naturally) as a complex number. Then $f*g$ is the **convolution** of the functions f and g.

Observe that $\operatorname{dom}(|f|*|g|) = \operatorname{dom}(f*g)$, and that $|f*g| \le |f|*|g|$ everywhere on their common domain, for all f and g.

Remark Note that I am here prepared to contemplate the convolution of f and g for arbitrary members of $\mathcal{L}^0_{\mathbb{C}}$, the space of almost-everywhere-defined measurable complex-valued functions, even though the domain of $f*g$ may be empty.

255F Elementary properties (a) Because integration is linear, we surely have

$$((f_1+f_2)*g)(x) = (f_1*g)(x) + (f_2*g)(x),$$

$$(f*(g_1+g_2))(x) = (f*g_1)(x) + (f*g_2)(x),$$

$$(cf*g)(x) = (f*cg)(x) = c(f*g)(x)$$

whenever the right-hand sides of the formulae are defined.

(b) If f and g are measurable complex-valued functions defined almost everywhere in $\mathbb{R}$, then $f*g = g*f$, in the strict sense that they have the same domain and the same value at each point of that common domain.

P Take $x \in \mathbb{R}$ and apply 255Ba to see that

$$(f*g)(x) = \int f(x-y)g(y)dy = \int f(x-(x-y))g(x-y)dy$$
$$= \int f(y)g(x-y)dy = (g*f)(x)$$

if either is defined. **Q**

(c) If f_1, f_2, g_1, g_2 are measurable complex-valued functions defined almost everywhere in $\mathbb{R}$, $f_1 =_{\text{a.e.}} f_2$ and $g_1 =_{\text{a.e.}} g_2$, then $f_1 * g_1 = f_2 * g_2$. **P** For every $x \in \mathbb{R}$ we shall have $f_1(x-y) = f_2(x-y)$ for almost every $y \in \mathbb{R}$, by 255Ac. Consequently $f_1(x-y)g_1(y) = f_2(x-y)g_2(y)$ for almost every y, and $(f_1 * g_1)(x) = (f_2 * g_2)(x)$ in the sense that if one of these is defined so is the other, and they are then equal. **Q**

It follows that if $u, v \in L^0_{\mathbb{C}}$, then we have a function $\theta(u,v)$ which is equal to $f*g$ whenever f, $g \in \mathcal{L}^0_{\mathbb{C}}$ are such that $f^\bullet = u$ and $g^\bullet = u$. Observe that $\theta(u,v) = \theta(v,u)$, and that $\theta(u_1+u_2,v)$ extends $\theta(u_1,v)+\theta(u_2,v)$, $\theta(cu,v)$ extends $c\theta(u,v)$ for all u, u_1, u_2, $v \in L^0_{\mathbb{C}}$ and $c \in \mathbb{C}$.

255G I have grouped 255Fa-255Fc together because they depend only on ideas up to and including 255Ac and 255Ba. Using the second halves of 255A and 255B we get much deeper. I begin with what seems to be the fundamental result.

Theorem Let f, g and h be measurable complex-valued functions defined almost everywhere in $\mathbb{R}$.

(a) Suppose that $\int h(x+y)f(x)g(y)d(x,y)$ exists in $\mathbb{C}$. Then

$$\begin{aligned}\int h(x)(f*g)(x)dx &= \int h(x+y)f(x)g(y)d(x,y)\\ &= \iint h(x+y)f(x)g(y)dxdy = \iint h(x+y)f(x)g(y)dydx\end{aligned}$$

provided that in the expression $h(x)(f*g)(x)$ we interpret the product as 0 if $h(x)=0$ and $(f*g)(x)$ is undefined.

(b) If, on a similar interpretation of $|h(x)|(|f|*|g|)(x)$, the integral $\int |h(x)|(|f|*|g|)(x)dx$ is finite, then $\int h(x+y)f(x)g(y)d(x,y)$ exists in $\mathbb{C}$.

proof Consider the functions

$$k_1(x,y) = h(x)f(x-y)g(y), \quad k_2(x,y) = h(x+y)f(x)g(y)$$

wherever these are defined. 255D tells us that k_1 and k_2 are measurable and defined almost everywhere. Now setting $\phi(x,y) = (x+y,y)$, we have $k_2 = k_1\phi$, so that

$$\int k_1(x,y)d(x,y) = \int k_2(x,y)d(x,y)$$

if either exists, by 255Bb.

If

$$\int h(x+y)f(x)g(y)d(x,y) = \int k_2$$

exists, then by Fubini's theorem we have

$$\int k_2 = \int k_1(x,y)d(x,y) = \int(\int h(x)f(x-y)g(y)dy)dx$$

so $\int h(x)f(x-y)g(y)dy$ exists almost everywhere, that is, $(f*g)(x)$ exists for almost every x such that $h(x) \neq 0$; on the interpretation I am using here, $h(x)(f*g)(x)$ exists almost everywhere, and

$$\begin{aligned}\int h(x)(f*g)(x)dx &= \int\Bigl(\int h(x)f(x-y)g(y)dy\Bigr)dx = \int k_1\\ &= \int k_2 = \int h(x+y)f(x)g(y)d(x,y)\\ &= \iint h(x+y)f(x)g(y)dxdy = \iint h(x+y)f(x)g(y)dydx\end{aligned}$$

by Fubini's theorem again.

If (on the same interpretation) $|h| \times (|f|*|g|)$ is integrable,

$$|k_1(x,y)| = |h(x)||f(x-y)||g(y)|$$

is measurable, and

$$\iint |h(x)||f(x-y)||g(y)|dydx = \int |h(x)|(|f|*|g|)(x)dx$$

is finite, so by Tonelli's theorem (252G, 252H) k_1 and k_2 are integrable.

255H Certain standard results are now easy.

Corollary If f, g are complex-valued functions which are integrable over $\mathbb{R}$, then $f*g$ is integrable, with

$$\int f*g=\int f\int g,\quad \int|f*g|\le\int|f|\int|g|.$$

proof In 255G, set $h(x)=1$ for every $x\in\mathbb{R}$; then

$$\int h(x+y)f(x)g(y)d(x,y)=\int f(x)g(y)d(x,y)=\int f\int g$$

by 253D, so

$$\int f*g=\int h(x)(f*g)(x)dx=\int h(x+y)f(x)g(y)d(x,y)=\int f\int g,$$

as claimed. Now

$$\int|f*g|\le\int|f|*|g|=\int|f|\int|g|.$$

255I Corollary For any measurable complex-valued functions f, g defined almost everywhere in $\mathbb{R}$, $f*g$ is measurable and has measurable domain.

proof Set $f_n(x)=f(x)$ if $x\in\operatorname{dom}f$, $|x|\le n$ and $|f(x)|\le n$, and 0 elsewhere in $\mathbb{R}$; define g_n similarly from g. Then f_n and g_n are integrable, $|f_n|\le|f|$ and $|g_n|\le|g|$ almost everywhere, $f=_{\text{a.e.}}\lim_{n\to\infty}f_n$ and $g=_{\text{a.e.}}\lim_{n\to\infty}g_n$. Consequently, by Lebesgue's Dominated Convergence Theorem,

$$\begin{aligned}(f*g)(x)&=\int f(x-y)g(y)dy=\int\lim_{n\to\infty}f_n(x-y)g_n(y)dy\\&=\lim_{n\to\infty}\int f_n(x-y)g_n(y)dy=\lim_{n\to\infty}(f_n*g_n)(x)\end{aligned}$$

for every $x\in\operatorname{dom}f*g$. But f_n*g_n is integrable, therefore measurable, for every n, so that $f*g$ must be measurable.

As for the domain of $f*g$,

$$\begin{aligned}x\in\operatorname{dom}(f*g)&\iff\int f(x-y)g(y)dy\text{ is defined in }\mathbb{C}\\&\iff\int|f(x-y)||g(y)|dy\text{ is defined in }\mathbb{R}\\&\iff\int|f_n(x-y)||g_n(y)|dy\text{ is defined in }\mathbb{R}\text{ for every }n\\&\qquad\text{and }\sup_{n\in\mathbb{N}}\int|f_n(x-y)||g_n(y)|dy<\infty.\end{aligned}$$

Because every $|f_n|*|g_n|$ is integrable, therefore measurable and with measurable domain,

$$\operatorname{dom}(f*g)=\{x:x\in\textstyle\bigcap_{n\in\mathbb{N}}\operatorname{dom}(|f_n|*|g_n|),\ \sup_{n\in\mathbb{N}}(|f_n|*|g_n|)(x)<\infty\}$$

is measurable.

255J Theorem Let f, g and h be complex-valued measurable functions, defined almost everywhere in $\mathbb{R}$, such that $f*g$ and $g*h$ are defined a.e. Suppose that $x\in\mathbb{R}$ is such that one of $(|f|*(|g|*|h|))(x)$, $((|f|*|g|)*|h|)(x)$ is defined in $\mathbb{R}$. Then $f*(g*h)$ and $(f*g)*h$ are defined and equal at x.

proof Set $k(y)=f(x-y)$ when this is defined, so that k is measurable and defined almost everywhere (255D).

(a) If $(|f|*(|g|*|h|))(x)$ is defined, this is $\int|k(y)|(|g|*|h|)(y)dy$, so by 255G we have

$$\int k(y)(g*h)(y)dy=\int k(y+z)g(y)h(z)d(y,z),$$

that is,

$$\begin{aligned}(f*(g*h))(x) &= \int f(x-y)(g*h)(y)dy = \int k(y)(g*h)(y)dy \\ &= \int k(y+z)g(y)h(z)d(y,z) = \iint k(y+z)g(y)h(z)dydz \\ &= \iint f(x-y-z)g(y)h(z)dydz = \int (f*g)(x-z)h(z)dz \\ &= ((f*g)*h)(x).\end{aligned}$$

(b) If $((|f|*|g|)*|h|)(x)$ is defined, this is

$$\begin{aligned}\int (|f|*|g|)(x-z)|h(z)|dz &= \iint |f(x-z-y)||g(y)||h(z)|dydz \\ &= \iint |k(y+z)||g(y)||h(z)|dydz.\end{aligned}$$

By 255D again, $(y,z)\mapsto k(y+z)$ is measurable, so we can apply Tonelli's theorem to see that $\int k(y+z)g(y)h(z)d(y,z)$ is defined, and is equal to $\int k(y)(g*h)(y)dy = (f*(g*h))(x)$ by 255Ga. On the other side, by the last two lines of the proof of (a), $\int k(y+z)g(y)h(z)d(y,z)$ is also equal to $((f*g)*h)(x)$.

255K I do not think we shall need an exhaustive discussion of the question of just when $(f*g)(x)$ is defined; this seems to be complicated. However there is a fundamental case in which we can be sure that $(f*g)(x)$ is defined everywhere.

Proposition Suppose that f, g are measurable complex-valued functions defined almost everywhere in $\mathbb{R}$, and that $f\in\mathcal{L}^p_{\mathbb{C}}$, $g\in\mathcal{L}^q_{\mathbb{C}}$ where p, $q\in[1,\infty]$ and $\frac{1}{p}+\frac{1}{q}=1$ (writing $\frac{1}{\infty}=0$ as usual). Then $f*g$ is defined everywhere in $\mathbb{R}$, is uniformly continuous, and

$$\begin{aligned}\sup_{x\in\mathbb{R}}|(f*g)(x)| &\le \|f\|_p\|g\|_q \text{ if } 1<p<\infty,\ 1<q<\infty, \\ &\le \|f\|_1 \operatorname{ess\,sup}|g| \text{ if } p=1,\ q=\infty, \\ &\le \operatorname{ess\,sup}|f|\cdot\|g\|_1 \text{ if } p=\infty,\ q=1.\end{aligned}$$

proof (a) (For an introduction to $\mathcal{L}^p$ spaces, see §244.) For any $x\in\mathbb{R}$, the function f_x, defined by setting $f_x(y)=f(x-y)$ whenever $x-y\in\operatorname{dom}f$, must also belong to $\mathcal{L}^p$, because $f_x=f\phi$ for an automorphism ϕ of the measure space. Consequently $(f*g)(x)=\int f_x\times g$ is defined, and of modulus at most $\|f\|_p\|g\|_q$ or $\|f\|_1\operatorname{ess\,sup}|g|$ or $\operatorname{ess\,sup}|f|\cdot\|g\|_1$, by 244Eb/244Pb and 243Fa/243K.

(b) To see that $f*g$ is uniformly continuous, argue as follows. Suppose first that $p<\infty$. Let $\epsilon>0$. Let $\eta>0$ be such that $(2+2^{1/p})\|g\|_q\eta\le\epsilon$. Then there is a bounded continuous function $h:\mathbb{R}\to\mathbb{C}$ such that $\{x:h(x)\ne0\}$ is bounded and $\|f-h\|_p\le\eta$ (244Hb/244Pb); let $M\ge1$ be such that $h(x)=0$ whenever $|x|\ge M-1$. Next, h is uniformly continuous, so there is a $\delta\in\,]0,1]$ such that $|h(x)-h(x')|\le M^{-1/p}\eta$ whenever $|x-x'|\le\delta$.

Suppose that $|x-x'|\le\delta$. Defining $h_x(y)=h(x-y)$, as before, we have

$$\int|h_x-h_{x'}|^p=\int|h(x-y)-h(x'-y)|^p dy=\int|h(t)-h(x'-x+t)|^p dt$$

(substituting $t=x-y$)

$$=\int_{-M}^{M}|h(t)-h(x'-x+t)|^p dt$$

(because $h(t)=h(x'-x+t)=0$ if $|t|\ge M$)

$$\le 2M(M^{-1/p}\eta)^p$$

(because $|h(t)-h(x'-x+t)|\le M^{-1/p}\eta$ for every t)

$$= 2\eta^p.$$

So $\|h_x - h_{x'}\|_p \le 2^{1/p}\eta$. On the other hand,

$$\int |h_x - f_x|^p = \int |h(x-y) - f(x-y)|^p dy = \int |h(y) - f(y)|^p dy,$$

so $\|h_x - f_x\|_p = \|h - f\|_p \le \eta$, and similarly $\|h_{x'} - f_{x'}\|_p \le \eta$. So

$$\|f_x - f_{x'}\|_p \le \|f_x - h_x\|_p + |h_x - h_{x'}\|_p + \|h_{x'} - f_{x'}\|_p \le (2 + 2^{1/p})\eta.$$

This means that

$$|(f*g)(x) - (f*g)(x')| = |\int f_x \times g - \int f_{x'} \times g| = |\int (f_x - f_{x'}) \times g|$$
$$\le \|f_x - f_{x'}|_p \|g\|_q \le (2 + 2^{1/p})\|g\|_q \eta \le \epsilon.$$

As ϵ is arbitrary, $f*g$ is uniformly continuous.

The argument here supposes that p is finite. But if $p = \infty$ then $q = 1$ is finite, so we can apply the method with g in place of f to show that $g*f$ is uniformly continuous, and $f*g = g*f$ by 255Fb.

