Marc A Rieffel

Chapter 6 - Product Measures and Fubini's Theorem

Note: Following Lemma 1.17 of Chapter 1 we should have had:

1.18 Corollary. The ring generated by a semiring P consists of the finite disjoint unions of elements of P.

Proof. By Lemma 1.17 we have

$$\oplus E_m - \oplus F_n = \oplus_m (E_m - \oplus_n F_n) = \oplus_m \oplus_{j=1}^{k_m} G_{mj}$$

where the E_m, F_n and G_{mj} are elements of P. Thus the collection of finite disjoint unions of elements of P is closed under taking differences. But this collection is also clearly closed under taking finite disjoint unions. It follows that it is closed under taking finite unions, since $E \cup F = E \oplus (F-E)$.//

The numbers of subsequent items in Chapter 1 must thus all be increased by 1. In Chapter 6 references to items in Chapter 1 will be numbered accordingly.

Also, pages 4.27 and 4.28 of Chapter 4 should be replaced by the following pages:

non-decreasing, and are clearly bounded above by c, and so they form a Cauchy sequence. But $f_n - f_m$ is either positive a.e. or negative a.e., depending on whether or not n is larger than m, and so we have

$$\|f_n - f_m\|_1 = \int |f_n - f_m| d|\mu|| = |\int (f_n - f_m) d|\mu|| = |\int f_n d|\mu| - \int f_m d|\mu||.$$

Thus f_n is a mean Cauchy sequence also. Since $\mathcal{L}^1$ is complete, there exists an $f \in \mathcal{L}^1$ to which the sequence f_n converges in mean, and hence in measure. But then by Corollary 3.27 there must be a subsequence of the f_n which converges to f a.e. Since the f_n are nondecreasing, it follows that the sequence f_n itself also converges to f a.e.//

<u>4.43 Corollary</u>. If f_n is a sequence of real-valued integrable functions which is non-decreasing a.e. and which converges a.e. to a function f, and if the sequence of the norms of the f_n is bounded, then f is integrable and the sequence f_n converges to f in mean. In particular, $\int f_n d\mu$ converges to $\int f d\mu$. The same result holds if instead the sequence f_n is non-increasing a.e.

<u>Proof</u>. By Theorem 4.42 the sequence f_n will converge a.e. and in mean to an integrable function h, which must of course equal f a.e.//

By making the appropriate definition of the integrals with respect to a non-negative measure of arbitrary non-negative extended real-valued functions which are measurable in a natural sense, we can conveniently state a corollary of the Monotone Convergence Theorem which is useful in certain situations. For a typical application of this corollary see the proof of Theorem 6.9.

<u>4.44 Definition</u>. Let μ be a non-negative measure, and let f be a non-negative extended real-valued function on X. Let E be the set where f takes the value ∞. We say that f is μ-<u>measurable</u> if E is μ-measurable and if f is μ-measurable in the usual sense on $X - E$. If f is μ-measurable, then we define the integral of f with respect to μ as follows. If $\mu(E) > 0$ then we set $\int f d\mu = \infty$. If $\mu(E) = 0$, then we can view f as being undefined on the null set E, and we then set $\int f d\mu = \infty$ if f is not integrable in the usual sense, whereas if f is integrable, then we let $\int f d\mu$ be the usual integral of f.

<u>4.45 Corollary</u>. If μ is a non-negative measure and if f_n is a non-decreasing sequence of μ-measurable non-negative extended real-valued functions which converges a.e. to a non-negative extended real-valued function f, then f is μ-measurable and the sequence $\int f_n d\mu$ converges to $\int f d\mu$.

<u>Proof</u>. We will let the reader verify that f must be μ-measurable, since in the applications we make of this corollary in these notes we will always know in advance that f is measurable. We remark next that if g and h are μ-measurable non-negative extended real-valued functions such that $g \geq h$ a.e. then it follows from Theorem 4.41 and Proposition 4.18e that $\int g d\mu \geq \int h d\mu$. Thus if $\int f_n d\mu = \infty$ for at least one n, then it is

clear that the sequence $\int f_n d\mu$ converges to $\infty = \int f d\mu$. On the other hand, if all of the f_n are integrable, then either the increasing sequence $\int f_n d\mu$ is unbounded, in which case it is clear again that it converges to $\infty = \int f d\mu$, or else it is bounded, in which case we can apply Corollary 4.43 to conclude that f also is integrable and that $\int f_n d\mu$ again converges to $\int f d\mu$.//

We remark that the corresponding statement for the case in which the sequence f_n is non-increasing is false in general unless at least one of the f_n is integrable (in which case Corollary 4.43 is applicable), for essentially the same reason as the fact that Proposition 2.17 is false unless at least one of the E_n has finite measure.

