PSH, KSP, KMJ, PBU

수리통계학 강의록

Mathematical Statistics Lecture Note

3rd DRAFT

September 25, 2019

Seoul National University
Department of Mathematics & Statistics

함께했던 날들과,
함께할 날들을 위해.
2018년의 끝자락, 관악에서.

Acronyms

WLOG 일반성을 잃지 않고 (Without Loss Of Generality)
TFAE 다음은 서로 동치이다. (The Followings Are Equivalent.)
ow. 그렇지 않으면 (OtherWise)
ae. 거의 어디서나 (Almost Everywhere)
LUBP Least Upper Bound Property
GLBP Greatest Lower Bound Property
MSP Monotone Sequence Property
NSP Nested Set Property
MCT 단조수렴정리 (Monotone Convergence Theorem)
DCT Lebesgue의 지배수렴정리 (Lebesgue's Dominated Convergence Theorem)
FTC 미적분학의 기본정리 (Funcdamental Theorem of Calculus)
MVT 평균값 정리 (Mean Value Theorem)
IVT 중간값 정리 (Intermediate Value Theorem)
INFT 역함수 정리 (INverse Function Theorem)
rv. 확률변수 (Random Variable) 혹은 확률벡터 (Random Vector)
PDF 확률밀도함수 (Probability Density Function)
PMF 확률질량함수 (Probability Mass Function)
CDF 누적분포함수 (Cumulative Distribution Function)
MGF 적률생성함수 (Moment Generating Function)
CF 특성함수 (Characteristic Function)
CGF 누율생성함수 (Cumulant Generating Function)
PGF 확률생성함수 (Probability Generating Function)
iid. Independent and Identically Distributed rv.
io. Infinitely Often

as. 거의 확실하게 (Almost Surely)

Contents

Chapter 0

Preliminary

0.1 Analysis

0.1.1 Inequalities

다음 부등식들은 각종 정리의 증명에서 유용하게 사용된다.

Theorem 0.1 (Young's inequality) 임의의 $x, y \geq 0$와 $1/p + 1/q = 1$인 $p, q > 1$에 대해 $xy \leq x^p/p + y^q/q$이다. 이때, 등호가 성립할 필요충분조건은 $x^p = y^q$인 것이다.

PROOF 임의의 $y \geq 0$를 고정하고 함수 $f : \mathbb{R} \to \mathbb{R}$를 $f : x \mapsto x^p/p + y^q/q - xy$로 두면 $f' : x \mapsto x^{p-1} - y$에서 f는 $y^{1/(1-p)} = y^{q/p}$에서 최솟값 $f(y^{q/p}) = (1/p + 1/q)y^q - y^{q/p+1} = 0$을 가진다. 여기서 마지막 등호는 $q/p + 1 = q$에서 성립하고, 증명은 이로써 충분하다. □

Theorem 0.2 (Hölder's inequality) 임의의 $x, y \in \mathbb{R}^n$와 $1/p + 1/q = 1$인 $p, q > 1$에 대해 $\sum_{i=1}^n |x_i y_i| \leq ||x||_p ||y||_q$이다. 이때, 등호가 성립할 필요충분조건은 모두 0은 아닌 적당한 $a, b \in \mathbb{R}$가 존재하여 각 $i \leq n$에 대해 $a|x_i|^p + b|y_i|^q = 0$인 것이다.

PROOF WLOG, 필요하다면 x, y를 각각 $|x|, |y|$로 대체하여 x, y의 모든 성분이 처음부터 음이 아니라 해도 된다. 한편, 만약 $||x||_p = 0$이면 $x = 0$에서 부등식이 자명하므로 $||x||_p > 0$이라 하고, 비슷한 이유로 $||y||_p > 0$이라 하자. 이제 $||x||_p = ||y||_q = 1$인 특별한 경우를 생각하면 Young의 부등식으로부터 $\sum_{i=1}^n x_i y_i \leq \sum_{i=1}^n (x_i^p/p + y_i^q/q) = ||x||_p^p/p + ||y||_q^q/q = 1/p + 1/q = 1$이다. 이어서 일관적인 경우에 $\widetilde{x} = x/||x||_p, \widetilde{y} = y/||y||_q$라 하면 $||\widetilde{x}||_p = ||\widetilde{y}||_q = 1$이므로 앞선 결과로부터 $\sum_{i=1}^n \widetilde{x}_i \widetilde{y}_i \leq 1$이 되어 곧 $\sum_{i=1}^n x_i y_i \leq ||x||_p ||y||_q$임을 안다.