255L The r-dimensional case I have written 255A-255K out as theorems about Lebesgue measure on $\mathbb{R}$. However they all apply equally well to Lebesgue measure on $\mathbb{R}^r$ for any $r \ge 1$, and the modifications required are so small that I think I need do no more than ask you to read through the arguments again, turning every $\mathbb{R}$ into an $\mathbb{R}^r$, and every $\mathbb{R}^2$ into an $(\mathbb{R}^r)^2$. In 255A and elsewhere, the measure μ_2 should be read either as Lebesgue measure on $\mathbb{R}^{2r}$ or as the product measure on $(\mathbb{R}^r)^2$; by 251N the two may be identified. There is a trivial modification required in part (b) of the proof; if $I_n = [a_n, b_n[$ then

$$\mu I_n = \mu(-I_n) = \prod_{i=1}^r \max(0, \beta_{ni} - \alpha_{ni}),$$

writing $a_n = (\alpha_{n1}, \dots, \alpha_{nr})$. In the proof of 255I, the functions f_n should be defined by saying that $f_n(x) = f(x)$ if $|f(x)| \le n$ and $\|x\| \le n$, 0 otherwise.

In quoting these results, therefore, I shall be uninhibited in referring to the paragraphs 255A-255K as if they were actually written out for general $r \ge 1$.

255M The case of $]-\pi, \pi]$ The same ideas also apply to the circle group S^1 and to the interval $]-\pi, \pi]$, but here perhaps rather more explanation is in order.

(a) The first thing to establish is the appropriate group operation. If we think of S^1 as the set $\{z : z \in \mathbb{C}, |z| = 1\}$, then the group operation is complex multiplication, and in the formulae above $x + y$ must be rendered as xy, while $x - y$ must be rendered as xy^{-1}. On the interval $]-\pi, \pi]$, the group operation is $+_{2\pi}$, where for $x, y \in]-\pi, \pi]$ I write $x +_{2\pi} y$ for whichever of $x + y$, $x + y + 2\pi$, $x + y - 2\pi$ belongs to $]-\pi, \pi]$. To see that this is indeed a group operation, one method is to note that it corresponds to multiplication on S^1 if we use the canonical bijection $x \mapsto e^{ix} :]-\pi, \pi] \to S^1$; another, to note that it corresponds to the operation on the quotient group $\mathbb{R}/2\pi\mathbb{Z}$. Thus in this interpretation of the ideas of 255A-255K, we shall wish to replace $x + y$ by $x +_{2\pi} y$, $-x$ by $-_{2\pi}x$, and $x - y$ by $x -_{2\pi} y$, where

$$-_{2\pi}x = -x \text{ if } x \in]-\pi, \pi[, \quad -_{2\pi}\pi = \pi,$$

and $x -_{2\pi} y$ is whichever of $x - y$, $x - y + 2\pi$, $x - y - 2\pi$ belongs to $]-\pi, \pi]$.

(b) As for the measure, the measure to use on $]-\pi, \pi]$ is just Lebesgue measure. Note that because $]-\pi, \pi]$ is Lebesgue measurable, there will be no confusion concerning the meaning of 'measurable subset', as the relatively measurable subsets of $]-\pi, \pi]$ are actually measured by Lebesgue measure on $\mathbb{R}$. Also we can identify the product measure on $]-\pi, \pi] \times]-\pi, \pi]$ with the subspace measure induced by Lebesgue measure on $\mathbb{R}^2$ (251R).

On S^1, we need the corresponding measure induced by the canonical bijection between S^1 and $]-\pi, \pi]$, which indeed is often called 'Lebesgue measure on S^1'. (We shall see in 265E that it is also equal to Hausdorff one-dimensional measure on S^1.) We are very close to the level at which it would become reasonable to move to S^1 and this measure (or its normalized version, in which it is reduced by a factor of 2π, so as to make S^1 a probability space). However, the elementary theory of Fourier series, which will be the principal application of this work in the present volume, is generally done on intervals in $\mathbb{R}$, so that formulae based on $]-\pi,\pi]$ are closer to the standard expressions. Henceforth, therefore, I will express the work in terms of $]-\pi,\pi]$.

(c) The result corresponding to 255A now takes a slightly different form, so I spell it out.

255N Theorem Let μ be Lebesgue measure on $]-\pi,\pi]$ and μ_2 Lebesgue measure on $]-\pi,\pi]\times]-\pi,\pi]$; write Σ, Σ_2 for their domains.

(a) For any $a\in\,]-\pi,\pi]$, the map $x\mapsto a+_{2\pi}x:\,]-\pi,\pi]\to\,]-\pi,\pi]$ is a measure space automorphism of $(]-\pi,\pi],\Sigma,\mu)$.

(b) The map $x\mapsto -_{2\pi}x:\,]-\pi,\pi]\to\,]-\pi,\pi]$ is a measure space automorphism of $(]-\pi,\pi],\Sigma,\mu)$.

(c) For any $a\in\,]-\pi,\pi]$, the map $x\mapsto a-_{2\pi}x:\,]-\pi,\pi]\to\,]-\pi,\pi]$ is a measure space automorphism of $(]-\pi,\pi],\Sigma,\mu)$.

(d) The map $(x,y)\mapsto(x+_{2\pi}y,y):\,]-\pi,\pi]^2\to\,]-\pi,\pi]^2$ is a measure space automorphism of $(]-\pi,\pi]^2$, $\Sigma_2,\mu_2)$.

(e) The map $(x,y)\mapsto(x-_{2\pi}y,y):\,]-\pi,\pi]^2\to\,]-\pi,\pi]^2$ is a measure space automorphism of $(]-\pi,\pi]^2$, $\Sigma_2,\mu_2)$.

proof (a) Set $\phi(x)=a+_{2\pi}x$. Then for any $E\subseteq\,]-\pi,\pi]$,

$$\phi[E]=((E+a)\cap\,]-\pi,\pi])\cup(((E+a)\cap\,]\pi,3\pi])-2\pi)\cup(((E+a)\cap\,]-3\pi,-\pi])+2\pi),$$

and these three sets are disjoint, so that

$$\begin{aligned}\mu\phi[E]&=\mu((E+a)\cap\,]-\pi,\pi])+\mu(((E+a)\cap\,]\pi,3\pi])-2\pi)\\&\qquad+\mu(((E+a)\cap\,]-3\pi,-\pi])+2\pi)\\&=\mu_L((E+a)\cap\,]-\pi,\pi])+\mu_L(((E+a)\cap\,]\pi,3\pi])-2\pi)\\&\qquad+\mu_L(((E+a)\cap\,]-3\pi,-\pi])+2\pi)\end{aligned}$$

(writing μ_L for Lebesgue measure on $\mathbb{R}$)

$$\begin{aligned}&=\mu_L((E+a)\cap\,]-\pi,\pi])+\mu_L((E+a)\cap\,]\pi,3\pi])+\mu_L((E+a)\cap\,]-3\pi,-\pi])\\&=\mu_L(E+a)=\mu_LE=\mu E.\end{aligned}$$

Similarly, $\mu\phi^{-1}[E]$ is defined and equal to μE for every $E\in\Sigma$, so that ϕ is an automorphism of $(]-\pi,\pi]$, $\Sigma,\mu)$.

(b) Of course this is quicker. Setting $\phi(x)=-_{2\pi}x$ for $x\in\,]-\pi,\pi]$, we have

$$\begin{aligned}\mu(\phi[E])&=\mu(\phi[E]\cap\,]-\pi,\pi[)=\mu(-(E\cap\,]-\pi,\pi[)\\&=\mu_L(-(E\cap\,]-\pi,\pi[))=\mu_L(E\cap\,]-\pi,\pi[)\\&=\mu(E\cap\,]-\pi,\pi[)=\mu E\end{aligned}$$

for every $E\in\Sigma$.

(c) This is just a matter of putting (a) and (b) together, as in 255A.

(d) We can argue as in (a), but with a little more elaboration. If $E\in\Sigma_2$, and $\phi(x,y)=(x+_{2\pi}y,y)$ for x, $y\in\,]-\pi,\pi]$, set $\psi(x,y)=(x+y,y)$ for x, $y\in\mathbb{R}$, and write $c=(2\pi,0)\in\mathbb{R}^2$, $H=\,]-\pi,\pi]^2$, $H'=H+c$, $H''=H-c$. Then for any $E\in\Sigma_2$,

$$\phi[E]=(\psi[E]\cap H)\cup((\psi[E]\cap H')-c)\cup((\psi[E]\cap H'')+c),$$

so

$$\begin{aligned}\mu_2\phi[E]&=\mu_2(\psi[E]\cap H)+\mu_2((\psi[E]\cap H')-c)+\mu_2((\psi[E]\cap H'')+c)\\&=\mu_L(\psi[E]\cap H)+\mu_L((\psi[E]\cap H')-c)+\mu_L((\psi[E]\cap H'')+c)\end{aligned}$$

(this time writing μ_L for Lebesgue measure on $\mathbb{R}^2$)

$$= \mu_L(\psi[E] \cap H) + \mu_L(\psi[E] \cap H') + \mu_L(\psi[E] \cap H'')$$
$$= \mu_L \psi[E] = \mu_L E = \mu_2 E.$$

In the same way, $\mu_2(\phi^{-1}[E]) = \mu_2 E$ for every $E \in \Sigma_2$, so ϕ is an automorphism of $(]-\pi,\pi]^2, \Sigma_2, \mu_2)$, as required.

(e) Finally, (e) is just a restatement of (d), as before.

255O Convolutions on $]-\pi,\pi]$ With the fundamental result established, the same arguments as in 255B-255K now yield the following. Write μ for Lebesgue measure on $]-\pi,\pi]$.

(a) Let f and g be measurable complex-valued functions defined almost everywhere in $]-\pi,\pi]$. Write $f*g$ for the function defined by the formula

$$(f*g)(x) = \int_{-\pi}^{\pi} f(x -_{2\pi} y)g(y)dy$$

whenever $x \in]-\pi,\pi]$ and the integral exists as a complex number. Then $f*g$ is the **convolution** of the functions f and g.

(b) If f and g are measurable complex-valued functions defined almost everywhere in $]-\pi,\pi]$, then $f*g = g*f$.

(c) Let f, g and h be measurable complex-valued functions defined almost everywhere in $]-\pi,\pi]$. Then (i)

$$\int_{-\pi}^{\pi} h(x)(f*g)(x)dx = \int_{]-\pi,\pi]^2} h(x +_{2\pi} y)f(x)g(y)d(x,y)$$

whenever the right-hand side exists and is finite, provided that in the expression $h(x)(f*g)(x)$ we interpret the product as 0 if $h(x) = 0$ and $(f*g)(x)$ is undefined.

(ii) If, on the same interpretation of $|h(x)|(|f|*|g|)(x)$, the integral $\int_{-\pi}^{\pi} |h(x)|(|f|*|g|)(x)dx$ is finite, then $\int_{]-\pi,\pi]^2} h(x +_{2\pi} y)f(x)g(y)d(x,y)$ exists in $\mathbb{C}$, so again we shall have

$$\int_{-\pi}^{\pi} h(x)(f*g)(x)dx = \int_{]-\pi,\pi]^2} h(x +_{2\pi} y)f(x)g(y)d(x,y).$$

(d) If f, g are complex-valued functions which are integrable over $]-\pi,\pi]$, then $f*g$ is integrable, with

$$\int_{-\pi}^{\pi} f*g = \int_{-\pi}^{\pi} f \int_{-\pi}^{\pi} g, \quad \int_{-\pi}^{\pi} |f*g| \le \int_{-\pi}^{\pi} |f| \int_{-\pi}^{\pi} |g|.$$

(e) Let f, g, h be complex-valued measurable functions defined almost everywhere in $]-\pi,\pi]$, such that $f*g$ and $g*h$ are also defined almost everywhere. Suppose that $x \in]-\pi,\pi]$ is such that one of $(|f|*(|g|*|h|))(x)$, $((|f|*|g|)*|h|)(x)$ is defined in $\mathbb{R}$. Then $f*(g*h)$ and $(f*g)*h$ are defined and equal at x.

(f) Suppose that $f \in \mathcal{L}^p_{\mathbb{C}}(\mu)$, $g \in \mathcal{L}^q_{\mathbb{C}}(\mu)$ where p, $q \in [1,\infty]$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then $f*g$ is defined everywhere in $]-\pi,\pi]$, and $\sup_{x\in]-\pi,\pi]} |(f*g)(x)| \le \|f\|_p \|g\|_q$, interpreting $\| \ \|_\infty$ as ess sup$| \ |$, as in 255K.

255X Basic exercises >(a) Let f, g be complex-valued functions defined almost everywhere in $\mathbb{R}$. Show that for any $x \in \mathbb{R}$, $(f*g)(x) = \int f(x+y)g(-y)dy$ if either is defined.

>(b) Let f and g be complex-valued functions defined almost everywhere in $\mathbb{R}$. (i) Show that if f and g are even functions, so is $f*g$. (ii) Show that if f is even and g is odd then $f*g$ is odd. (iii) Show that if f and g are odd then $f*g$ is even.

(c) Suppose that f, g are real-valued measurable functions defined almost everywhere in $\mathbb{R}^r$ and such that $f > 0$ a.e., $g \ge 0$ a.e. and $\{x : g(x) > 0\}$ is not negligible. Show that $f*g > 0$ everywhere in dom$(f*g)$.

>(d) Suppose that $f : \mathbb{R} \to \mathbb{C}$ is a bounded differentiable function and that f' is bounded. Show that for any integrable complex-valued function g on $\mathbb{R}$, $f * g$ is differentiable and $(f * g)' = f' * g$ everywhere. (*Hint*: 123D.)

(e) A complex-valued function g defined almost everywhere in $\mathbb{R}$ is **locally integrable** if $\int_a^b g$ is defined in $\mathbb{C}$ whenever $a < b$ in $\mathbb{R}$. Suppose that g is such a function and that $f : \mathbb{R} \to \mathbb{C}$ is a differentiable function, with continuous derivative, such that $\{x : f(x) \neq 0\}$ is bounded. Show that $(f * g)' = f' * g$ everywhere.

>(f) Set $\phi_\delta(x) = \exp(-\frac{1}{\delta^2 - x^2})$ if $|x| < \delta$, 0 if $|x| \geq \delta$, as in 242Xi. Set $\alpha_\delta = \int \phi_\delta$, $\psi_\delta = \alpha_\delta^{-1}\phi_\delta$. Let f be a locally integrable complex-valued function on $\mathbb{R}$. (i) Show that $f * \psi_\delta$ is a smooth function defined everywhere on $\mathbb{R}$ for every $\delta > 0$. (ii) Show that $\lim_{\delta \downarrow 0}(f * \psi_\delta)(x) = f(x)$ for almost every $x \in \mathbb{R}$. (*Hint*: 223Yg.) (iii) Show that if f is integrable then $\lim_{\delta \downarrow 0} \int |f - f * \psi_\delta| = 0$. (*Hint*: use (ii) and 245H(a-ii) *or* look first at the case $f = \chi[a, b]$ and use 242O, noting that $\int |f * \psi_\delta| \leq \int |f|$.) (iv) Show that if f is uniformly continuous and defined everywhere in $\mathbb{R}$ then $\lim_{\delta \downarrow 0} \sup_{x \in \mathbb{R}} |f(x) - (f * \psi_\delta)(x)| = 0$.