4.45 Theorem. (Fatou's Lemma). If μ is a non-negative measure, and if f_n is a sequence of non-negative integrable functions, then

$$\int_E (\lim\inf f_n) d\mu \leq \lim\inf \int f_n d\mu.$$

In particular, if the right hand is finite, then $\lim\inf f_n$ is integrable.

Proof. Let $g_n = \inf\{f_i : n \leq i < \infty\}$. Now g_n is the limit as m goes to ∞ of the decreasing sequence $\inf\{f_i : n \leq i \leq m\}$. But from Proposition 3.7 it follows that the infinum of a finite number of measurable functions is measurable. Thus g_n is measurable for each n. Since $\lim\inf f_n = \lim g_n$, we see that $\lim\inf f_n$ is also measurable. Now g_n is a non-decreasing sequence, and so $\lim \int g_n d\mu = \int \lim\inf f_n d\mu$ by Corollary 4.44. But $g_n \leq f_n$, and so $\int g_n d\mu \leq \int f_n d\mu$, for all n. Because the sequence $\int g_n d\mu$ is non-decreasing, we obtain the inequality

$$\lim\inf \int f_n d\mu \geq \lim \int g_n d\mu = \int \lim\inf f_n d\mu$$

as desired.//

Chapter 6 - Product Measures and Fubini's Theorem

Throughout this chapter we will let (X, S, μ) and (Y, T, ν) denote measure spaces with μ and ν non-negative (extended) real-valued. The objective of this chapter is to show how to define a measure, $\mu \otimes \nu$, on the σ-ring of subsets of $X \times Y$ generated by the sets $E \times F$, $E \in S$, $F \in T$, which has the property that

$$(\mu \otimes \nu)(E \times F) = \mu(E)\nu(F)$$

for all $E \in S$, $F \in T$. We then derive some relations between integration with respect to $\mu \otimes \nu$ and integration with respect to μ and ν.

A. Product Measures

6.1 Definition. We will let $S \times T$ denote the family of subsets of $X \times Y$ of the form $E \times F$ where $E \in S$ and $F \in T$. The elements of $S \times T$ will be called the measurable rectangles of $X \times Y$. We will denote the σ-ring generated by $S \times T$ by $S \otimes T$. The pair $(X \times Y, S \otimes T)$ will be called the product measurable space of the measurable spaces (X, S) and (Y, T).

If we let ρ be the function on $S \times T$ defined by $\rho(E \times F) = \mu(E)\nu(F)$ (with the convention that $0 \cdot \infty = 0$), the objective of this section is to show that we can extend ρ to a measure on $S \otimes T$. We do this by showing that the results of Chapter 1, particularly Theorem 1.33, are applicable to the set function ρ. In order to do this it will be useful to have the definition and some properties of "sections".

6.2 Definition. If $A \subseteq X \times Y$ and if $x \in X$, then the x-section of A, which is denoted by ${}^{x}A$, is defined by

$$^{x}A = \{y \in Y : (x, y) \in A\}.$$

Note that an x-section is a subset of Y. Similarly, if $y \in Y$, then the <u>y-section of A</u>, denoted by A^y, is defined by

$$A^y = \{x \in X : (x, y) \in A\}.$$

It is, of course, a subset of X.

<u>6.3 Proposition</u>. Let A_n be a sequence of subsets of $X \times Y$. Then $^{x}(\bigcup_{n=1}^{\infty} A_n) = \bigcup_{n=1}^{\infty} {}^{x}A_n$, $^{x}(\bigcap_{n=1}^{\infty} A_n) = \bigcap_{n=1}^{\infty} {}^{x}A_n$, and $^{x}(A_1 - A_2) = {}^{x}A_1 - {}^{x}A_2$. Similar results hold for y-sections.

<u>Proof</u>. Let $_{x}j$ denote the function from Y to $X \times Y$ defined by $_{x}j(y) = (x, y)$. For any $A \subseteq X \times Y$ it is clear that $^{x}A = ({}_{x}j)^{-1}(A)$. The proposition now follows immediately from the fact that the formation of preimages preserves unions, intersections and differences.//

<u>6.4 Definition</u>. If h is a function on $X \times Y$ and if $x \in X$, then the <u>x-section of h</u>, which is denoted by ^{x}h, is the function on Y defined by $(^{x}h)(y) = h(x, y)$. The definition of <u>y-sections of h</u>, which we denote by h^y, is similar.

<u>6.5 Proposition</u>. a) If g and h are B-valued functions on $X \times Y$ and if r and s are scalars, then for all $x \in X$ we have

$$^{x}(rg+sh) = r(^{x}g) + s(^{x}h).$$

b) If h_n is a sequence of functions on $X \times Y$ which converges to h pointwise, then for every x the sequence $^{x}h_n$ converges to ^{x}h pointwise.

c) If $A \subseteq X\ Y$, then ${}^x(\chi_A) = \chi_{({}^xA)}$.