다음으로, 등호조건을 보이기 위해 $\sum_{i=1}^n |x_i y_i| = ||x||_p ||y||_q$라 하자. 만약 $||x||_p = 0$이면 $x = 0$에서 각 $i \leq n$에 대해 $|x_i|^p + 0 \cdot |y_i|^q = 0$이 되어 더 이상 보일 것이 없으므로 $||x||_p \neq 0$

이라 하고, 비슷한 이유에서 $||y||_q \neq 0$이라 하자. 이제 $\widetilde{x} = |x|/||x||_p, \widetilde{y} = |y|/||y||_q$라 하면 $\sum_{i=1}^n \widetilde{x}_i\widetilde{y}_i = (\sum_{i=1}^n x_iy_i)/||x||_p||y||_q = 1 = ||\widetilde{x}||_p^p/p + ||\widetilde{y}||_q^q/q = \sum_{i=1}^n (\widetilde{x}_i^p/p + \widetilde{y}_i^q/q)$이므로 Young의 부등식으로부터 각 $i \leq n$에 대해 $\widetilde{x}_i\widetilde{y}_i = \widetilde{x}_i^p/p + \widetilde{y}_i^q/q$이다. 이는 다시 Young의 부등식의 등호조건으로부터 각 $i \leq n$에 대해 $\widetilde{x}_i^p = \widetilde{y}_i^q$임을 뜻하고, 따라서 $||y||_q^q|x_i|^p - ||x||_p^p|y_i|^q = 0$이다. 역으로 모두 0은 아닌 적당한 $a, b \in \mathbb{R}$가 존재하여 각 $i \leq n$에 대해 $a|x_i|^p + b|y_i|^q = 0$인 경우에 등호가 성립함은 자명하므로 증명은 이로써 충분한다. □

Corollary 0.3 (Hölder's inequatliy) 수열 $\{x_i\}, \{y_i\}$와 $1/p + 1/q = 1$인 $p, q > 1$에 대해 $\sum_{i=1}^\infty |x_i|^q, \sum_{i=1}^\infty |y_i|^p < \infty$이면 $\sum_{i=1}^\infty |x_iy_i| < \infty$이고 $\sum_{i=1}^\infty |x_iy_i| \leq (\sum_{i=1}^\infty |x_i|^p)^{1/p}(\sum_{i=1}^\infty |y_i|^q)^{1/q}$이다.

Proof 임의의 $k \in \mathbb{N}$에 대해 앞서 보인 Hölder의 부등식으로부터

$$\begin{aligned}
\sum_{i=1}^k |x_iy_i| &\leq \left(\sum_{i=1}^k |x_i|^p\right)^{1/p}\left(\sum_{i=1}^k |y_i|^q\right)^{1/q} \\
&\leq \left(\sum_{i=1}^\infty |x_i|^p\right)^{1/p}\left(\sum_{i=1}^\infty |y_i|^q\right)^{1/q} \\
&< \infty
\end{aligned}$$

이므로 MSP에서 $\sum_{i=1}^\infty |x_iy_i|$는 수렴하고 $\sum_{i=1}^\infty |x_iy_i| \leq (\sum_{i=1}^\infty |x_i|^p)^{1/p}(\sum_{i=1}^\infty |y_i|^q)^{1/q}$이다. □

Hölder의 부등식에서 조건으로 주어진 '$1/p + 1/q = 1$인 실수 $p, q > 1$' 는 다른 부등식에서도 종종 조건으로 주어지곤 하기에 특별히 이러한 $p, q \in \mathbb{R}$에 대해 **Hölder conjugate**라는 이름이 붙어있다.

Theorem 0.4 (Minkowski's inequality) 임의의 $x, y \in \mathbb{R}^n$와 $p \geq 1$에 대해 $||x+y||_p \leq ||x||_p + ||y||_p$이다. 이때, 등호가 성립한 필요충분조건은 모두 0은 아닌 적당한 $a, b \in \mathbb{R}$가 존재하여 각 $i \leq n$에 대해 $a|x_i|^p + b|y_i|^q = 0$인 것이다.