>(g) For $\alpha > 0$, set $g_\alpha(t) = \dfrac{1}{\Gamma(\alpha)} t^{\alpha - 1}$ for $t > 0$, 0 for $t \leq 0$. Show that $g_\alpha * g_\beta = g_{\alpha+\beta}$ for all α, $\beta > 0$. (*Hint*: 252Yf.)

>(h) Let μ be Lebesgue measure on $\mathbb{R}$. For u, v, $w \in L^0_{\mathbb{C}} = L^0_{\mathbb{C}}(\mu)$, say that $u * v = w$ if $f * g$ is defined almost everywhere and $(f * g)^\bullet = w$ whenever f, $g \in \mathcal{L}^0_{\mathbb{C}}(\mu)$, $f^\bullet = u$ and $g^\bullet = w$. (i) Show that $(u_1 + u_2) * v = u_1 * v + u_2 * v$ whenever u_1, u_2, $v \in L^0_{\mathbb{C}}$ and $u_1 * v$ and $u_2 * v$ are defined in this sense. (ii) Show that $u * v = v * u$ whenever u, $v \in L^0(\mathbb{C})$ and either $u * v$ or $v * u$ is defined. (iii) Show that if u, v, $w \in L^0_{\mathbb{C}}$, $u * v$ and $v * w$ are defined, and either $|u| * (|v| * |w|)$ or $(|u| * |v|) * |w|$ is defined, then then $u * (v * w) = (u * v) * w$ are defined and equal.

>(i) Let μ be Lebesgue measure on $\mathbb{R}$. (i) Show that $u * v$, as defined in 255Xh, belongs to $L^1_{\mathbb{C}}(\mu)$ whenever u, $v \in L^1_{\mathbb{C}}(\mu)$. (ii) Show that $L^1_{\mathbb{C}}$ is a commutative Banach algebra under $*$ (definition: 2A4J).

(j)(i) Show that if h is an integrable function on $\mathbb{R}^2$, then $(Th)(x) = \int h(x-y, y)dy$ exists for almost every $x \in \mathbb{R}$, and that $\int (Th)(x)dx = \int h(x, y)d(x, y)$. (ii) Write μ_2 for Lebesgue measure on $\mathbb{R}^2$, μ for Lebesgue measure on $\mathbb{R}$. Show that there is a linear operator $\tilde{T} : L^1(\mu_2) \to L^1(\mu)$ defined by setting $\tilde{T}(h^\bullet) = (Th)^\bullet$ for every integrable function h on $\mathbb{R}^2$. (iii) Show that in the language of 253E and 255Xh, $\tilde{T}(u \otimes v) = u * v$ for all u, $v \in L^1(\mu)$.

>(k) For $\boldsymbol{a}$, $\boldsymbol{b} \in \mathbb{C}^{\mathbb{Z}}$ set $(\boldsymbol{a} * \boldsymbol{b})(n) = \sum_{i \in \mathbb{Z}} \boldsymbol{a}(n - i)\boldsymbol{b}(i)$ whenever $\sum_{i \in \mathbb{Z}} |\boldsymbol{a}(n - i)\boldsymbol{b}(i)| < \infty$. Show that
(i) $\boldsymbol{a} * \boldsymbol{b} = \boldsymbol{b} * \boldsymbol{a}$;
(ii) $\sum_{i \in \mathbb{Z}} \boldsymbol{c}(i)(\boldsymbol{a} * \boldsymbol{b})(i) = \sum_{i,j \in \mathbb{Z}} \boldsymbol{c}(i + j)\boldsymbol{a}(i)\boldsymbol{b}(j)$ if $\sum_{i,j \in \mathbb{Z}} |\boldsymbol{c}(i + j)\boldsymbol{a}(i)\boldsymbol{b}(j)| < \infty$;
(iii) if $\boldsymbol{a}$, $\boldsymbol{b} \in \ell^1(\mathbb{Z})$ then $\boldsymbol{a} * \boldsymbol{b} \in \ell^1(\mathbb{Z})$ and $\|\boldsymbol{a} * \boldsymbol{b}\|_1 \leq \|\boldsymbol{a}\|_1 \|\boldsymbol{b}\|_1$;
(iv) If $\boldsymbol{a}$, $\boldsymbol{b} \in \ell^2(\mathbb{Z})$ then $\boldsymbol{a} * \boldsymbol{b} \in \ell^\infty(\mathbb{Z})$ and $\|\boldsymbol{a} * \boldsymbol{b}\|_\infty \leq \|\boldsymbol{a}\|_2 \|\boldsymbol{b}\|_2$;
(v) if $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c} \in \mathbb{C}^{\mathbb{Z}}$ and $(|\boldsymbol{a}| * (|\boldsymbol{b}| * |\boldsymbol{c}|))(n)$ is well-defined, then $(\boldsymbol{a} * (\boldsymbol{b} * \boldsymbol{c}))(n) = ((\boldsymbol{a} * \boldsymbol{b}) * \boldsymbol{c})(n)$.

255Y Further exercises (a) Let f be a complex-valued function which is integrable over $\mathbb{R}$. (i) Let x be any point of the Lebesgue set of f. Show that for any $\epsilon > 0$ there is a $\delta > 0$ such that $|f(x) - (f * g)(x)| \leq \epsilon$ whenever $g : \mathbb{R} \to [0, \infty[$ is a function which is non-decreasing on $]-\infty, 0]$, non-decreasing on $[0, \infty[$, and has $\int g = 1$ and $\int_{-\delta}^{\delta} g \geq 1 - \delta$. (ii) Show that for any $\epsilon > 0$ there is a $\delta > 0$ such that $\|f - f * g\|_1 \leq \epsilon$ whenever $g : \mathbb{R} \to [0, \infty[$ is a function which is non-decreasing on $]-\infty, 0]$, non-decreasing on $[0, \infty[$, and has $\int g = 1$ and $\int_{-\delta}^{\delta} g \geq 1 - \delta$.

(b) Let f be a complex-valued function which is integrable over $\mathbb{R}$. Show that, for almost every $x \in \mathbb{R}$,

$$\lim_{a \to \infty} \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{f(y)}{1 + a^2(x-y)^2} dy, \quad \lim_{a \to \infty} \frac{1}{a} \int_x^{\infty} f(y) e^{-a(y-x)} dy,$$

$$\lim_{\sigma \downarrow 0} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} f(y) e^{-(y-x)^2/2\sigma^2} dy$$

all exist and are equal to $f(x)$. (*Hint*: 263G.)

(c) Set $f(x)=1$ for all $x\in\mathbb{R}$, $g(x)=\frac{x}{|x|}$ for $0<|x|\le 1$ and 0 otherwise, $h(x)=\tanh x$ for all $x\in\mathbb{R}$. Show that $f*(g*h)$ and $(f*g)*h$ are both defined (and constant) everywhere, and are different.

(d) Discuss what can happen if, in the context of 255J, we know that $(|f|*(|g|*|h|))(x)$ is defined, but have no information on the domain of $f*g$.

(e) Suppose that $p\in[1,\infty[$ and that $f\in\mathcal{L}^p_{\mathbb{C}}(\mu)$, where μ is Lebesgue measure on $\mathbb{R}^r$. For $a\in\mathbb{R}^r$ set $(S_af)(x)=f(a+x)$ whenever $a+x\in\operatorname{dom} f$. Show that $S_af\in\mathcal{L}^p_{\mathbb{C}}(\mu)$, and that for every $\epsilon>0$ there is a $\delta>0$ such that $\|S_af-f\|_p\le\epsilon$ whenever $|a|\le\delta$.

(f) Suppose that $p,\,q\in\,]1,\infty[$ and $\frac{1}{p}+\frac{1}{q}=1$. Take $f\in\mathcal{L}^p_{\mathbb{C}}(\mu)$ and $g\in\mathcal{L}^q_{\mathbb{C}}(\mu)$, where μ is Lebesgue measure on $\mathbb{R}^r$. Show that $\lim_{\|x\|\to\infty}(f*g)(x)=0$. (*Hint*: use 244Hb.)

(g) Repeat 255Ye and 255K, this time taking μ to be Lebesgue measure on $]-\pi,\pi]$, and setting $(S_af)(x)=f(a+_{2\pi}x)$ for $a\in\,]-\pi,\pi]$; show that in the new version of 255K, $(f*g)(\pi)=\lim_{x\downarrow-\pi}(f*g)(x)$.

(h) Let μ be Lebesgue measure on $\mathbb{R}$. For $a\in\mathbb{R}$, $f\in\mathcal{L}^0=\mathcal{L}^0(\mu)$ set $(S_af)(x)=f(a+x)$ whenever $a+x\in\operatorname{dom} f$.

(i) Show that $S_af\in\mathcal{L}^0$ for every $f\in\mathcal{L}^0$.

(ii) Show that we have a map $\tilde S_a:L^0\to L^0$ defined by setting $\tilde S_a(f^\bullet)=(S_af)^\bullet$ for every $f\in\mathcal{L}^0$.

(iii) Show that $\tilde S_a$ is a Riesz space isomorphism and is a homeomorphism for the topology of convergence in measure; moreover, that $\tilde S_a(u\times v)=\tilde S_au\times\tilde S_av$ for all $u,\,v\in L^0$.

(iv) Show that $\tilde S_{a+b}=\tilde S_a\tilde S_b$ for all $a,\,b\in\mathbb{R}$.

(v) Show that $\lim_{a\to0}\tilde S_au=u$ for the topology of convergence in measure, for every $u\in L^0$.

(vi) Show that if $1\le p\le\infty$ then $\tilde S_a\restriction L^p$ is an isometric isomorphism of the Banach lattice L^p.

(vii) Show that if $p\in[1,\infty[$ then $\lim_{a\to0}\|\tilde S_au-u\|_p=0$ for every $u\in L^p$.

(viii) Show that if $A\subseteq L^1$ is uniformly integrable and $M\ge0$, then $\{\tilde S_au:u\in A,\,|a|\le M\}$ is uniformly integrable.

(ix) Suppose that $u,\,v\in L^0$ are such that $u*v$ is defined in L^0 in the sense of 255Xh. Show that $\tilde S_a(u*v)=(\tilde S_au)*v=u*(\tilde S_av)$ for every $a\in\mathbb{R}$.

(i) Prove 255Nd from 255Na by the method used to prove 255Ad from 255Aa, rather than by quoting 255Ad.

(j) Let μ be Lebesgue measure on $\mathbb{R}$, and $\phi:\mathbb{R}\to\mathbb{R}$ a convex function; let $\bar\phi:L^0\to L^0=L^0(\mu)$ be the associated operator (see 241I). Show that if $u\in L^1=L^1(\mu)$, $v\in L^0$ are such that $u\ge0$, $\int u=1$ and $u*v$, $u*\bar\phi(v)$ are both defined in the sense of 255Xh, then $\bar\phi(u*v)\le u*\bar\phi(v)$. (*Hint*: 233I.)

(k) Let μ be Lebesgue measure on $\mathbb{R}$, and $p\in[1,\infty]$. Let $f\in\mathcal{L}^1_{\mathbb{C}}(\mu)$, $g\in\mathcal{L}^p_{\mathbb{C}}(\mu)$. Show that $f*g\in\mathcal{L}^p_{\mathbb{C}}(\mu)$ and that $\|f*g\|_p\le\|f\|_1\|g\|_p$. (*Hint*: argue from 255Yj, as in 244M.)

(l) Suppose that $p,\,q,\,r\in\,]1,\infty[$ and that $\frac{1}{p}+\frac{1}{q}=1+\frac{1}{r}$. Let μ be Lebesgue measure on $\mathbb{R}$. (i) Show that

$$\int f\times g\le\|f\|_p^{1-p/r}\|g\|_q^{1-q/r}\Bigl(\int f^p\times g^q\Bigr)^{1/r}$$

whenever $f,\,g\ge0$ and $f\in\mathcal{L}^p(\mu)$, $g\in\mathcal{L}^q(\mu)$. (*Hint*: set $p'=p/(p-1)$, etc.; $f_1=f^{p/q'}$, $g_1=g^{q/p'}$, $h=(f^p\times g^q)^{1/r}$. Use 244Xc to see that $\|f_1\times g_1\|_{r'}\le\|f_1\|_{q'}\|g_1\|_{p'}$, so that $\int f_1\times g_1\times h\le\|f_1\|_{q'}\|g_1\|_{p'}\|h\|_r$.) (ii) Show that $f*g$ is defined a.e. and that $\|f*g\|_r\le\|f\|_p\|g\|_q$ for all $f\in\mathcal{L}^p(\mu)$, $g\in\mathcal{L}^q(\mu)$. (*Hint*: take $f,\,g\ge0$. Use (i) to see that $(f*g)(x)^r\le\|f\|_p^{r-p}\|g\|_q^{r-q}\int f(y)^pg(x-y)^qdy$, so that $\|f*g\|_r^r\le\|f\|_p^{r-p}\|g\|_q^{r-q}\int f(y)^p\|g\|_q^qdy$.) (This is **Young's inequality**.)

(m) Repeat the results of this section for the group $(S^1)^r$, where $r\ge2$, given its product measure.

(n) Let G be a group and μ a σ-finite measure on G such that (α) for every $a\in G$, the map $x\mapsto ax$ is an automorphism of (G,μ) (β) the map $(x,y)\mapsto(x,xy)$ is an automorphism of (G^2,μ_2), where μ_2 is the c.l.d.

product measure on $G\times G$. For f, $g\in\mathcal{L}^0_{\mathbb{C}}(\mu)$ write $(f*g)(x)=\int f(y)g(y^{-1}x)dy$ whenever this is defined. Show that

(i) if f, g, $h\in\mathcal{L}^0_{\mathbb{C}}(\mu)$ and $\int h(xy)f(x)g(y)d(x,y)$ is defined in $\mathbb{C}$, then $\int h(x)(f*g)(x)dx$ exists and is equal to $\int h(xy)f(x)g(y)d(x,y)$, provided that in the expression $h(x)(f*g)(x)$ we interpret the product as 0 if $h(x)=0$ and $(f*g)(x)$ is undefined;

(ii) if f, $g\in\mathcal{L}^1_{\mathbb{C}}(\mu)$ then $f*g\in\mathcal{L}^1_{\mathbb{C}}(\mu)$ and $\int f*g=\int f\int g$, $\|f*g\|_1\le\|f\|_1\|g\|_1$;

(iii) if f, g, $h\in\mathcal{L}^1_{\mathbb{C}}(\mu)$ then $f*(g*h)=(f*g)*h$.

(See Halmos 50, §59.)

(o) Repeat 255Yn for counting measure on any group G.

255 Notes and comments I have tried to set this section out in such a way that it will be clear that the basis of all the work here is 255A, and the crucial application is 255G. I hope that if and when you come to look at general topological groups (for instance, in Chapter 44), you will find it easy to trace through the ideas in any abelian topological group for which you can prove a version of 255A. For non-abelian groups, of course, rather more care is necessary, especially as in some important examples we no longer have $\mu\{x^{-1}:x\in E\}=\mu E$ for every E; see 255Yn-255Yo for a little of what can be done without using topological ideas.