Similar results hold for y-sections.

Proof. The proposition follows immediately from the observation that if h is a function on $X\times Y$ then ${}^xh = h\circ {}_xj$, where ${}_xj$ is the function defined in the proof of Proposition 6.3.//

6.6 Proposition. a) If $A \in S \otimes T$, then ${}^xA \in T$ for every $x \in X$, and $A^y \in S$ for every $y \in Y$.

b) If h is an $(S \otimes T)$-measurable function on $X\times Y$, then ${}^xh$ is T-measurable for every $x \in X$ and h^y is S-measurable for every $y \in Y$.

Proof. Let R be the set of those $A \in S \otimes T$ such that ${}^xA \in T$ for every $x \in X$ and $A^y \in S$ for every $y \in Y$. Clearly $S\times T \subseteq R$, and R is a σ-ring by Proposition 6.3. Therefore $R = S \otimes T$. Turning to the second part of the proposition, it follows from parts a) and c) of Proposition 6.5 that x-sections of simple $S \otimes T$-measurable functions are simple T-measurable functions, and similarly for y-sections. The second part of the proposition now follows from part b) of Proposition 6.5 and the definition of measurable functions.//

6.7 Lemma. $S\times T$ is a semiring.

Proof. This is a consequence of the following easily verified set theoretic equalities.

a) $\bigcap_{n=1}^{\infty} (E_n\times F_n) = (\bigcap_{n=1}^{\infty} E_n)\times(\bigcap_{n=1}^{\infty} F_n)$.

b) $E_1 \times F_1 - E_2 \times F_2 = [(E_1 - E_2) \times F_1] \oplus [(E_1 \cap E_2) \times (F_1 - F_2)]$

$$= [E_1 \times (F_1 - F_2)] \oplus [(E_1 - E_2) \times (F_1 \cap F_2)].//$$

6.8 Theorem. Let μ and ν be arbitrary non-negative measures. Then the set function ρ defined on $S \times T$ by $\rho(E \times F) = \mu(E)\nu(F)$ (with the convention that $0 \cdot \infty = 0$) is a premeasure, and so extends to a non-negative measure on $S \otimes T$.

Proof. We must show that ρ is countably additive. Suppose that $E \times F = \bigoplus_{n=1}^{\infty} E_n \times F_n$, where $E \times F$ and the $E_n \times F_n$ are in $S \times T$. We must show that $\rho(E \times F) = \sum_{n=1}^{\infty} \rho(E_n \times F_n)$. For each m let $A_m = \bigoplus_{n=1}^{m} E_n \times F_n$, and let $A = E \times F$. Then A_m increases to A, and so ${}^x\chi_{A_m}$ increases to ${}^x\chi_A$ for each x. If we use Definition 4.44 for the integral of non-negative functions which are not necessarily integrable, and apply Corollary 4.45, we find that $\int {}^x\chi_{A_m} d\nu$ increases to $\int {}^x\chi_A d\nu$ for each x. Evaluating these integrals, we find that $\sum_{n=1}^{m} \chi_{E_n}(x)\nu(F_n)$ increases to $\chi_E(x)\nu(F)$ for each x (where we must again use the convention that $0 \cdot \infty = 0$). Note that these functions may now take the value ∞, but are clearly measurable in the sense of Definition 4.44. Furthermore it is easily seen that if we apply the definition of the integral given in Definition 4.44, then we find that $\int (\chi_E \nu(F)) d\mu = \rho(E \times F)$ and $\int (\sum_{n=1}^{m} \chi_{E_n} \nu(F_n)) d\mu = \sum_{n=1}^{m} \rho(E_n \times F_m)$. Applying Corollary 4.45 again, we thus conclude that $\sum_{n=1}^{m} \rho(E_n \times F_m)$ converges to $\rho(E \times F)$ as m goes to ∞, so that ρ is a premeasure. Theorem 1.33 is now applicable.//

We invite the reader to contemplate the possibility of a proof of Theorem 6.8 which does not use integration theory (in the form of the Monotone Convergence Theorem) but only results of Chapter 1.

6.9 Definition. The measure obtained from ρ by applying Theorem 1.33 is called the product of the measures μ and ν. We will denote it by $\mu \otimes \nu$. The measure space $(X \times Y, S \otimes T, \mu \otimes \nu)$ is called the product of the measure spaces (X, S, T) and (Y, T, ν).

In order to know that the extension of ρ to a measure on $S \otimes T$ is unique (by applying Theorem 1.39) we must know that ρ is σ-finite.