Proof 만약 $p = 1$이면 $||x+y||_1 = \sum_{i=1}^n |x_i + y_i| \leq \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| = ||x||_1 + ||y||_1$에서 부등식이 자명하므로 $p > 1$이라 하자. 또한 $||x+y||_p = 0$인 경우에도 부등식이 자명하므로 $||x+y||_p > 0$이라 하자. 그렇다면 Hölder의 부등식으로부터 $1/p + 1/q = 1$인 $q > 1$에 대해

$$\begin{aligned}
||x+y||_p^p &= \sum_{i=1}^n |x_i + y_i|^p \\
&\leq \sum_{i=1}^n |x_i||x_i + y_i|^{p-1} + \sum_{i=1}^n |y_i||x_i + y_i|^{p-1} \\
&\leq ||x||_p\left(\sum_{i=1}^n |x_i + y_i|^{(p-1)q}\right)^{1/q} + ||y||_p\left(\sum_{i=1}^n |x_i + y_i|^{(p-1)q}\right)^{1/q}
\end{aligned}$$

$$= (||x||_p + ||y||_p)\left(\sum_{i=1}^n |x_i + y_i|^p\right)^{(p-1)/p}$$
$$= (||x||_p + ||y||_p)||x+y||_p^{p-1}$$

이므로 $||x+y||_p \leq ||x||_p + ||y||_p$ 임을 안다.

다음으로, 등호조건을 보이기 위해 $||x+y||_p = ||x||_p + ||y||_p$ 라 하자. 만약 $x = 0$이면 각 $i \leq n$에 대해 $|x_i|^p + 0 \cdot |y_i|^p = 0$에서 더 이상 보일 것이 없으므로 $x \neq 0$이라 하고, 비슷한 이유에서 $y \neq 0$이라 하자. 그렇다면 위의 증명과정과 Hölder의 부등식의 등호조건으로부터 적당한 $a_1, a_2, b_1, b_2 \in \mathbb{R}$가 존재하여 $a_1|x_i|^p + b_1|x_i + y_i|^{(p-1)q} = a_1|x_i|^p + b_1|x_i + y_i|^p = 0$이고 $a_2|y_i|^p + b_2|x_i + y_i|^{(p-1)q} = a_2|y_i|^p + b_2|x_i + y_i|^p = 0$이다. 여기서 만약 $a_1 = 0$이면 $x + y = 0$이 되어 $||x||_p + ||y||_p = ||x+y||_p = 0$에서 $x = y = 0$의 모순이 발생하므로 $a_1 \neq 0$이고 $b_1 = 0$이면 $x = 0$의 모순이 발생하므로 $b_1 \neq 0$이다. 비슷하게 $a_2, b_2 \neq 0$이므로 이상을 종합하면 $a_1b_2|x_i|^p - a_2b_1|y_i|^p = 0$이고 여기서 $a_1b_2, a_2b_1 \neq 0$이다. 역으로 모두 0은 아닌 적당한 $a, b \in \mathbb{R}$가 존재하여 각 $i \leq n$에 대해 $a|x_i|^p + b|y_i|^q = 0$인 경우에 등호가 성립함은 자명하므로 증명은 이로써 충분하다. □

Corollary 0.5 (Minkowski's inequatliy) 수열 $\{x_i\}, \{y_i\}$와 $p \geq 1$에 대해 $\sum_{i=1}^\infty |x_i|^q, \sum_{i=1}^\infty |y_i|^p < \infty$이면 $\sum_{i=1}^\infty |x_i + y_i|^p < \infty$이고 $(\sum_{i=1}^\infty |x_i + y_i|^p)^{1/p} \leq (\sum_{i=1}^\infty |x_i|^p)^{1/p} + (\sum_{i=1}^\infty |y_i|^q)^{1/q}$이다.

Proof 임의의 $k \in \mathbb{N}$에 대해 앞서 보인 Minkowski의 부등식으로부터

$$\left(\sum_{i=1}^k |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^k |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^k |y_i|^q\right)^{1/q}$$
$$\leq \left(\sum_{i=1}^\infty |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^\infty |y_i|^q\right)^{1/q}$$
$$< \infty$$

이므로 MSP에서 $\sum_{i=1}^\infty |x_i + y_i|^p$는 수렴하고 $(\sum_{i=1}^\infty |x_i + y_i|^p)^{1/p} \leq (\sum_{i=1}^\infty |x_i|^p)^{1/p} + (\sum_{i=1}^\infty |y_i|^q)^{1/q}$이다. □

0.1.2 Norms

노름은 벡터공간에서 정의되는 '크기' 의 추상화이다. 이미 해석개론 전반부에서 어느정도 논하였을 것이므로 여기서는 앞으로 사용될 몇몇 특별한 노름들을 간단히 소개하고 다변수 해석학에 필요한 작용소 노름과 행렬 노름에 대해 집중적으로 살펴본다.