The critical point in 255A is the move from the one-dimensional results in 255Aa-255Ac, which are just the translation- and reflection-invariance of Lebesgue measure, to the two-dimensional results in 255Ac-255Ad. And the living centre of the argument, as I present it, is the fact that the shear transformation ϕ is an automorphism of the structure $(\mathbb{R}^2,\Sigma_2)$. The actual calculation of $\mu_2\phi[E]$, assuming that it is measurable, is an easy application of Fubini's and Tonelli's theorems and the translation-invariance of μ. It is for this step that we absolutely need the topological properties of Lebesgue measure. I should perhaps remind you that the fact that ϕ is a homeomorphism is not sufficient; in 134I I described a homeomorphism of the unit interval which does not preserve measurability, and it is easy to adapt this to produce a homeomorphism $\psi:\mathbb{R}^2\to\mathbb{R}^2$ such that $\psi[E]$ is not always measurable for measurable E. The argument of 255A is dependent on the special relationships between all three of the measure, topology and group structure of $\mathbb{R}$.

I have already indulged in a few remarks on what ought, or ought not, to be 'obvious' (255C). But perhaps I can add that such results as 255B and the later claim, in the proof of 255K, that a reflected version of a function in $\mathcal{L}^p$ is also in $\mathcal{L}^p$, can only be trivial consequences of results like 255A if every step in the construction of the integral is done in the abstract context of general measure spaces. Even though we are here working exclusively with the Lebesgue integral, the argument will become untrustworthy if we have at any stage in the definition of the integral even mentioned that we are thinking of Lebesgue measure. I advance this as a solid reason for defining 'integration' on abstract measure spaces from the beginning, as I did in Volume 1. Indeed, I suggest that generally in pure mathematics there are good reasons for casting arguments into the forms appropriate to the arguments themselves.

I am writing this book for readers who are interested in proofs, and as elsewhere I have written the proofs of this section out in detail. But most of us find it useful to go through some material in 'advanced calculus' mode, by which I mean starting with a formula such as

$$(f*g)(x)=\int f(x-y)g(y)dy,$$

and then working out consequences by formal manipulations, for instance

$$\int h(x)(f*g)(x)dx=\iint h(x)f(x-y)g(y)dydx=\iint h(x+y)f(x)g(y)dydx,$$

without troubling about the precise applicability of the formulae to begin with. In some ways this formula-driven approach can be more truthful to the structure of the subject than the careful analysis I habitually present. The exact hypotheses necessary to make the theorems strictly true are surely secondary, in such contexts as this section, to the pattern formed by the ensemble of the theorems, which can be adequately and elegantly expressed in straightforward formulae. Of course I do still insist that we cannot properly appreciate the structure, nor safely use it, without mastering the ideas of the proofs – and as I have said elsewhere, I believe that mastery of ideas necessarily includes mastery of the formal details, at least in the sense of being able to reconstruct them fairly fluently on demand.

Throughout the main exposition of this section, I have worked with functions rather than equivalence classes of functions. But all the results here have interpretations of great importance for the theory of the 'function spaces' of Chapter 24. In 255Xh and the succeeding exercises, I have pointed to a definition of convolution as an operator from a subset of $L^0 \times L^0$ to L^0. It is an interesting point that if $u, v \in L^0$ then $u*v$ can be interpreted as a *function*, not as a member of L^0 (255Fc). Thus 255H can be regarded as saying that $u*v \in \mathcal{L}^1$ for $u, v \in L^1$. We cannot quite say that convolution is a bilinear operator from $L^1 \times L^1$ to $\mathcal{L}^1$, because $\mathcal{L}^1$, as I define it, is not strictly speaking a linear space. If we want a bilinear operator, then we have to regard convolution as a function from $L^1 \times L^1$ to L^1. But when we look at convolution as a function on $L^2 \times L^2$, for instance, then our functions $u*v$ are defined everywhere (255K), and indeed are continuous functions vanishing at ∞ (255Ye-255Yf). So in this case it seems more appropriate to regard convolution as a bilinear operator from $L^2 \times L^2$ to some space of continuous functions, and not as an operator from $L^2 \times L^2$ to L^∞. For an example of an interesting convolution which is not naturally representable in terms of an operator on L^p spaces, see 255Xg.

Because convolution acts as a continuous bilinear operator from $L^1(\mu) \times L^1(\mu)$ to $L^1(\mu)$, where μ is Lebesgue measure on $\mathbb{R}$, Theorem 253F tells us that it must correspond to a linear operator from $L^1(\mu_2)$ to $L^1(\mu)$, where μ_2 is Lebesgue measure on $\mathbb{R}^2$. This is the operator $\tilde{T}$ of 255Xj.

So far in these notes I have written as though we were concerned only with Lebesgue measure on $\mathbb{R}$. However many applications of the ideas involve $\mathbb{R}^r$ or $]-\pi,\pi]$ or S^1. The move to $\mathbb{R}^r$ should be elementary. The move to S^1 does require a re-formulation of the basic result 255A/255N. It should also be clear that there will be no new difficulties in moving to $]-\pi,\pi]^r$ or $(S^1)^r$. Moreover, we can also go through the whole theory for the groups $\mathbb{Z}$ and $\mathbb{Z}^r$, where the appropriate measure is now counting measure, so that $L^0_{\mathbb{C}}$ becomes identified with $\mathbb{C}^{\mathbb{Z}}$ or $\mathbb{C}^{\mathbb{Z}^r}$ (255Xk, 255Yo).

Version of 6.8.15

256 Radon measures on $\mathbb{R}^r$

In the next section, and again in Chapters 27 and 28, we need to consider the principal class of measures on Euclidean spaces. For a proper discussion of this class, and the interrelationships between the measures and the topologies involved, we must wait until Volume 4. For the moment, therefore, I present definitions adapted to the case in hand, warning you that the correct generalizations are not quite obvious. I give the definition (256A) and a characterization (256C) of Radon measures on Euclidean spaces, and theorems on the construction of Radon measures as indefinite integrals (256E, 256J), as image measures (256G) and as product measures (256K). In passing I give a version of Lusin's theorem concerning measurable functions on Radon measure spaces (256F).

Throughout this section, r and s will be integers greater than or equal to 1.

256A Definitions Let ν be a measure on $\mathbb{R}^r$ and Σ its domain.

(a) ν is a **topological measure** if every open set belongs to Σ. Note that in this case every Borel set, and in particular every closed set, belongs to Σ.

(b) ν is **locally finite** if every bounded set has finite outer measure.

(c) If ν is a topological measure, it is **inner regular with respect to the compact sets** if

$$\nu E = \sup\{\nu K : K \subseteq E \text{ is compact}\}$$

for every $E \in \Sigma$. (Because ν is a topological measure, and compact sets are closed (2A2Ec), νK is defined for every compact set K.)

(d) ν is a **Radon measure** if it is a complete locally finite topological measure which is inner regular with respect to the compact sets.

256B It will be convenient to be able to call on the following elementary facts.

Lemma Let ν be a Radon measure on $\mathbb{R}^r$, and Σ its domain.

(a) ν is σ-finite.

(b) For any $E \in \Sigma$ and any $\epsilon > 0$ there are a closed set $F \subseteq E$ and an open set $G \supseteq E$ such that $\nu(G \setminus F) \le \epsilon$.

(c) For every $E \in \Sigma$ there is a set $H \subseteq E$, expressible as the union of a sequence of compact sets, such that $\nu(E \setminus H) = 0$.

(d) Every continuous real-valued function on $\mathbb{R}^r$ is Σ-measurable.

(e) If $h : \mathbb{R}^r \to \mathbb{R}$ is continuous and has bounded support, then h is ν-integrable.

proof (a) For each $n \in \mathbb{N}$, $B(\mathbf{0}, n) = \{x : \|x\| \le n\}$ is a closed bounded set, therefore Borel. So if ν is a Radon measure on $\mathbb{R}^r$, $\langle B(\mathbf{0}, n)\rangle_{n\in\mathbb{N}}$ is a sequence of sets of finite measurea covering $\mathbb{R}^r$.

(b) Set $E_n = \{x : x \in E, n \le \|x\| < n + 1\}$ for each n. Then $\nu E_n < \infty$, so there is a compact set $K_n \subseteq E_n$ such that $\nu K_n \ge \nu E_n - 2^{-n-2}\epsilon$. Set $F = \bigcup_{n\in\mathbb{N}} K_n$; then

$$\nu(E \setminus F) = \sum_{n=0}^{\infty} \nu(E_n \setminus K_n) \le \frac{1}{2}\epsilon.$$

Also $F \subseteq E$ and F is closed because

$$F \cap B(\mathbf{0}, n) = \bigcup_{i\le n} K_i \cap B(\mathbf{0}, n)$$

is closed for each n.

In the same way, there is a closed set $F' \subseteq \mathbb{R}^r \setminus E$ such that $\nu((\mathbb{R}^r \setminus E) \setminus F') \le \frac{1}{2}\epsilon$. Setting $G = \mathbb{R}^r \setminus F'$, we see that G is open, that $G \supseteq E$ and that $\nu(G \setminus E) \le \frac{1}{2}\epsilon$, so that $\nu(G \setminus F) \le \epsilon$, as required.

(c) By (b), we can choose for each $n \in \mathbb{N}$ a closed set $F_n \subseteq E$ such that $\nu(E \setminus F_n) \le 2^{-n}$. Set $H = \bigcup_{n\in\mathbb{N}} F_n$; then $H \subseteq E$ and $\nu(E \setminus H) = 0$, and also $H = \bigcup_{m,n\in\mathbb{N}} B(\mathbf{0}, m) \cap F_n$ is a countable union of compact sets.

(d) If $h : \mathbb{R}^r \to \mathbb{R}$ is continuous, all the sets $\{x : h(x) > a\}$ are open, so belong to Σ.

(e) By (d), h is measurable. Now we are supposing that there is some $n \in \mathbb{N}$ such that $h(x) = 0$ whenever $x \notin B(\mathbf{0}, n)$. Since $B(\mathbf{0}, n)$ is compact (2A2F), h is bounded on $B(\mathbf{0}, n)$ (2A2G), and we have $|h| \le \gamma\chi B(\mathbf{0}, n)$ for some γ; since $\nu B(\mathbf{0}, n)$ is finite, h is ν-integrable.

256C Theorem A measure ν on $\mathbb{R}^r$ is a Radon measure iff it is the completion of a locally finite measure defined on the σ-algebra $\mathcal{B}$ of Borel subsets of $\mathbb{R}^r$.

proof (a) Suppose first that ν is a Radon measure. Write Σ for its domain.

(i) Set $\nu_0 = \nu\restriction\mathcal{B}$. Then ν_0 is a measure with domain $\mathcal{B}$, and it is locally finite because $\nu_0 B(\mathbf{0}, n) = \nu B(\mathbf{0}, n)$ is finite for every n. Let $\hat{\nu}_0$ be the completion of ν_0 (212C).

(ii) If $\hat{\nu}_0$ measures E, there are E_1, $E_2 \in \mathcal{B}$ such that $E_1 \subseteq E \subseteq E_2$ and $\nu_0(E_2 \setminus E_1) = 0$. Now $E \setminus E_1 \subseteq E_2 \setminus E_1$ must be ν-negligible; as ν is complete, $E \in \Sigma$ and

$$\nu E = \nu E_1 = \nu_0 E_1 = \hat{\nu}_0 E.$$

(iii) If $E \in \Sigma$, then by 256Bc there is a Borel set $H \subseteq E$ such that $\nu(E \setminus H) = 0$. Equally, there is a Borel set $H' \subseteq \mathbb{R}^r \setminus E$ such that $\nu((\mathbb{R}^r \setminus E) \setminus H') = 0$, so that we have $H \subseteq E \subseteq \mathbb{R}^r \setminus H'$ and

$$\nu_0((\mathbb{R}^r \setminus H') \setminus H) = \nu((\mathbb{R}^r \setminus H') \setminus H) = 0.$$

So $\hat{\nu}_0 E$ is defined and equal to $\nu_0 E_1 = \nu E$.

This shows that $\nu = \hat{\nu}_0$ is the completion of the locally finite Borel measure $\nu\restriction\mathcal{B}$. And this is true for any Radon measure ν on $\mathbb{R}^r$.

(b) For the rest of the proof, I suppose that ν_0 is a locally finite measure with domain $\mathcal{B}$ and ν is its completion. Write Σ for the domain of ν. We say that a subset of $\mathbb{R}^r$ is a $\mathbf{K}_\sigma$ **set** if it is expressible as the union of a sequence of compact sets. Note that every K_σ set is a Borel set, so belongs to Σ. Set

$$\mathcal{A} = \{E : E \in \Sigma, \text{ there is a } \mathrm{K}_\sigma \text{ set } H \subseteq E \text{ such that } \nu(E \setminus H) = 0\},$$

$$\Sigma = \{E : E \in \mathcal{A}, \mathbb{R}^r \setminus E \in \mathcal{A}\}.$$

(c)(i) Every open set is itself a K_σ set, so belongs to $\mathcal{A}$. **P** Let $G \subseteq \mathbb{R}^r$ be open. If $G = \emptyset$ then G is compact and the result is trivial. Otherwise, let $\mathcal{I}$ be the set of closed intervals of the form $[q, q']$, where q, $q' \in \mathbb{Q}^r$, which are included in G. Then all the members of $\mathcal{I}$ are closed and bounded, therefore compact. If $x \in G$, there is a $\delta > 0$ such that $B(x, \delta) = \{y : \|y - x\| \le \delta\} \subseteq G$; now there is an $I \in \mathcal{I}$ such that $x \in I \subseteq B(x, \delta)$. Thus $G = \bigcup \mathcal{I}$. But $\mathcal{I}$ is countable, so G is K_σ. **Q**

(ii) Every closed subset of $\mathbb{R}$ is K_σ, so belongs to $\mathcal{A}$. **P** If $F \subseteq \mathbb{R}$ is closed, then $F = \bigcup_{n \in \mathbb{N}} F \cap B(\mathbf{0}, n)$; but every $F \cap B(\mathbf{0}, n)$ is closed and bounded, therefore compact. **Q**

(iii) If $\langle E_n \rangle_{n \in \mathbb{N}}$ is any sequence in $\mathcal{A}$, then $E = \bigcup_{n \in \mathbb{N}} E_n$ belongs to $\mathcal{A}$. **P** For each $n \in \mathbb{N}$ we have a countable family $\mathcal{K}_n$ of compact subsets of E_n such that $\nu(E_n \setminus \bigcup \mathcal{K}_n) = 0$; now $\mathcal{K} = \bigcup_{n \in \mathbb{N}} \mathcal{K}_n$ is a countable family of compact subsets of E, and $E \setminus \bigcup \mathcal{K} \subseteq \bigcup_{n \in \mathbb{N}} (E_n \setminus \bigcup \mathcal{K}_n)$ is ν-negligible. **Q**