6.10 Proposition. If μ and ν are both σ-finite, then so is ρ.

Proof. Let $E \times F \in S \times T$. Since μ and ν are assumed σ-finite, we have $E = \bigcup_{m=1}^{\infty} E_m$ and $F = \bigcup_{n=1}^{\infty} F_n$ where the E_m and the F_n have finite measure. Then $E \times F \subseteq \bigcup_{m,n} E_m \times F_n$ and $\rho(E_m \times F_n) < \infty$ for all m and n.//

From Proposition 6.10 and Theorem 1.39 we immediately obtain:

6.11 Proposition. If μ and ν are σ-finite then so is $\mu \otimes \nu$, and it is the only measure on $S \otimes T$ which has the property that $(\mu \otimes \nu)(E \times F) = \mu(E)\nu(F)$ for all $E \in S$ and $F \in T$.

6.12 Corollary. If μ and ν are arbitrary non-negative measures, and if E and F are σ-finite elements of S and T respectively, then the product of the restrictions of μ and ν to E and F respectively coincides with the restriction of $\mu \otimes \nu$ to $E \times F$.

<u>6.13 Proposition</u>. If μ and ν are both finite (or totally finite, or totally σ-finite), then so is $\mu \otimes \nu$.

<u>Proof</u>. Note that every element of $S \otimes T$ is contained in a measurable rectangle (since the collection of sets having this property is clearly a σ-ring containing $S \times T$). So if $A \in S \otimes T$, let $A \subseteq E \times F$ for some $E \in S$ and $F \in T$. Then if μ and ν are finite $(\mu \otimes \nu)(A) \leq (\mu \otimes \nu)(E \times F)$ $= \mu(E)\nu(F) < \infty$. In particular if $\mu(X) < \infty$ and $\nu(Y) < \infty$ then $(\mu \otimes \nu)(X \times Y) < \infty$, and similarly in the σ-finite case.//

B. Fubini's Theorem

In this section we examine the relation between integration with respect to $\mu \otimes \nu$ and integration with respect to μ and ν. The main result is known as Fubini's theorem.

In order to prove the main theorems we will need to make some finiteness assumptions. For this reason it is sufficient in the preliminary lemmas to assume that the measures μ and ν are finite.

Let R denote the collection of all finite disjoint unions of elements of $S \times T$. Then from Corollary 1.18 we see that R is just the ring generated by $S \times T$. The first step in the proof of Fubini's theorem is to prove a version of it just for this ring.

<u>6.14 Lemma</u>. Let μ and ν be finite. If A is any element of R (the ring generated by $S \times T$) then

1) ${}^x\chi_A$ and $\chi_A{}^y$ are measurable, in fact integrable, functions with respect to ν and μ respectively,
2) $\int {}^x\chi_A d\nu$ and $\int \chi_A{}^y d\mu$ are measurable, in fact integrable, functions of x and y respectively, and
3) $\int(\int {}^x\chi_A d\nu) d\mu(x) = \int \chi_A d(\mu \otimes \nu) = \int(\int {}^y\chi_A d\mu) d\nu(y)$.

<u>Proof</u>. Let $A = \bigoplus_{i=1}^{n} E_i \times F_i$ where $E_i \otimes F_i \in S \times T$ for each i. It follows that $\chi_A(x, y) = \sum_{i=1}^{n} \chi_{E_i}(x)\chi_{F_i}(y)$ for all x and y. From this it is clear that each section is an ISF, which proves 1). It is also clear that $\int {}^x\chi_A d\nu = \sum_{i=1}^{n} \chi_{E_i}(x)\nu(F_i)$, and similarly for the integral with respect

to μ, and so they too are both ISF, proving 2). Finally, it is now clear that both iterated integrals turn out to be $\sum_{i=1}^{n} \mu(E_i)\nu(F_i)$, that is, $\int \chi_A d(\mu \otimes \nu)$.//

The main step in the proof of Fubini's theorem is to extend Lemma 6.14 to the case in which A is an arbitrary element of $S \otimes T$. But it is sufficient to still work with finite measures.

<u>6.15 Key Lemma</u>. Let μ and ν be finite and let $A \in S \otimes T$. Then

1) ${}^x\chi_A$ and χ_A^y are measurable, in fact integrable, functions with respect to μ and ν respectively.
2) $\int {}^x\chi_A d\nu$ and $\int \chi_A^y d\mu$ are measurable, in fact integrable, functions of x and y respectively.
3) $\int(\int {}^x\chi_A d\nu) d\mu(x) = \int \chi_A d(\mu \otimes \nu) = \int(\int \chi_A^y d\mu) d\nu(y)$.

<u>Proof</u>. By Proposition 6.5c) and 6.6a), the functions in 1) are characteristic functions of measurable sets, and so, since μ and ν are finite measures, they are integrable. This proves property 1).