Definition 0.6 벡터공간 V에서 정의된 함수 $||\cdot|| : \mathsf{V} \to \mathbb{R}_0^+$가 임의의 $\lambda \in \mathbb{R}$와 임의의 $v, w \in \mathsf{V}$에 대해

i. (양의 동차성) $||\lambda v|| = |\lambda|||v||$.
ii. (삼각부등식) $||v+w|| \leq ||v|| + ||w||$.
iii. (양의 정부호성) $||v|| = 0$일 필요충분조건은 $v = 0$인 것이다.

를 만족하면 이때의 함수 $||\cdot||$를 V 위의 **노름 (norm)**이라 하고, tuple $(\mathsf{V}, ||\cdot||)$를 **노름공간 (norm space)**이라 한다.

첫째로 소개할 노름은 p−노름으로 사실상 유클리드 공간에서의 표준으로 자리잡은 노름이다. 단순한 정의임에도 불구하고 그 기저에 깔린 훌륭한 기하학적 직관과 이가 가지는 좋은 성질들 덕분에 유클리드 공간을 넘어 수열공간이나 함수공간으로 그 정의가 확장되어 사용되지만, 무한차원 벡터공간에서는 조금 귀찮은 일들이 생기는 까닭에 여기서는 n차원 벡터공간에 한하여 p−노름을 도입한다. 이후 함수공간에서의 p−노름에 대해서는 측도론을 배우면서 L^p 공간의 개념을 통해 엄밀히 논의할 것이다. 특별히 정하지 않는 이상, $\mathbb{R}^n$에는 2−노름이 장착된 것으로 생각한다.

Definition 0.7 n차원 벡터공간 V와 $p \geq 1$에 대해 **p−노름 (p−norm)**을 $||\cdot||_p : \mathsf{V} \to \mathbb{R}_0^+$로 쓰고 $||\cdot||_p : v \mapsto (\sum_{i=1}^n |v_i|^p)^{1/p}$로 정의한다. 여기서 $v_1, \cdots, v_n$은 V의 고정된 기저에 대한 v의 좌표이다.

Proposition 0.8 n차원 벡터공간 V와 $p \geq 1$에 대해 p−노름은 노름이다. 따라서 $(\mathsf{V}, ||\cdot||_p)$는 노름공간을 이룬다.

Proof 양의 동차성과 양의 정부호성은 p−노름의 정의로부터 명백하고, 삼각부등식도 Minkowski의 부등식으로부터 자명하다. □

다음으로 소개할 노름은 p−노름만큼은 아니지만 각종 증명에서 이따금씩 등장하는 최댓값 노름이다. 아래의 정리에서 볼 수 있듯이 표기가 p−노름과 유사한 것은 우연의 일치가 아니며, 이러한 이유로 최댓값 노름을 p−노름의 $p = \infty$인 특별한 경우로 보는 경우도 있다.

Definition 0.9 n차원 벡터공간 V에 대해 **최댓값 노름 (maximum norm)**을 $||\cdot||_\infty : \mathsf{V} \to \mathbb{R}_0^+$로 쓰고 $||\cdot||_\infty : v \mapsto \max_{i=1}^n |v_i|$로 정의한다. 여기서 $v_1, \cdots, v_n$은 V의 고정된 기저에 대한 v의 좌표이다.

Proposition 0.10 n차원 벡터공간 V에 대해 최댓값 노름은 노름이다. 따라서 $(\mathsf{V}, ||\cdot||_\infty)$는 노름공간을 이룬다.

PROOF 양의 동차성과 양의 정부호성은 최댓값 노름의 정의로부터 명백하므로 삼각부등식만 보이면 된다. 이를 위해 임의의 $v, w \in \mathsf{V}$를 생각하면 각 $i \le n$에 대해 $v_i + w_i \le ||v||_\infty + ||w||_\infty$이므로 $||v+w||_\infty \le ||v||_\infty + ||w||_\infty$가 성립한다. □

Theorem 0.11 n차원 벡터공간 V에 속하는 임의의 $v \in \mathsf{V}$에 대해 $p \to \infty$이면 $||v||_p \to ||v||_\infty$이다.