(iv) If $\langle E_n \rangle_{n \in \mathbb{N}}$ is any sequence in $\mathcal{A}$, then $F = \bigcap_{n \in \mathbb{N}} E_n \in \mathcal{A}$. **P** For each $n \in \mathbb{N}$, let $\langle K_{ni} \rangle_{i \in \mathbb{N}}$ be a sequence of compact subsets of E_n such that $\nu(E_n \setminus \bigcup_{i \in \mathbb{N}} K_{ni}) = 0$. Set $K'_{nj} = \bigcup_{i \le j} K_{ni}$ for each j, so that

$$\nu(E_n \cap H) = \lim_{j \to \infty} \nu(K'_{nj} \cap H)$$

for every $H \in \Sigma$. Now, for each m, $n \in \mathbb{N}$, choose $j(m, n)$ such that

$$\nu(E_n \cap B(\mathbf{0}, m) \cap K'_{n,j(m,n)}) \ge \nu(E_n \cap B(\mathbf{0}, m)) - 2^{-(m+n)}.$$

Set $K_m = \bigcap_{n \in \mathbb{N}} K'_{n,j(m,n)}$; then K_m is closed (being an intersection of closed sets) and bounded (being a subset of $K'_{0,j(m,0)}$), therefore compact. Also $K_m \subseteq F$, because $K'_{n,j(m,n)} \subseteq E_n$ for each n, and

$$\nu(F \cap B(\mathbf{0}, m) \setminus K_m) \le \sum_{n=0}^{\infty} \nu(E_n \cap B(\mathbf{0}, m) \setminus K'_{n,j(m,n)}) \le \sum_{n=0}^{\infty} 2^{-(m+n)} = 2^{-m+1}.$$

Consequently $H = \bigcup_{m \in \mathbb{N}} K_m$ is a K_σ subset of F and

$$\nu(F \cap B(\mathbf{0}, m) \setminus H) \le \inf_{k \ge m} \nu(F \cap B(\mathbf{0}, k) \setminus H_k) = 0$$

for every m, so $\nu(F \setminus H) = 0$ and $F \in \mathcal{A}$. **Q**

(d) Σ is a σ-algebra of subsets of $\mathbb{R}$. **P** (i) $\emptyset$ and its complement are open, so belong to $\mathcal{A}$ and therefore to Σ. (ii) If $E \in \Sigma$ then both $\mathbb{R}^r \setminus E$ and $\mathbb{R}^r \setminus (\mathbb{R}^r \setminus E) = E$ belong to $\mathcal{A}$, so $\mathbb{R}^r \setminus E \in \Sigma$. (iii) Let $\langle E_n \rangle_{n \in \mathbb{N}}$ be a sequence in Σ with union E. By (a-iii) and (a-iv),

$$E \in \mathcal{A}, \quad \mathbb{R}^r \setminus E = \bigcap_{n \in \mathbb{N}} (\mathbb{R}^r \setminus E_n) \in \mathcal{A},$$

so $E \in \Sigma$. **Q**

(e) By (c-i) and (c-ii), every open set belongs to Σ; consequently every Borel set belongs to Σ and therefore to $\mathcal{A}$. Now if E is any member of Σ, there is a Borel set $E_1 \subseteq E$ such that $\nu(E \setminus E_1) = 0$ and a K_σ set $H \subseteq E_1$ such that $\nu(E_1 \setminus H) = 0$. Express H as $\bigcup_{n \in \mathbb{N}} K_n$ where every K_n is compact; then

$$\nu E = \nu H = \lim_{n \to \infty} \nu(\bigcup_{i \le n} K_i) \le \sup_{K \subseteq E \text{ is compact}} \nu K \le \nu E$$

because $\bigcup_{i \in n} K_i$ is a compact subset of E for every n.

(f) Thus ν is inner regular with respect to the compact sets. But of course it is complete (being the completion of ν_0) and a locally finite topological measure (because ν_0 is); so it is a Radon measure. This completes the proof.

256D Proposition If ν and ν' are two Radon measures on $\mathbb{R}^r$, the following are equiveridical:
(i) $\nu = \nu'$;
(ii) $\nu K = \nu' K$ for every compact set $K \subseteq \mathbb{R}^r$;
(iii) $\nu G = \nu' G$ for every open set $G \subseteq \mathbb{R}^r$;
(iv) $\int h \, d\nu = \int h \, d\nu'$ for every continuous function $h : \mathbb{R}^r \to \mathbb{R}$ with bounded support.

proof (a)(i)⇒(iv) is trivial.

(b)(iv)⇒(iii) If (iv) is true, and $G \subseteq \mathbb{R}^r$ is an open set, then for each $n \in \mathbb{N}$ set

$$h_n(x) = \min(1, 2^n \inf_{y \in \mathbb{R}^r \setminus (G \cap B(\mathbf{0}, n))} \|y - x\|)$$

for $x \in \mathbb{R}^r$. Then h_n is continuous (in fact $|h_n(x) - h_n(x')| \le 2^n\|x - x'\|$ for all x, $x' \in \mathbb{R}^r$) and zero outside $B(\mathbf{0}, n)$, so $\int h_n d\nu = \int h_n d\nu'$. Next, $\langle h_n(x)\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence converging to $\chi G(x)$ for every $x \in \mathbb{R}^r$. So

$$\nu G = \lim_{n\to\infty} \int h_n d\nu = \lim_{n\to\infty} \int h_n d\nu' = \nu' G,$$

by 135Ga. As G is arbitrary, (iii) is true.

(c)(iii)⇒(ii) If (iii) is true, and $K \subseteq \mathbb{R}^r$ is compact, let n be so large that $\|x\| < n$ for every $x \in K$. Set $G = \{x : \|x\| < n\}$, $H = G \setminus K$. Then G and H are open and G is bounded, so $\nu G = \nu' G$ is finite, and

$$\nu K = \nu G - \nu H = \nu' G - \nu' H = \nu' K.$$

As K is arbitrary, (ii) is true.

(d)(ii)⇒(i) If ν, ν' agree on the compact sets, then

$$\nu E = \sup_{K\subseteq E \text{ is compact}} \nu K = \sup_{K\subseteq E \text{ is compact}} \nu' K = \nu' E$$

for every Borel set E. So $\nu\restriction\mathcal{B} = \nu'\restriction\mathcal{B}$, where $\mathcal{B}$ is the algebra of Borel sets. But since ν and ν' are both the completions of their restrictions to $\mathcal{B}$, they are identical.

256E It is I suppose time I gave some examples of Radon measures. However it will save a few lines if I first establish some basic constructions. You may wish to glance ahead to 256H at this point.

Theorem Let ν be a Radon measure on $\mathbb{R}^r$, with domain Σ, and f a non-negative Σ-measurable function defined on a ν-conegligible subset of $\mathbb{R}^r$. Suppose that f is **locally integrable** in the sense that $\int_E f d\nu < \infty$ for every bounded set E. Then the indefinite-integral measure ν' on $\mathbb{R}^r$ defined by saying that

$$\nu' E = \int_E f d\nu \text{ whenever } E \cap \{x : x \in \operatorname{dom} f,\ f(x) > 0\} \in \Sigma$$

is a Radon measure on $\mathbb{R}^r$.

proof For the construction of ν', see 234I-234L. Indefinite-integral measures, as I define them, are always complete (234I). ν' is locally finite because f is locally integrable. ν' is a topological measure because every open set belongs to Σ and therefore to the domain Σ' of ν'. To see that ν' is inner regular with respect to the compact sets, take any set $E \in \Sigma'$, and set $E' = \{x : x \in E \cap \operatorname{dom} f,\ f(x) > 0\}$. Then $E' \in \Sigma$, so there is a set $H \subseteq E'$, expressible as the union of a sequence of compact sets, such that $\nu(E' \setminus H) = 0$. In this case

$$\nu'(E \setminus H) = \int_{E\setminus H} f d\nu = 0.$$

Let $\langle K_n\rangle_{n\in\mathbb{N}}$ be a sequence of compact sets with union H; then

$$\nu' E = \nu' H = \lim_{n\to\infty} \nu'(\bigcup_{i\le n} K_i) \le \sup_{K\subseteq E \text{ is compact}} \nu' K \le \nu' E.$$

As E is arbitrary, ν' is inner regular with respect to the compact sets.

256F Theorem Let ν be a Radon measure on $\mathbb{R}^r$, and Σ its domain. Let $f : D \to \mathbb{R}$ be a Σ-measurable function, where $D \subseteq \mathbb{R}^r$. Then for every $\epsilon > 0$ there is a closed set $F \subseteq \mathbb{R}^r$ such that $\nu(\mathbb{R}^r \setminus F) \le \epsilon$ and $f\restriction F$ is continuous.

proof By 121I, there is a Σ-measurable function $h : \mathbb{R}^r \to \mathbb{R}$ extending f. Enumerate $\mathbb{Q}$ as $\langle q_n\rangle_{n\in\mathbb{N}}$. For each $n \in \mathbb{N}$ set $E_n = \{x : h(x) \le q_n\}$, $E'_n = \{x : h(x) > q_n\}$ and use 256Bb to choose closed sets $F_n \subseteq E_n$, $F'_n \subseteq E'_n$ such that $\nu(E_n \setminus F_n) \le 2^{-n-2}\epsilon$ and $\nu(E'_n \setminus F'_n) \le 2^{-n-2}\epsilon$. Set $F = \bigcap_{n\in\mathbb{N}}(F_n \cup F'_n)$; then F is closed and

$$\nu(\mathbb{R}^r \setminus F) \le \sum_{n=0}^{\infty} \nu(\mathbb{R}^r \setminus (F_n \cup F'_n)) \le \sum_{n=0}^{\infty} \nu(E_n \setminus F_n) + \nu(E'_n \setminus F'_n) \le \epsilon.$$

I claim that $h\restriction F$ is continuous. **P** Suppose that $x \in F$ and $\delta > 0$. Then there are m, $n \in \mathbb{N}$ such that

$$h(x) - \delta \le q_m < h(x) \le q_n \le h(x) + \delta.$$

This means that $x \in E'_m \cap E_n$; consequently $x \notin F_m \cup F'_n$. Because $F_m \cup F'_n$ is closed, there is an $\eta > 0$ such that $y \notin F_m \cup F'_n$ whenever $\|y - x\| \le \eta$. Now suppose that $y \in F$ and $\|y - x\| \le \eta$. Then

$y \in (F_m \cup F'_m) \cap (F_n \cup F'_n)$ and $y \notin F_m \cup F'_n$, so $y \in F'_m \cap F_n \subseteq E'_m \cap E_n$ and $q_m < h(y) \le q_n$. Consequently $|h(y) - h(x)| \le \delta$. As x and δ are arbitrary, $h \restriction F$ is continuous. **Q** It follows that $f \restriction F = (h \restriction F) \restriction D$ is continuous, as required.

256G Theorem Let ν be a Radon measure on $\mathbb{R}^r$, with domain Σ, and suppose that $\phi : \mathbb{R}^r \to \mathbb{R}^s$ is measurable in the sense that all its coordinates are Σ-measurable. If the image measure $\nu' = \nu\phi^{-1}$ (234D) is locally finite, it is a Radon measure.

proof Write Σ' for the domain of ν'. If $\phi = (\phi_1, \dots, \phi_s)$, then

$$\phi^{-1}[\{y : \eta_j \le \alpha\}] = \{x : \phi_j(x) \le \alpha\} \in \Sigma,$$

so $\{y : \eta_j \le \alpha\} \in \Sigma'$ for every $j \le s$, $\alpha \in \mathbb{R}$, where I write $y = (\eta_1, \dots, \eta_s)$ for $y \in \mathbb{R}^s$. Consequently every Borel subset of $\mathbb{R}^s$ belongs to Σ' (121J), and ν' is a topological measure. It is complete by 234Eb.

The point is of course that ν' is inner regular with respect to the compact sets. **P** Suppose that $F \in \Sigma'$ and that $\gamma < \nu' F$. For each $j \le s$, there is a closed set $H_j \subseteq \mathbb{R}^r$ such that $\phi_j \restriction H_j$ is continuous and $\nu(\mathbb{R}^r \setminus H_j) < \frac{1}{s}(\nu' F - \gamma)$, by 256F. Set $H = \bigcap_{j \le s} H_j$; then H is closed and $\phi \restriction H$ is continuous and

$$\nu(\mathbb{R}^r \setminus H) < \nu' F - \gamma = \nu\phi^{-1}[F] - \gamma,$$

so that $\nu(\phi^{-1}[F] \cap H) > \gamma$. Let $K \subseteq \phi^{-1}[F] \cap H$ be a compact set such that $\nu K \ge \gamma$, and set $L = \phi[K]$. Because $K \subseteq H$ and $\phi \restriction H$ is continuous, L is compact (2A2Eb). Of course $L \subseteq F$, and

$$\nu' L = \nu\phi^{-1}[L] \ge \nu K \ge \gamma.$$

As F and γ are arbitrary, ν' is inner regular with respect to the compact sets. **Q**

Since ν' is locally finite by the hypothesis of the theorem, it is a Radon measure.

256H Examples I come at last to the promised examples.

(a) Lebesgue measure on $\mathbb{R}^r$ is a Radon measure. (It is a topological measure by 115G, and inner regular with respect to the compact sets by 134Fb.)

(b) A point-supported measure on $\mathbb{R}^r$ is a Radon measure iff it is locally finite. **P** Let μ be a point-supported measure on $\mathbb{R}^r$. If it is a Radon measure, then of course it is locally finite. If it is locally finite, then surely it is a complete topological measure, since it measures every subset of $\mathbb{R}^r$. Let $h : \mathbb{R}^r \to [0, \infty]$ be such that $\mu E = \sum_{x \in E} h(x)$ for every $E \subseteq \mathbb{R}^r$. Take any $E \subseteq \mathbb{R}^r$. Then

$$\begin{aligned} \mu E &= \sum_{x \in E} h(x) = \sup_{I \subseteq E \text{ is finite}} \sum_{x \in I} h(x) \\ &= \sup_{I \subseteq E \text{ is finite}} \mu I \le \sup_{K \subseteq E \text{ is compact}} \mu K \le \mu E \end{aligned}$$

so $\mu E = \sup_{K \subseteq E \text{ is compact}} \mu K$; thus μ is inner regular with respect to the compact sets and is a Radon measure. **Q**

(c) Now we come to a new idea. Recall that the Cantor set C (134G) is a closed Lebesgue negligible subset of $[0, 1]$, and that the Cantor function (134H) is a non-decreasing continuous function $f : [0, 1] \to [0, 1]$ such that $f(0) = 0$, $f(1) = 1$ and f is constant on each of the intervals composing $[0, 1] \setminus C$. It follows that if we set $g(x) = \frac{1}{2}(x + f(x))$ for $x \in [0, 1]$, then $g : [0, 1] \to [0, 1]$ is a continuous permutation such that the Lebesgue measure of $g[C]$ is $\frac{1}{2}$ (134I); consequently $g^{-1} : [0, 1] \to [0, 1]$ is continuous. Now extend g to a permutation $h : \mathbb{R} \to \mathbb{R}$ by setting $h(x) = x$ for $x \in \mathbb{R} \setminus [0, 1]$. Then h and h^{-1} are continuous. Note that $h[C] = g[C]$ has Lebesgue measure $\frac{1}{2}$.