Note that because the integrands in property 2) are characteristic functions and the measures are finite, the two functions in property 2) are bounded. Thus if we can prove that these functions are measurable, it will follow that they are integrable.

Let M be the collection of all sets $A \in S \otimes T$ for which properties 1), 2) and 3) are true. We wish to show that $M = S \otimes T$. Let R be the set of finite disjoint unions of elements of $S \times T$. (so that by Corollary 1.18 R is the ring generated by $S \times T$). Then Lemma 6.14 says exactly that

$R \subseteq M$. Let us investigate further the properties of M.

Suppose that A_n is a sequence of elements of M, and that A_n decreases to a set A. Then $A \in S \otimes T$ and so property 1) holds for A. Now ${}^x\chi_{A_n}$ and $\chi^y_{A_n}$ will decrease to ${}^x\chi_A$ and χ^y_A, and so we can apply Corollary 4.43 of the Monotone Convergence Theorem to conclude that

$$\int {}^x\chi_{A_n} d\nu \text{ decreases to } \int {}^x\chi_A d\nu \text{ and}$$

$$\int \chi^y_{A_n} d\mu \text{ decreases to } \int \chi^y_A d\mu.$$

Since property 2) holds for the sets A_n, it follows that the functions $\int {}^x\chi_A d\nu$ and $\int \chi^y_A d\mu$ are measurable. Thus property 2) holds for A also. Using Corollary 4.43 again, we see that

$$\int(\int {}^x\chi_{A_n} d\nu) d\mu(x) \text{ decreases to } \int(\int {}^x\chi_A d\nu) d\mu(x) \text{ and}$$

$$\int(\int \chi^y_{A_n} d\mu) d\nu(y) \text{ decreases to } \int(\int \chi^y_A d\mu) d\nu(y).$$

Since property 3) holds for the sets A_n, the left hand sides are just $\int \chi_{A_n} d(\mu \otimes \nu)$. But $\int \chi_{A_n} d(\mu \otimes \nu)$ decreases to $\int \chi_A d(\mu \otimes \nu)$ by Proposition 4.43, and so property 3) holds for A also. A similar proof shows that if instead the sequence A_n increases to A, then again $A \in M$.

We are thus led to make the following definition:

<u>6.16 Definition</u>. A collection M of sets is called a <u>monotone class</u> if it is closed under the formation of countable increasing unions and countable decreasing intersections.

Thus what we have shown above is that the collection M of the proof of Lemma 6.15 is a monotone class which contains the ring generated by the semiring S×T. It is then clear that the proof of Lemma 6.15 will be completed once we have proven the following lemma:

6.17 Lemma (The Lemma on Monotone Classes). Let M be a monotone class and let P be a semiring. If M contains the ring generated by P, then M contains the σ-ring generated by P.

Proof. Note first that the intersection of any collection of monotone classes is again a monotone class, so that any collection of sets is contained in a smallest monotone class, which it is said to generate. Thus we can (and will) assume that M is the monotone class generated by the ring, R, generated by P. Thus we wish to show that M coincides with the σ-ring, S(P), generated by P. For this it suffices to show that M is a σ-ring, and for this it suffices to show that M is closed under taking differences and finite unions (since $\bigcup_{n=1}^{\infty} F_n = \bigcup_{m=1}^{\infty} (\bigcup_{n=1}^{m} F_n)$ which is an increasing union).

Thus for each $E \in M$ define a subset, K(E), of M by

$$K(E) = \{F \in M : E - F, F - E, \text{ and } E \cup F \text{ are in } M\}.$$

What we need to show is that $K(E) = M$ for all $E \in M$. We divide the proof of this fact into 6 short steps.

1) If $F \in K(E)$ then $E \in K(F)$. This is clear from the definitions of K(E) and K(F).

2) If $E \in R$ then $R \subseteq K(E)$. This is clear from the definition of a ring.

3) $K(E)$ is a monotone class for every $E \in M$. To see this, suppose that F_n is a sequence of elements of $K(E)$ which increases to a set F. Then $F_n - E$ increases to $F - E$, $E - F_n$ decreases to $E - F$, and $F_n \cup E$ increases to $F \cup E$, so that $F - E$, $E - F$ and $E \cup F$ are in M. Thus $F \in K(E)$. A similar argument works if instead the sequence F_n decreases to F.

4) If $E \in R$ then $K(E) = M$. This follows from 2) and 3) and the fact that we have assumed that M is the monotone class generated by R.

5) $R \subseteq K(E)$ for all $E \in M$. This follows from 4) and 1).