Let ν_1 be the indefinite-integral measure defined from Lebesgue measure μ on $\mathbb{R}$ and the function $2\chi(h[C])$; that is, $\nu_1 E = 2\mu(E \cap h[C])$ whenever this is defined. By 256E, ν_1 is a Radon measure, and $\nu_1 h[C] = \nu_1 \mathbb{R} = 1$. Let ν be the measure $\nu_1 (h^{-1})^{-1}$, that is, $\nu E = \nu_1 h[E]$ for just those $E \subseteq \mathbb{R}$ such that $h[E] \in \operatorname{dom} \nu_1$. Then ν is a Radon probability measure on $\mathbb{R}$, by 256G, and $\nu C = 1$, $\nu(\mathbb{R} \setminus C) = \mu C = 0$.

256I Remarks (a) The measure ν of 256Hc, sometimes called **Cantor measure**, is a classic example, and as such has many constructions, some rather more natural than the one I use here (see 256Xk, and also

264Ym below). But I choose the method above because it yields directly, without further investigation or any appeal to more advanced general theory, the fact that ν is a Radon measure.

(b) The examples above are chosen to represent the extremes under the 'Lebesgue decomposition' described in 232I. If ν is a (totally finite) Radon measure on $\mathbb{R}^r$, we can use 232Ib to express its restriction $\nu\restriction\mathcal{B}$ to the Borel σ-algebra as $\nu_p+\nu_{ac}+\nu_{cs}$, where ν_p is the 'point-mass' or 'atomic' part of $\nu\restriction\mathcal{B}$, ν_{ac} is the 'absolutely continuous' part (with respect to Lebesgue measure), and ν_{cs} is the 'atomless singular part'. In the example of 256Hb, we have $\nu\restriction\mathcal{B}=\nu_p$; in 256E, if we start from Lebesgue measure, we have $\nu\restriction\mathcal{B}=\nu_{ac}$; and in 256Hc we have $\nu\restriction\mathcal{B}=\nu_{cs}$.

256J Absolutely continuous Radon measures It is worth pausing a moment over the indefinite-integral measures described in 256E.

Proposition Let ν be a Radon measure on $\mathbb{R}^r$, and write μ for Lebesgue measure on $\mathbb{R}^r$. Then the following are equiveridical:

(i) ν is an indefinite-integral measure over μ;

(ii) $\nu E=0$ whenever E is a Borel subset of $\mathbb{R}^r$ and $\mu E=0$.

In this case, if $g\in\mathcal{L}^0(\mu)$ and $\int_E g\,d\mu=\nu E$ for every Borel set $E\subseteq\mathbb{R}^r$, then g is a Radon-Nikodým derivative of ν with respect to μ in the sense of 232Hf.

proof (a)(i)⇒(ii) If f is a Radon-Nikodým derivative of ν with respect to μ, then of course

$$\nu E=\int_E fd\mu=0$$

whenever $\mu E=0$.

(ii)⇒(i) If $\nu E=0$ for every μ-negligible Borel set E, then νE is defined and equal to 0 for every μ-negligible set E, because ν is complete and any μ-negligible set is included in a μ-negligible Borel set. Consequently $\operatorname{dom}\nu$ includes the domain Σ of μ, since every Lebesgue measurable set is expressible as the union of a Borel set and a negligible set.

For each $n\in\mathbb{N}$ set $E_n=\{x:n\le\|x\|<n+1\}$, so that $\langle E_n\rangle_{n\in\mathbb{N}}$ is a partition of $\mathbb{R}^r$ into bounded Borel sets. Set $\nu_nE=\nu(E\cap E_n)$ for every Lebesgue measurable set E and every $n\in\mathbb{N}$. Now ν_n is absolutely continuous with respect to μ (232Ba), so by the Radon-Nikodým theorem (in the form 232F) there is a μ-integrable function f_n such that $\int_E f_nd\mu=\nu_nE$ for every Lebesgue measurable set E. Because $\nu_nE\ge 0$ for every $E\in\Sigma$, $f_n\ge 0$ a.e.; because $\nu_n(\mathbb{R}^r\setminus E_n)=0$, $f_n=0$ a.e. on $\mathbb{R}^r\setminus E_n$. Now if we set

$$f=\max(0,\sum_{n=0}^\infty f_n),$$

f will be defined μ-a.e. and we shall have

$$\int_E fd\mu=\sum_{n=0}^\infty\int_E f_nd\mu=\sum_{n=0}^\infty\nu(E\cap E_n)=\nu E$$

for every Borel set E, so that the indefinite-integral measure ν' defined by f and μ agrees with ν on the Borel sets. Since this ensures that ν' is locally finite, ν' is a Radon measure, by 256E, and is equal to ν, by 256D. Accordingly ν is an indefinite-integral measure over μ.

(b) As in (a-ii) above, g must be locally integrable and the indefinite-integral measure defined by g agrees with ν on the Borel sets, so is identical with ν.

256K Products The class of Radon measures on Euclidean spaces is stable under a wide variety of operations, as we have already seen; in particular, we have the following.

Theorem Let ν_1, ν_2 be Radon measures on $\mathbb{R}^r$ and $\mathbb{R}^s$ respectively. Let λ be their c.l.d. product measure on $\mathbb{R}^r\times\mathbb{R}^s$. Then λ is a Radon measure.

Remark When I say that λ is 'Radon' according to the definition in 256A, I am of course identifying $\mathbb{R}^r\times\mathbb{R}^s$ with $\mathbb{R}^{r+s}$, as in 251M-251N.

proof I hope the following notation will seem natural. Write Σ_1, Σ_2 for the domains of ν_1, ν_2; $\mathcal{B}_r$, $\mathcal{B}_s$ for the Borel σ-algebras of $\mathbb{R}^r$, $\mathbb{R}^s$; Λ for the domain of λ; and $\mathcal{B}$ for the Borel σ-algebra of $\mathbb{R}^{r+s}$.

Because each ν_i is the completion of its restriction to the Borel sets (256C), λ is the product of $\nu_1 \restriction \mathcal{B}_r$ and $\nu_2 \restriction \mathcal{B}_s$ (251T). Because $\nu_1 \restriction \mathcal{B}_r$ and $\nu_2 \restriction \mathcal{B}_s$ are σ-finite (256Ba, 212Ga), λ must be the completion of its restriction to $\mathcal{B}_r \widehat{\otimes} \mathcal{B}_s$, which by 251M is identified with $\mathcal{B}$. Setting $Q_n = \{(x, y) : \|x\| \le n, \|y\| \le n\}$ we have

$$\lambda Q_n = \nu_1\{x : \|x\| \le n\} \cdot \nu_2\{y : \|y\| \le n\} < \infty$$

for every n, while every bounded subset of $\mathbb{R}^{r+s}$ is included in some Q_n. So $\lambda \restriction \mathcal{B}$ is locally finite, and its completion λ is a Radon measure, by 256C.

256L Remark We see from 253I that if ν_1 and ν_2 are Radon measures on $\mathbb{R}^r$ and $\mathbb{R}^s$ respectively, and both are indefinite-integral measures over Lebesgue measure, then their product measure on $\mathbb{R}^{r+s}$ is also an indefinite-integral measure over Lebesgue measure.

***256M** For the sake of applications in §286 below, I include another result, which is in fact one of the fundamental properties of Radon measures, as will appear in §414.

Proposition Let ν be a Radon measure on $\mathbb{R}^r$, and D any subset of $\mathbb{R}^r$. Let Φ be a non-empty upwards-directed family of non-negative continuous functions from D to $\mathbb{R}$. For $x \in D$ set $g(x) = \sup_{f \in \Phi} f(x)$ in $[0, \infty]$. Then

(a) $g : D \to [0, \infty]$ is lower semi-continuous, therefore Borel measurable;

(b) $\int_D g \, d\nu = \sup_{f \in \Phi} \int_D f d\nu$.

proof (a) For any $u \in [-\infty, \infty]$,

$$\{x : x \in D, \, g(x) > u\} = \bigcup_{f \in \Phi} \{x : x \in D, \, f(x) > u\}$$

is an open set for the subspace topology on D (2A3C), so is the intersection of D with a Borel subset of $\mathbb{R}^r$. This is enough to show that g is Borel measurable (121B-121C).

(b) Accordingly $\int_D g \, d\nu$ will be defined in $[0, \infty]$, and of course $\int_D g \, d\nu \ge \sup_{f \in \Phi} \int_D f d\nu$.

For the reverse inequality, observe that there is a countable set $\Psi \subseteq \Phi$ such that $g(x) = \sup_{f \in \Psi} f(x)$ for every $x \in D$. **P** For $a \in \mathbb{Q}$, q, $q' \in \mathbb{Q}^r$ set

$$\Phi_{aqq'} = \{f : f \in \Phi, \, f(y) > a \text{ whenever } y \in D \cap [q, q']\},$$

interpreting $[q, q']$ as in 115G. Choose $f_{aqq'} \in \Phi_{aqq'}$ if $\Phi_{aqq'}$ is not empty, and arbitrarily in Φ otherwise; and set $\Psi = \{f_{aqq'} : a \in \mathbb{Q}, \, q, \, q' \in \mathbb{Q}^r\}$, so that Ψ is a countable subset of Φ. If $x \in D$ and $b < g(x)$, there is an $a \in \mathbb{Q}$ such that $b \le a < g(x)$; there is an $\hat{f} \in \Phi$ such that $\hat{f}(x) > a$; because $\hat{f}$ is continuous, there are q, $q' \in \mathbb{Q}^r$ such that $q \le x \le q'$ and $\hat{f}(y) \ge a$ whenever $y \in D \cap [q, q']$; so that $\hat{f} \in \Phi_{aqq'}$, $\Phi_{aqq'} \ne \emptyset$, $f_{aqq'} \in \Phi_{aqq'}$ and $\sup_{f \in \Psi} f(x) \ge f_{aqq'}(x) \ge b$. As b is arbitrary, $g(x) = \sup_{f \in \Psi} f(x)$. **Q**

Let $\langle f_n \rangle_{n \in \mathbb{N}}$ be a sequence running over Ψ. Because Φ is upwards-directed, we can choose $\langle f'_n \rangle_{n \in \mathbb{N}}$ in Φ inductively in such a way that $f'_{n+1} \ge \max(f'_n, f_n)$ for every $n \in \mathbb{N}$. So $\langle f'_n \rangle_{n \in \mathbb{N}}$ is a non-decreasing sequence in Φ and $\sup_{n \in \mathbb{N}} f'_n(x) \ge \sup_{f \in \Psi} f(x) = g(x)$ for every $x \in D$. By B.Levi's theorem,

$$\int_D g \, d\nu \le \sup_{n \in \mathbb{N}} \int_D f'_n d\nu \le \sup_{f \in \Phi} \int_D f d\nu,$$

and we have the required inequality.

256X Basic exercises >(a) Let ν be a measure on $\mathbb{R}^r$. (i) Show that it is locally finite, in the sense of 256Ab, iff for every $x \in \mathbb{R}^r$ there is a $\delta > 0$ such that $\nu^* B(x, \delta) < \infty$. (*Hint*: the sets $B(\mathbf{0}, n)$ are compact.) (ii) Show that in this case ν is σ-finite.

>(b) Let ν be a Radon measure on $\mathbb{R}^r$ and $\mathcal{G}$ a non-empty upwards-directed family of open sets in $\mathbb{R}^r$. (i) Show that $\nu(\bigcup \mathcal{G}) = \sup_{G \in \mathcal{G}} \nu G$. (*Hint*: observe that if $K \subseteq \bigcup \mathcal{G}$ is compact, then $K \subseteq G$ for some $G \in \mathcal{G}$.) (ii) Show that $\nu(E \cap \bigcup \mathcal{G}) = \sup_{G \in \mathcal{G}} \nu(E \cap G)$ for every set E which is measured by ν.

>(c) Let ν be a Radon measure on $\mathbb{R}^r$ and $\mathcal{F}$ a non-empty downwards-directed family of closed sets in $\mathbb{R}^r$ such that $\inf_{F \in \mathcal{F}} \nu F < \infty$. (i) Show that $\nu(\bigcap \mathcal{F}) = \inf_{F \in \mathcal{F}} \nu F$. (*Hint*: apply 256Xb(ii) to $\mathcal{G} = \{\mathbb{R}^r \setminus F : F \in \mathcal{F}\}$.) (ii) Show that $\nu(E \cap \bigcap \mathcal{F}) = \inf_{F \in \mathcal{F}} \nu(E \cap F)$ for every E in the domain of ν.

>(d) Show that a Radon measure ν on $\mathbb{R}^r$ is atomless iff $\nu\{x\}=0$ for every $x\in\mathbb{R}^r$. (*Hint*: apply 256Xc with $\mathcal{F}=\{F:F\subseteq E$ is closed, not negligible$\}$.)

(e) Let ν_1, ν_2 be Radon measures on $\mathbb{R}^r$, and α_1, $\alpha_2\in\,]0,\infty[$. Set $\Sigma=\operatorname{dom}\nu_1\cap\operatorname{dom}\nu_2$, and for $E\in\Sigma$ set $\nu E=\alpha_1\nu_1E+\alpha_2\nu_2E$. Show that ν is a Radon measure on $\mathbb{R}^r$. Show that ν is an indefinite-integral measure over Lebesgue measure iff ν_1, ν_2 are, and that in this case a linear combination of of Radon-Nikodým derivatives of ν_1 and ν_2 is a Radon-Nikodým derivative of ν.

>(f) Let ν be a Radon measure on $\mathbb{R}^r$. (i) Show that there is a unique closed set $F\subseteq\mathbb{R}^r$ such that, for open sets $G\subseteq\mathbb{R}^r$, $\nu G>0$ iff $G\cap F\neq\emptyset$. (F is called the **support** of ν.) (ii) Generally, a set $A\subseteq\mathbb{R}^r$ is called **self-supporting** if $\nu^*(A\cap G)>0$ whenever $G\subseteq\mathbb{R}^r$ is an open set meeting A. Show that for every closed set $F\subseteq\mathbb{R}^r$ there is a unique self-supporting closed set $F'\subseteq F$ such that $\nu(F\setminus F')=0$.

>(g) Show that a measure ν on $\mathbb{R}$ is a Radon measure iff it is a Lebesgue-Stieltjes measure as described in 114Xa. Show that in this case ν is an indefinite-integral measure over Lebesgue measure iff the function $x\mapsto\nu[a,x]:[a,b]\to\mathbb{R}$ is absolutely continuous whenever $a\le b$ in $\mathbb{R}$.

(h) Let ν be a Radon measure on $\mathbb{R}^r$. Let C_k be the space of continuous real-valued functions on $\mathbb{R}^r$ with bounded supports. Show that for every ν-integrable function f and every $\epsilon>0$ there is a $g\in C_k$ such that $\int|f-g|d\nu\le\epsilon$. (*Hint*: use arguments from 242O, but in (a-i) of the proof there start with *closed* intervals I.)

(i) Let ν be a Radon measure on $\mathbb{R}^r$, and ν^* the corresponding outer measure. Show that $\nu A=\inf\{\nu G:G\supseteq A$ is open$\}$ for every set $A\subseteq\mathbb{R}^r$.