6) $K(E) = M$ for all $E \in M$. This follows from 5) and 3).//

The next step in the proof of Fubini's theorem is to extend Lemma 6.15 to the case of non-negative measurable functions. We do not need to assume that μ and ν are finite any more, but we do need to assume that they are σ-finite (see exercise). Because we will want to use Corollary 4.43 of the Monotone convergence theorem we need the following fact:

<u>6.18 Proposition</u>. If f is a non-negative measurable function, then there is an increasing sequence of non-negative simple measurable functions which converges to f pointwise.

<u>Proof</u>. Define f_n by

$$f_n(x) = \begin{cases} (k-1)/2^n & \text{if } (k-1)/2^n \leq f(x) < k/2^n \quad k = 1,\dots,n2^n \\ n & \text{if } n \leq f(x). \end{cases}$$

<u>6.19 Theorem (Fubini, Tonelli)</u>. Let f be a non-negative $S \otimes T$-measurable function. Then the following conditions are equivalent:

1) f is integrable,

2) ${}^x f$ is integrable for almost all $x \in X$, and $x \mapsto \int {}^x f d\nu$ is an (almost everywhere defined) μ-integrable function on X,

3) f^y is integrable for almost all $y \in Y$, and $y \mapsto \int f^y d\mu$ is an (almost everywhere defined) ν-integrable function on Y.

If any of these three conditions holds then

$$\int(\int {}^x f d\nu)d\mu = \int f d(\mu \otimes \nu) = \int(\int f^y d\mu)d\nu.$$

<u>Proof</u>. Since f is a measurable function, $C(f)$ is an $S \otimes T$-measurable set. Since we are assuming that the measures μ and ν are σ-finite, it is easy to see that there is a rectangle $E \times F \in S \times T$ and a sequence $E_n \times F_n$ of elements of $S \times T$ such that $C(f) \subseteq E \times F$, the sequence $E_n \times F_n$ increases to $E \times F$, and $\mu(E_n) < \infty$, $\nu(F_n) < \infty$ for all n. Using Proposition 6.18 let f_n be an increasing sequence of non-negative simple $S \otimes T$-measurable functions which converges to f pointwise. If we let $g_n = \chi_{E_n \times F_n} f_n$ for each n, then the g_n form a sequence of non-negative ISF increasing to f pointwise.

For the moment fix n, and consider μ, ν and $\mu \otimes \nu$ restricted to E_n, F_n and $E_n \times F_n$ respectively. According to Proposition 6.13 the product of the restrictions of μ and ν to E_n and F_n respectively coincides with $\mu \otimes \nu$ restricted to $E_n \times F_n$, and all these restrictions are finite. We can thus apply key Lemma 6.15 to conclude that conditions 2) and 3) and the equality between integrals in Theorem 6.19 are true for the characteristic function of any measurable subset of $E_n \times F_n$, in fact true even with the qualification "almost" omitted. Now since $C(g_n) \subseteq E_n \times F_n$ and g_n is just a finite sum of characteristic functions of measurable subsets of $E_n \times F_n$, it follows that conditions 2) and 3) and the equality between integrals in Theorem 6.19 are true for g_n also, again even with the qualification "almost" omitted. In particular, $\int {}^x g_n d\nu$ is an integrable function of x for each n, and the $\int {}^x g_n d\nu$ form an increasing sequence of non-negative integrable functions.

Now if condition 1) holds, then

$$\int(\int {}^x g_n d\nu)d\mu = \int g_n d(\mu \otimes \nu) \leq \int f d(\mu \otimes \nu)$$

for all n. On the other hand, if condition 2) holds, then $\int {}^x g_n d\nu \leq \int {}^x f d\nu$ a.e., and so

$$\int(\int {}^x g_n d\nu)d\mu \leq \int(\int {}^x f d\nu)d\mu$$

for all n. Thus in either case the sequence of the L^1-norms of the functions $\int {}^x g_n d\nu$ is bounded above, and so we can apply the Monotone Convergence Theorem (Theorem 4.42) to conclude that the sequence $\int {}^x g_n d\nu$ converges a.e. and in mean to an integrable function, h. Let N be the null set off of which $\int {}^x g_n d\nu$ converges to $h(x)$ for all x. If $x \notin N$ then $\int {}^x g_n d\nu \leq h(x)$ for all n, and so the sequence of the norms of the ${}^x g_n$ is bounded. But ${}^x g_n$ increases to ${}^x f$, and so ${}^x f$ is integrable and $\int {}^x f d\nu = \lim \int {}^x g_n d\nu$ for all $x \notin N$ by Corollary 4.43. (Thus we see that N is exactly the set of those $x \in X$ for which ${}^x f$ is not integrable.)