(j) Let ν, ν' be two Radon measures on $\mathbb{R}^r$, and suppose that $\nu I=\nu' I$ for every half-open interval $I\subseteq\mathbb{R}^r$ (definition: 115Ab). Show that $\nu=\nu'$.

(k) Let ν be Cantor measure (256Hc). (i) Show that if C_n is the nth set used in the construction of the Cantor set, so that C_n consists of 2^n intervals of length 3^{-n}, then $\nu I=2^{-n}$ for each of the intervals I composing C_n. (ii) Let λ be the usual measure on $\{0,1\}^{\mathbb{N}}$ (254J). Define $\phi:\{0,1\}^{\mathbb{N}}\to\mathbb{R}$ by setting $\phi(x)=\frac{2}{3}\sum_{n=0}^{\infty}3^{-n}x(n)$ for each $x\in\{0,1\}^{\mathbb{N}}$. Show that ϕ is a bijection between $\{0,1\}^{\mathbb{N}}$ and C. (iii) Show that if $\mathcal{B}$ is the Borel σ-algebra of $\mathbb{R}$, then $\{\phi^{-1}[E]:E\in\mathcal{B}\}$ is precisely the σ-algebra of subsets of $\{0,1\}^{\mathbb{N}}$ generated by the sets $\{x:x(n)=i\}$ for $n\in\mathbb{N}$, $i\in\{0,1\}$. (iv) Show that ϕ is an isomorphism between $(\{0,1\}^{\mathbb{N}},\lambda)$ and (C,ν_C), where ν_C is the subspace measure on C induced by ν.

(l) Let ν and ν' be two Radon measures on $\mathbb{R}^r$. Show that ν' is an indefinite-integral measure over ν iff $\nu'E=0$ whenever $\nu E=0$, and in this case a function f is a Radon-Nikodým derivative of ν' with respect to ν iff $\int_E fd\nu=\nu'E$ for every Borel set E.

256Y Further exercises (a) Let ν be a Radon measure on $\mathbb{R}^r$, and X any subset of $\mathbb{R}^r$; let ν_X be the subspace measure on X and Σ_X its domain, and give X its subspace topology. Show that ν_X has the following properties: (i) ν_X is complete and locally determined; (ii) every open subset of X belongs to Σ_X; (iii) $\nu_XE=\sup\{\nu_XF:F\subseteq E$ is closed in $X\}$ for every $E\in\Sigma_X$; (iv) whenever $\mathcal{G}$ is a non-empty upwards-directed family of open subsets of X, $\nu_X(\bigcup\mathcal{G})=\sup_{G\in\mathcal{G}}\nu_XG$; (v) every point of X belongs to an open set of finite measure.

(b) Let ν be a Radon measure on $\mathbb{R}^r$, with domain Σ, and $f:\mathbb{R}^r\to\mathbb{R}$ a function. Show that the following are equiveridical: (i) f is Σ-measurable; (ii) for every non-negligible set $E\in\Sigma$ there is a non-negligible $F\in\Sigma$ such that $F\subseteq E$ and $f\restriction F$ is continuous; (iii) for every set $E\in\Sigma$, $\nu E=\sup_{K\in\mathcal{K}_f,K\subseteq E}\nu K$, where $\mathcal{K}_f=\{K:K\subseteq\mathbb{R}^r$ is compact, $f\restriction K$ is continuous$\}$. (*Hint*: for (ii)$\Rightarrow$(i), apply 215B(iv) to $\mathcal{K}_f$.)

(c) Take ν, X, ν_X and Σ_X as in 256Ya. Suppose that $f:X\to\mathbb{R}$ is a function. Show that f is Σ_X-measurable iff for every non-negligible measurable set $E\subseteq X$ there is a non-negligible measurable $F\subseteq E$ such that $f\restriction F$ is continuous.

(d)(i) Let λ be the usual measure on $\{0,1\}^{\mathbb{N}}$. Define $\psi:\{0,1\}^{\mathbb{N}}\to\{0,1\}^{\mathbb{N}}$ by setting $\psi(x)(i)=x(i+1)$ for $x\in\{0,1\}^{\mathbb{N}}$ and $j\in\mathbb{N}$. Show that ψ is inverse-measure-preserving. (ii) Define $\theta:\mathbb{R}\to\mathbb{R}$ by setting $\theta(t)=<3t>=3t-\lfloor 3t\rfloor$ for $t\in\mathbb{R}$. Show that θ is inverse-measure-preserving for Cantor measure as defined in 256Hc.

(e) Let $\langle\nu_n\rangle_{n\in\mathbb{N}}$ be a sequence of Radon measures on $\mathbb{R}^r$. Show that there is a Radon measure ν on $\mathbb{R}^r$ such that every ν_n is an indefinite-integral measure over ν. (*Hint*: find a sequence $\langle\alpha_n\rangle_{n\in\mathbb{N}}$ of strictly positive numbers such that $\sum_{n=0}^{\infty}\alpha_n\nu_n B(\mathbf{0},k)<\infty$ for every k, and set $\nu=\sum_{n=0}^{\infty}\alpha_n\nu_n$, using the idea of 256Xe.)

(f) A set $G\subseteq\mathbb{R}^{\mathbb{N}}$ is **open** if for every $x\in G$ there are $n\in\mathbb{N}$, $\delta>0$ such that

$$\{y:y\in\mathbb{R}^{\mathbb{N}},\ |y(i)-x(i)|<\delta\text{ for every }i\le n\}\subseteq G.$$

The **Borel σ-algebra** of $\mathbb{R}^{\mathbb{N}}$ is the σ-algebra $\mathcal{B}$ of subsets of $\mathbb{R}^{\mathbb{N}}$ generated, in the sense of 111Gb, by the family $\mathfrak{T}$ of open sets. (i) Show that $\mathfrak{T}$ is a topology (2A3A). (ii) Show that a filter $\mathcal{F}$ on $\mathbb{R}^{\mathbb{N}}$ converges to $x\in\mathbb{R}^{\mathbb{N}}$ iff $\pi_i[[\mathcal{F}]]\to x(i)$ for every $i\in\mathbb{N}$, where $\pi_i(y)=y(i)$ for $i\in\mathbb{N}$, $y\in\mathbb{R}^{\mathbb{N}}$. (iii) Show that $\mathcal{B}$ is the σ-algebra generated by sets of the form $\{x:x\in\mathbb{R}^{\mathbb{N}},\ x(i)\le a\}$, where i runs over $\mathbb{N}$ and a runs over $\mathbb{R}$. (iv) Show that if $\alpha_i\ge 0$ for every $i\in\mathbb{N}$, then $\{x:|x(i)|\le\alpha_i\ \forall\ i\in\mathbb{N}\}$ is compact. (v) Show that any open set in $\mathbb{R}^{\mathbb{N}}$ is the union of a sequence of closed sets. (*Hint*: look at sets of the form $\{x:q_i\le x(i)\le q_i'\ \forall\ i\le n\}$, where q_i, $q_i'\in\mathbb{Q}$ for $i\le n$.) (vi) Show that if ν_0 is any probability measure with domain $\mathcal{B}$, then its completion ν is inner regular with respect to the compact sets, and therefore may be called a 'Radon measure on $\mathbb{R}^{\mathbb{N}}$'. (*Hint*: show that there are compact sets of measure arbitrarily close to 1, and therefore that every open set, and every closed set, includes a K_σ set of the same measure.)

256 Notes and comments Radon measures on Euclidean spaces are very special, and the results of this section do not give clear pointers to the direction the theory takes when applied to other kinds of topological space. With the material here you could make a stab at developing a theory of Radon measures on complete separable metric spaces, provided you use 256Xa as the basis for your definition of 'locally finite'. These are the spaces for which a version of 256C is true. (See 256Yf.) But for generalizations to other types of topological space, and for the more interesting parts of the theory on $\mathbb{R}^r$, I must ask you to wait for Volume 4. My purpose in introducing Radon measures here is strictly limited; I wish only to give a basis for §257 and §271 sufficiently solid not to need later revision. In fact I think that all we really need are the Radon probability measures.

The chief technical difficulty in the definition of 'Radon measure' here lies in the insistence on completeness. It may well be that for everything studied in this volume, it would be simpler to look at locally finite measures with domain the algebra of Borel sets. This would involve us in a number of circumlocutions when dealing with Lebesgue measure itself and its derivates, since Lebesgue measure is defined on a larger σ-algebra; but the serious objection arises in the more advanced theory, when non-Borel sets of various kinds become central. Since my aim in this book is to provide secure foundations for the study of all aspects of measure theory, I ask you to take a little extra trouble now in order to avoid the possibility of having to re-work all your ideas later. The extra trouble arises, for instance, in 256D, 256Xe and 256Xj; since different Radon measures are defined on different σ-algebras, we have to check that two Radon measures which agree on the compact sets, or on the open sets, have the same domains. On the credit side, some of the power of 256G arises from the fact that the Radon image measure $\nu\phi^{-1}$ is defined on the whole σ-algebra $\{F:\phi^{-1}[F]\in\operatorname{dom}(\nu)\}$, not just on the Borel sets.

The further technical point that Radon measures are expected to be locally finite gives less difficulty; its effect is that from most points of view there is little difference between a general Radon measure and a totally finite Radon measure. The extra condition which obviously has to be put into the hypotheses of such results as 256E and 256G is no burden on either intuition or memory.

In effect, we have two definitions of Radon measures on Euclidean spaces: they are the inner regular locally finite topological measures, and they are also the completions of the locally finite Borel measures. The equivalence of these definitions is Theorem 256C. The latter definition is the better adapted to 256K, and the former to 256G. The 'inner regularity' of the basic definition refers to compact sets; we also have forms of inner regularity with respect to closed sets (256Bb) and K_σ sets (256Bc), and a complementary notion of 'outer regularity' with respect to open sets (256Xi).

Version of 14.8.13

257 Convolutions of measures

The ideas of this chapter can be brought together in a satisfying way in the theory of convolutions of Radon measures, which will be useful in §272 and again in §285. I give just the definition (257A) and the central property (257B) of the convolution of totally finite Radon measures, with a few corollaries and a note on the relation between convolution of functions and convolution of measures (257F).

257A Definition Let $r \geq 1$ be an integer and ν_1, ν_2 two totally finite Radon measures on $\mathbb{R}^r$. Let λ be the product measure on $\mathbb{R}^r \times \mathbb{R}^r$; then λ also is a (totally finite) Radon measure, by 256K. Define $\phi : \mathbb{R}^r \times \mathbb{R}^r \to \mathbb{R}^r$ by setting $\phi(x, y) = x + y$; then ϕ is continuous, therefore measurable in the sense of 256G. The **convolution** of ν_1 and ν_2, $\nu_1 * \nu_2$, is the image measure $\lambda\phi^{-1}$; by 256G, this is a Radon measure.

Note that if ν_1 and ν_2 are Radon probability measures, then λ and $\nu_1 * \nu_2$ are also probability measures.

257B Theorem Let $r \geq 1$ be an integer, and ν_1 and ν_2 two totally finite Radon measures on $\mathbb{R}^r$; let $\nu = \nu_1 * \nu_2$ be their convolution, and λ their product on $\mathbb{R}^r \times \mathbb{R}^r$. Then for any real-valued function h defined on a subset of $\mathbb{R}^r$,

$$\int h(x + y)\lambda(d(x, y)) \text{ exists} = \int h(x)\nu(dx)$$

if either integral is defined in $[-\infty, \infty]$.

proof Apply 235J with $J(x, y) = 1$, $\phi(x, y) = x + y$ for all x, $y \in \mathbb{R}^r$.

257C Corollary Let $r \geq 1$ be an integer, and ν_1, ν_2 two totally finite Radon measures on $\mathbb{R}^r$; let $\nu = \nu_1 * \nu_2$ be their convolution, and λ their product on $\mathbb{R}^r \times \mathbb{R}^r$; write Λ for the domain of λ. Let h be a Λ-measurable function defined λ-almost everywhere in $\mathbb{R}^r$. Suppose that any one of the integrals

$$\iint |h(x + y)|\nu_1(dx)\nu_2(dy), \quad \iint |h(x + y)|\nu_2(dy)\nu_1(dx), \quad \int h(x + y)\lambda(d(x, y))$$

exists and is finite. Then h is ν-integrable and

$$\int h(x)\nu(dx) = \iint h(x + y)\nu_1(dx)\nu_2(dy) = \iint h(x + y)\nu_2(dy)\nu_1(dx).$$

proof Put 257B together with Fubini's and Tonelli's theorems (252H).

257D Corollary If ν_1 and ν_2 are totally finite Radon measures on $\mathbb{R}^r$, then $\nu_1 * \nu_2 = \nu_2 * \nu_1$.

proof For any Borel set $E \subseteq \mathbb{R}^r$, apply 257C to $h = \chi E$ to see that

$$\begin{aligned}(\nu_1 * \nu_2)(E) &= \iint \chi E(x + y)\nu_1(dx)\nu_2(dy) = \iint \chi E(x + y)\nu_2(dy)\nu_1(dx) \\ &= \iint \chi E(y + x)\nu_2(dy)\nu_1(dx) = (\nu_2 * \nu_1)(E).\end{aligned}$$

Thus $\nu_1 * \nu_2$ and $\nu_2 * \nu_1$ agree on the Borel sets of $\mathbb{R}^r$; because they are both Radon measures, they must be identical (256D).

257E Corollary If ν_1, ν_2 and ν_3 are totally finite Radon measures on $\mathbb{R}^r$, then $(\nu_1 * \nu_2) * \nu_3 = \nu_1 * (\nu_2 * \nu_3)$.

proof For any Borel set $E \subseteq \mathbb{R}^r$, apply 257B to $h = \chi E$ to see that

$$\begin{aligned}((\nu_1 * \nu_2) * \nu_3)(E) &= \iint \chi E(x + z)(\nu_1 * \nu_2)(dx)\nu_3(dz) \\ &= \iiint \chi E(x + y + z)\nu_1(dx)\nu_2(dy)\nu_3(dz)\end{aligned}$$

(because $x \mapsto \chi E(x + z)$ is Borel measurable for every z)

$$= \iint \chi E(x+y)\nu_1(dx)(\nu_2 * \nu_3)(dy)$$

(because $(x,y) \mapsto \chi E(x+y)$ is Borel measurable, so $y \mapsto \int \chi E(x+y)\nu_1(dx)$ is $(\nu_2 * \nu_3)$-integrable)

$$= (\nu_1 * (\nu_2 * \nu_3))(E).$$

Thus $(\nu_1 * \nu_2) * \nu_3$ and $\nu_1 * (\nu_2 * \nu_3)$ agree on the Borel sets of $\mathbb{R}^r$; because they are both Radon measures, they must be identical.