It follows that $\int {}^x f d\nu = h(x)$ for all $x \in N$, and in particular that the function $\int {}^x f d\nu$ (defined off of N) is integrable. Thus if it is condition 1) which holds, then we see that we have shown that it follows that condition 2) holds also.

Now since $\int {}^x f d\nu = h(x)$ for $x \notin N$, it follows that the sequence of functions $\int {}^x g_n d\nu$ converges to $\int {}^x f d\nu$ in mean, so that

$$\int (\int {}^x f d\nu) d\mu = \lim \int (\int {}^x g_n d\nu) d\mu = \lim \int g_n d(\mu \otimes \nu).$$

Suppose now that it is condition 2) which holds. Then we see from the above equation that the sequence of L^1-norms of the g_n is bounded above, and so, since g_n increases to f, that f is integrable and g_n converges to f in mean by Corollary 4.43. In particular, condition 1) is seen to follow. Finally if either condition, so both, hold, we see that

$$\int f d(\mu \otimes \nu) = \lim \int g_n d(\mu \otimes \nu) = \int (\int {}^x f d\nu) d\mu.$$

Of course an entirely parallel argument shows that conditions 1) and 3) are equivalent and that the second equality between integrals holds.//

We remark that the above proof could have been shortened somewhat by allowing functions to take the value ∞, and in particular by defining $\int {}^x f d\nu$ to be ∞ whenever ${}^x f$ is not integrable. But this seems to obscure somewhat the role of the null set N, which is an essential aspect of this theorem. Furthermore, the final version of Fubini's theorem which we shall give involves functions with values in a Banach space and again there will be a null set, N, on which ${}^x f$ is not integrable. But in this setting it is no longer natural to set $\int {}^{x}f d\nu$ equal to ∞ for $x \in N$. Thus even in the present setting we prefer simply to say that $\int {}^{x}f d\nu$ is undefined for $x \in N$.

We also remark that the part of the above theorem which is usually called Tonelli's theorem is the fact that conditions 2) or 3) imply condition 1). For applications this is by far the most useful method of trying to show that a measurable function on a product space is integrable. Notice that Tonelli's theorem is immediately applicable to functions with values in a Banach space, since if f is such a function and is measurable, then to show that f is integrable it suffices by Theorem 4.41 to show that the non-negative measurable function $\|f(\cdot)\|$ is integrable. A counter-example for Tonelli's theorem when one of the measures μ and ν is not σ-finite can be found in exercise .

Notation which is more commonly used in stating Fubini's theorem than that which we have used above is as follows:

6.20 Definition. If f is a measurable function on $X \times Y$, if $\int {}^{x}f d\nu$ is defined a.e. and measurable, and if $\int(\int {}^{x}f d\nu)d\mu$ exists, then we will write $\int f(x, y)d\nu(y)d\mu(x)$ or $\int f d\nu d\mu$ instead of $\int(\int {}^{x}f d\nu)d\mu$. Under similar conditions we will write $\int f(x, y)d\mu(x)d\nu(y)$ or $\int f d\mu d\nu$ instead of $\int(\int f^{y}d\mu)d\nu$. These expressions are called iterated integrals to distinguish them from $\int f d(\mu \otimes \nu)$, which is called the double integral of f.

From Fubini's theorem we obtain the following convenient characterization of null sets with respect to $\mu \otimes \nu$:

6.21 Corollary. If $C \in N(\mu \otimes \nu)$ (the collection of null sets for $\mu \otimes \nu$), then ${}^{x}C \in N(\nu)$ for almost all $x \in X$ and $C^{y} \in N(\mu)$ for almost all $y \in Y$. Conversely if $A \in S \otimes T$ and ${}^{x}A \in N(\nu)$ a.e. (or $A^{y} \in N(\mu)$ a.e.), then $A \in N(\mu \otimes \nu)$.

Proof. Since $C \in N(\mu \otimes \nu)$, there exists $A \in S \otimes T$ such that $C \subseteq A$ and $(\mu \otimes \nu)(A) = 0$. By Theorem 6.19 we have

$$\int(\int {}^{x}\chi_{A} d\nu) d\mu = \int \chi_{A} d(\mu \otimes \nu) = (\mu \otimes \nu)(A) = 0.$$

Thus $\nu({}^{x}A) = \int {}^{x}\chi_{A} d\nu = 0$ for almost all $x \in X$. Since ${}^{x}C \subseteq {}^{x}A$ for all $x \in X$ we are done. The converse is clear.//

We now come to our final version of Fubini's theorem, which involves functions with values in a Banach space.

6.22 Theorem (Fubini). If μ and ν are σ-finite non-negative measures and if f is a B-valued $(\mu \otimes \nu)$-integrable function, then

1) ${}^{x}f$ and f^{y} are integrable functions for almost all $x \in X$ and almost all $y \in Y$,
2) $\int {}^{x}f d\nu$ and $\int f^{y} d\mu$ are (almost everywhere defined) integrable functions of x and y respectively, and
3) $\iint f d\nu d\mu = \int f d(\mu \otimes \nu) = \iint f d\mu d\nu$.