257F Theorem Suppose that ν_1 and ν_2 are totally finite Radon measures on $\mathbb{R}^r$ which are indefinite-integral measures over Lebesgue measure μ. Then $\nu_1 * \nu_2$ also is an indefinite-integral measure over μ; if f_1 and f_2 are Radon-Nikodým derivatives of ν_1, ν_2 respectively, then $f_1 * f_2$ is a Radon-Nikodým derivative of $\nu_1 * \nu_2$.

proof By 255H/255L, $f_1 * f_2$ is integrable with respect to μ, with $\int f_1 * f_2 d\mu = 1$, and of course $f_1 * f_2$ is non-negative. If $E \subseteq \mathbb{R}^r$ is a Borel set,

$$\int_E f_1 * f_2 d\mu = \iint \chi E(x+y) f_1(x) f_2(y)\mu(dx)\mu(dy)$$

(255G)

$$= \iint \chi E(x+y) f_2(y)\nu_1(dx)\mu(dy)$$

(because $x \mapsto \chi E(x+y)$ is Borel measurable)

$$= \iint \chi E(x+y)\nu_1(dx)\nu_2(dy)$$

(because $(x,y) \mapsto \chi E(x+y)$ is Borel measurable, so $y \mapsto \int \chi E(x+y)\nu_1(dx)$ is ν_2-integrable)

$$= (\nu_1 * \nu_2)(E).$$

So $f_1 * f_2$ is a Radon-Nikodým derivative of ν with respect to μ, by 256J.

257X Basic exercises >(a) Let $r \geq 1$ be an integer. Let δ_0 be the Dirac measure on $\mathbb{R}^r$ concentrated at 0. Show that δ_0 is a Radon probability measure on $\mathbb{R}^r$ and that $\delta_0 * \nu = \nu$ for every totally finite Radon measure on $\mathbb{R}^r$.

(b) Let μ and ν be totally finite Radon measures on $\mathbb{R}^r$, and E any set measured by their convolution $\mu * \nu$. Show that $\int \mu(E-y)\nu(dy)$ is defined in $[0,\infty]$ and equal to $(\mu * \nu)(E)$.

(c) Let $\nu_1, \dots, \nu_n$ be totally finite Radon measures on $\mathbb{R}^r$, and let ν be the convolution $\nu_1 * \dots * \nu_n$ (using 257E to see that such a bracketless expression is legitimate). Show that

$$\int h(x)\nu(dx) = \int \dots \int h(x_1 + \dots + x_n)\nu_1(dx_1)\dots\nu_n(dx_n)$$

for every ν-integrable function h.

(d) Let ν_1 and ν_2 be totally finite Radon measures on $\mathbb{R}^r$, with supports F_1, F_2 (256Xf). Show that the support of $\nu_1 * \nu_2$ is $\overline{\{x+y : x \in F_1,\, y \in F_2\}}$.

>(e) Let ν_1 and ν_2 be totally finite Radon measures on $\mathbb{R}^r$, and suppose that ν_1 has a Radon-Nikodým derivative f with respect to Lebesgue measure μ. Show that $\nu_1 * \nu_2$ has a Radon-Nikodým derivative g, where $g(x) = \int f(x-y)\nu_2(dy)$ for μ-almost every $x \in \mathbb{R}^r$.

(f) Suppose that ν_1, ν_2, ν_1' and ν_2' are totally finite Radon measures on $\mathbb{R}^r$, and that ν_1', ν_2' are absolutely continuous with respect to ν_1, ν_2 respectively. Show that $\nu_1' * \nu_2'$ is absolutely continuous with respect to $\nu_1 * \nu_2$.

257Y Further exercises (a) Let M be the space of countably additive functionals defined on the algebra $\mathcal{B}$ of Borel subsets of $\mathbb{R}$, with its norm $\|\nu\| = |\nu|(\mathbb{R})$ (see 231Yh). (i) Show that we have a unique bilinear operator $* : M \times M \to M$ such that $(\mu_1 \restriction \mathcal{B}) * (\mu_2 \restriction \mathcal{B}) = (\mu_1 * \mu_2) \restriction \mathcal{B}$ for all totally finite Radon measures μ_1, μ_2 on $\mathbb{R}$. (ii) Show that $*$ is commutative and associative. (iii) Show that $\|\nu_1 * \nu_2\| \le \|\nu_1\| \|\nu_2\|$ for all ν_1, $\nu_2 \in M$, so that M is a Banach algebra under this multiplication. (iv) Show that M has a multiplicative identity. (v) Show that $L^1(\mu)$ can be regarded as a closed subalgebra of M, where μ is Lebesgue measure on $\mathbb{R}^r$ (cf. 255Xi).

(b) Let us say that a **Radon measure on** $]-\pi, \pi]$ is a complete measure ν on $]-\pi, \pi]$ such that (i) every Borel subset of $]-\pi, \pi]$ belongs to the domain Σ of μ (ii) for every $E \in \Sigma$ there are Borel sets E_1, E_2 such that $E_1 \subseteq E \subseteq E_2$ and $\nu(E_2 \setminus E_1) = 0$ (iii) every compact subset of $]-\pi, \pi]$ has finite measure. Show that for any two totally finite Radon measures ν_1, ν_2 on $]-\pi, \pi]$ there is a unique totally finite Radon measure ν on $]-\pi, \pi]$ such that

$$\int h(x)\nu(dx) = \int h(x +_{2\pi} y)\nu_1(dx)\nu_2(dy)$$

for every ν-integrable function h, where $+_{2\pi}$ is defined as in 255Ma.

257 Notes and comments Of course convolution of functions and convolution of measures are very closely connected; the obvious link being 257F, but the correspondence between 255G and 257B is also very marked. In effect, they give us the same notion of convolution $u * v$ when u, v are positive members of L^1 and $u * v$ is interpreted in L^1 rather than as a function (257Ya). But we should have to go rather deeper than the arguments here to find ideas in the theory of convolution of measures to correspond to such results as 255K. I will return to questions of this type in §444 in Volume 4.

All the theorems of this section can be extended to general abelian locally compact Hausdorff topological groups; but for such generality we need much more advanced ideas (see §444), and for the moment I leave only the suggestion in 257Yb that you should try to adapt the ideas here to $]-\pi, \pi]$ or S^1.

Version of 10.11.06

Concordance

I list here the section and paragraph numbers which have (to my knowledge) appeared in print in references to this chapter, and which have since been changed.

251N Paragraph numbers in the second half of §251, referred to in editions of Volumes 3 and 4 up to and including 2006, and in Bogachev 07, have been changed, so that 251M-251S are now 251N-251T.

252Yf Exercise This exercise, referred to in the first edition of Volume 1, has been moved to 252Ym.

254Yh Exercise This exercise, referred to in the 2013 edition of Volume 4, has been moved to 254Ye.

References

Bogachev V.I. [07] *Measure theory.* Springer, 2007.

Version of 17.12.12

References for Volume 2

Alexits G. [78] (ed.) *Fourier Analysis and Approximation Theory.* North-Holland, 1978 (Colloq. Math. Soc. Janos Bolyai 19).

Antonov N.Yu. [96] 'Convergence of Fourier series', East J. Approx. 7 (1996) 187-196. [§286 *notes*.]

Arias de Reyna J. [02] *Pointwise Convergence of Fourier Series.* Springer, 2002 (Lecture Notes in Mathematics 1785). [§286 *notes*.]

Baker R. [04] '"Lebesgue measure" on $\mathbb{R}^\infty$, II', Proc. Amer. Math. Soc. 132 (2004) 2577-2591. [254Yb.]

Bergelson V., March P. & Rosenblatt J. [96] (eds.) *Convergence in Ergodic Theory and Probability.* de Gruyter, 1996.

Bogachev V.I. [07] *Measure theory.* Springer, 2007.

du Bois-Reymond P. [1876] 'Untersuchungen über die Convergenz und Divergenz der Fouriersche Darstellungformeln', Abh. Akad. München 12 (1876) 1-103. [§282 *notes*.]

Bourbaki N. [66] *General Topology.* Hermann/Addison-Wesley, 1968. [2A5F.]

Bourbaki N. [87] *Topological Vector Spaces.* Springer, 1987. [2A5E.]

Carleson L. [66] 'On convergence and growth of partial sums of Fourier series', Acta Math. 116 (1966) 135-157. [§282 *notes*, §286 *intro.*, §286 *notes*.]

Clarkson J.A. [1936] 'Uniformly convex spaces', Trans. Amer. Math. Soc. 40 (1936) 396-414. [244O.]

Defant A. & Floret K. [93] *Tensor Norms and Operator Ideals*, North-Holland, 1993. [§253 *notes*.]

Doob J.L. [53] *Stochastic Processes.* Wiley, 1953.

Dudley R.M. [89] *Real Analysis and Probability.* Wadsworth & Brooks/Cole, 1989. [§282 *notes*.]

Dunford N. & Schwartz J.T. [57] *Linear Operators I.* Wiley, 1957 (reprinted 1988). [§244 *notes*, 2A5J.]

Enderton H.B. [77] *Elements of Set Theory.* Academic, 1977. [§2A1.]

Engelking R. [89] *General Topology.* Heldermann, 1989 (Sigma Series in Pure Mathematics 6). [2A5F.]

Etemadi N. [96] 'On convergence of partial sums of independent random variables', pp. 137-144 in Bergelson March & Rosenblatt 96. [272V.]

Evans L.C. & Gariepy R.F. [92] *Measure Theory and Fine Properties of Functions.* CRC Press, 1992. [263Ec, §265 *notes*.]

Federer H. [69] *Geometric Measure Theory.* Springer, 1969 (reprinted 1996). [262C, 263Ec, §264 *notes*, §265 *notes*, §266 *notes*.]

Feller W. [66] *An Introduction to Probability Theory and its Applications,* vol. II. Wiley, 1966. [Chap. 27 *intro.*, 274H, 275Xc, 285N.]

Fremlin D.H. [74] *Topological Riesz Spaces and Measure Theory.* Cambridge U.P., 1974. [§232 *notes*, 241F, §244 *notes*, §245 *notes*, §247 *notes*.]

Fremlin D.H. [93] 'Real-valued-measurable cardinals', pp. 151-304 in Judah 93. [232H.]

Haimo D.T. [67] (ed.) *Orthogonal Expansions and their Continuous Analogues.* Southern Illinois University Press, 1967.

Hall P. [82] *Rates of Convergence in the Central Limit Theorem.* Pitman, 1982. [274H.]

Halmos P.R. [50] *Measure Theory.* Van Nostrand, 1950. [§251 *notes*, §252 *notes*, 255Yn.]

Halmos P.R. [60] *Naive Set Theory.* Van Nostrand, 1960. [§2A1.]

Hanner O. [56] 'On the uniform convexity of L^p and l^p', Arkiv för Matematik 3 (1956) 239-244. [244O.]

Henle J.M. [86] *An Outline of Set Theory.* Springer, 1986. [§2A1.]

Hoeffding W. [63] 'Probability inequalities for sums of bounded random variables', J. Amer. Statistical Association 58 (1963) 13-30. [272W.]

Hunt R.A. [67] 'On the convergence of Fourier series', pp. 235-255 in Haimo 67. [§286 *notes*.]

Jorsbøe O.G. & Mejlbro L. [82] *The Carleson-Hunt Theorem on Fourier Series.* Springer, 1982 (Lecture Notes in Mathematics 911). [§286 *notes*.]

Judah H. [93] (ed.) Proceedings of the Bar-Ilan Conference on Set Theory and the Reals, 1991. Amer. Math. Soc. (Israel Mathematical Conference Proceedings 6), 1993.

Kelley J.L. [55] *General Topology.* Van Nostrand, 1955. [2A5F.]

Kelley J.L. & Namioka I. [76] *Linear Topological Spaces.* Springer, 1976. [2A5C.]

Kirszbraun M.D. [1934] 'Über die zusammenziehenden und Lipschitzian Transformationen', Fund. Math. 22 (1934) 77-108. [262C.]

Kolmogorov A.N. [1926] 'Une série de Fourier-Lebesgue divergente partout', C. R. Acad. Sci. Paris 183 (1926) 1327-1328. [§282 *notes*.]

Komlós J. [67] 'A generalization of a problem of Steinhaus', Acta Math. Acad. Sci. Hung. 18 (1967) 217-229. [276H.]

Körner T.W. [88] *Fourier Analysis.* Cambridge U.P., 1988. [§282 *notes*.]

Köthe G. [69] *Topological Vector Spaces I.* Springer, 1969. [2A5C, 2A5E, 2A5J.]

Krivine J.-L. [71] *Introduction to Axiomatic Set Theory.* D. Reidel, 1971. [§2A1.]

Lacey M.T. [05] *Carleson's Theorem: Proof, Complements, Variations.* `http://arxiv.org/pdf/math/0307008v4.pdf`. [§286 *notes*.]

Lacey M.T. & Thiele C.M. [00] 'A proof of boundedness of the Carleson operator', Math. Research Letters 7 (2000) 1-10. [§286 *intro.*, 286H.]

Lebesgue H. [72] *Oeuvres Scientifiques.* L'Enseignement Mathématique, Institut de Mathématiques, Univ. de Genève, 1972. [Chap. 27 *intro.*]

Liapounoff A. [1901] 'Nouvelle forme du théorème sur la limite de probabilité', Mém. Acad. Imp. Sci. St-Pétersbourg 12(5) (1901) 1-24. [274Xh.]

Lighthill M.J. [59] *Introduction to Fourier Analysis and Generalised Functions.* Cambridge U.P., 1959. [§284 *notes*.]

Lindeberg J.W. [1922] 'Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung', Math. Zeitschrift 15 (1922) 211-225. [274H, §274 *notes*.]

Lipschutz S. [64] *Set Theory and Related Topics.* McGraw-Hill, 1964 (Schaum's Outline Series). [§2A1.]

Loève M. [77] *Probability Theory I.* Springer, 1977. [Chap. 27 *intro.*, 274H.]

Luxemburg W.A.J. & Zaanen A.C. [71] *Riesz Spaces I.* North-Holland, 1971. [241F.]

Mozzochi C.J. [71] *On the Pointwise Convergence of Fourier Series.* Springer, 1971 (Lecture Notes in Mathematics 199). [§286 *notes*.]

Naor A. [04] 'Proof of the uniform convexity lemma', `http://www.cims.nyu.edu/~naor/homepage files/inequality.pdf`, 26.2.04. [244O.]

Rényi A. [70] *Probability Theory.* North-Holland, 1970. [274H.]

Roitman J. [90] *An Introduction to Set Theory.* Wiley, 1990. [§2A1.]

Roselli P. & Willem M. [02] 'A convexity inequality', Amer. Math. Monthly 109 (2002) 64-70. [244Ym.]

Saks S. [1924] 'Sur les nombres dérivés des fonctions', Fund

Schipp F. [78] 'On Carleson's method', pp. 679-695 in Alexits 78. [§286 *notes*.]

Semadeni Z. [71] *Banach spaces of continuous functions I.* Polish Scientific Publishers, 1971. [§253 *notes*.]

Shiryayev A.N. [84] *Probability.* Springer, 1984. [285N.]

Steele J.M. [86] 'An Efron-Stein inequality of nonsymmetric statistics', Annals of Statistics 14 (1986) 753-758. [274Ya.]

Zaanen A.C. [83] *Riesz Spaces II.* North-Holland, 1983. [241F.]

Zygmund A. [59] *Trigonometric Series.* Cambridge U.P, 1959. [§244 *notes*, §282 *notes*, 284Xk.]