Proof. Since f is $(\mu \otimes \nu)$-integrable, $\|f(\cdot)\|$ is a non-negative $(\mu \otimes \nu)$-integrable function. Choose (as in the proof of Theorem 4.41) a sequence, f_n, of simple $S \otimes T$-measurable functions which converges to f a.e. and such that $\|f_n\| \leq 2\|f\|$ for all n. Then by the Lebesgue Dominated

Convergence Theorem (Theorem 4.40) the sequence f_n is a mean Cauchy sequence of ISF which converges to f in mean. In particular,

$$\int f d(\mu \otimes \nu) = \lim \int f_n d(\mu \otimes \nu).$$

Now by Theorem 6.19, ${}^x(\|f\|) = \|f(x, \cdot)\| = \|{}^x f\|$ is a ν-integrable function on Y for almost all $x \in X$, say for $x \in N_1 \in N(\mu)$. It follows from Theorem 4.41 that part 1) is proved. Let C be the null set off of which f_n converges to f pointwise. By the definition of the f_n's the sequence ${}^x f_n$ converges a.e. to ${}^x f$ for all $x \notin N_2$, where $N_2 \in N(\mu)$ is the set of those $x \in X$ for which $\nu({}^x C) \neq 0$ (see Corollary 6.21). Therefore, since $\|{}^x f_n\| \leq 2\|{}^x f\|$, we can use the Lebesgue Dominated Convergence Theorem again to conclude that ${}^x f_n$ and ${}^x f$ are ν-integrable and that $\int {}^x f_n d\nu$ converges to $\int {}^x f d\nu$ for all $x \notin N_1 \cup N_2$. Now from key Lemma 6.15 it follows (as in the proof of Theorem 6.19) that the present theorem is true for ISF, and so in particular, $\int {}^x f_n d\nu$ is a measurable (and, in fact, integrable) function on X. Since

$$\left\|\int {}^x f_n d\nu\right\| \leq 2 \int {}^x(\|f\|) d\nu$$

for all $x \notin N_1$ and all n, and since by Theorem 6.19 the right hand side is a μ-integrable function of x, we can again apply the Lebesgue Dominated Convergence Theorem to conclude that $\int {}^x f d\nu$ is a μ-integrable function on X (proving part 2)), and that

$$\iint f d\nu d\mu = \lim \iint f_n d\nu d\mu.$$

But as was noted before, Fubini's Theorem holds for the ISF, and so

$$\int f d(\mu \otimes \nu) = \lim \int f_n d(\mu \otimes \nu) = \lim \iint f_n d\nu d\mu.$$

Thus part 3) and the theorem are proved.//

We remark again that in most applications of this theorem one will first have to invoke Tonelli's theorem to show that f is integrable.

We conclude this chapter by indicating how to extend Fubini's theorem to the case in which μ and ν are complex-valued measures. To begin with, when μ is a real measure, define μ^+ and μ^- by

$$\mu^+ = \frac{|\mu| + \mu}{2} \quad \text{and} \quad \mu^- = \frac{|\mu| - \mu}{2}.$$

Then μ^+ and μ^- are easily seen to be non-negative measures and $\mu = \mu^+ - \mu^-$. Thus if both μ and ν are real measures it is natural to define $\mu \otimes \nu$ by

$$\mu \otimes \nu = \mu^+ \otimes \nu^+ - \mu^+ \otimes \nu^- - \mu^- \otimes \nu^+ + \mu^- \otimes \nu^-.$$

It is also easy to show that this is what the product measure should be, in the sense that it does the right thing on rectangles, and that by linearity Fubini's theorem is true with this definition of $\mu \otimes \nu$.

Similarly, if μ is a complex measure, define μ_r and μ_i by

$$\mu_r = \frac{\mu + \bar{\mu}}{2} \quad \text{and} \quad \mu_i = \frac{\mu - \bar{\mu}}{2i}$$

(where the bar denotes complex conjugation and $\bar{\mu}$ is defined by $\bar{\mu}(E) = (\mu(E))^-$). Then μ_r and μ_i are real measures and $\mu = \mu_r + i\mu_i$. Thus if μ and ν are both complex measures, we define $\mu \otimes \nu$ by

$$\mu \otimes \nu = \mu_r \otimes \nu_r - \mu_i \otimes \nu_i + i(\mu_r \otimes \nu_i + \mu_i \otimes \nu_r).$$

Again it is easily seen that this is the appropriate product measure and that Fubini's theorem holds for it.