PROOF 만약 $v = 0$이면 정리가 자명하므로 $v \neq 0$이라 하자. 또한, 간결한 논의를 위해 v의 모든 성분이 0이 아니라 하자. (다른 경우에 대해서도 이와 비슷하게 하면 된다.) 그렇다면 $||v||_\infty > 0$이고 L'Hôspital의 법칙으로부터

$$\begin{aligned}
\lim_{p\to\infty} \frac{||v||_p}{||v||_\infty} &= \lim_{p\to\infty} \left[\sum_{i=1}^{n} \left(\frac{|v_i|}{||v||_\infty} \right)^p \right]^{1/p} \\
&= \lim_{p\to\infty} \exp\left(\frac{1}{p} \log \sum_{i=1}^{n} \left(\frac{|v_i|}{||v||_\infty} \right)^p \right) \\
&= \exp\left(\lim_{p\to\infty} \frac{1}{p} \log \sum_{i=1}^{n} \left(\frac{|v_i|}{||v||_\infty} \right)^p \right) \\
&= \exp\left(\lim_{p\to\infty} \sum_{i=1}^{n} \left(\frac{|v_i|}{||v||_\infty} \right)^p \log\left(\frac{|v_i|}{||v||_\infty} \right) \Big/ \sum_{j=1}^{n} \left(\frac{|v_j|}{||v||_\infty} \right)^p \right) \\
&= \exp\left(-\sum_{i=1}^{n} \left[\lim_{p\to\infty} \left(\frac{|v_i|}{||v||_\infty} \right)^p \log\left(\frac{||v||_\infty}{|v_i|} \right) \Big/ \sum_{j=1}^{n} \left(\frac{|v_j|}{||v||_\infty} \right)^p \right] \right)
\end{aligned}$$

인데, 각 $i \le n$에 대해 만약 $|v_i| = ||v||_\infty$이면 $\log(|v_i|/||v||_\infty) = 0$이고 그렇지 않으면 $|v_i| < |v_k|$인 $k \le n$가 적어도 하나 존재하여

$$\begin{aligned}
0 &\le \lim_{p\to\infty} \left(\frac{|v_i|}{||v||_\infty} \right)^p \log\left(\frac{||v||_\infty}{|v_i|} \right) \Big/ \sum_{j=1}^{n} \left(\frac{|v_j|}{||v||_\infty} \right)^p \\
&\le \lim_{p\to\infty} \left(\frac{|v_i|}{||v||_\infty} \right)^p \log\left(\frac{||v||_\infty}{|v_i|} \right) \Big/ \left(\frac{|v_k|}{||v||_\infty} \right)^p \\
&= \lim_{p\to\infty} \frac{\log(||v||_\infty/|v_i|)}{(|v_k|/|v_i|)^p} \\
&= 0
\end{aligned}$$

이므로 이상으로부터 $p \to \infty$이면 $||v||_p/||v||_\infty \to 1$임을 알고, 증명은 이로써 충분하다. □

무한차원 벡터공간으로 넘어가기 전에, 유한차원 벡터공간에 한해서만 성립하는 중요한 결과 하나를 소개한다. 이는 유한차원에서 사용할 노름의 선택을 단순한 개인의 취향 문제로 만들어버리는 강력한 결과이다.

Definition 0.12 벡터공간 V 위의 두 노름 $||\cdot||, ||\cdot||'$에 대해 $C_1 \le C_2$인 $C_1, C_2 > 0$가 존재하여 임의의 $v \in \mathsf{V}$에 대해 $C_1||v|| \le ||v||' \le C_2||v||$이면 이때 $||\cdot||$와 $||\cdot||'$가 **equivalent**하다고 하고 $||\cdot|| \sim ||\cdot||'$으로 쓴다.

Proposition 0.13 벡터공간 V 위의 노름간의 equivalence는 동치관계이다. 즉, V 위의 노름 $||\cdot||, ||\cdot||', ||\cdot||''$에 대해 다음이 성립한다.

i. (반사성) $||\cdot|| \sim ||\cdot||$.
ii. (대칭성) $||\cdot|| \sim ||\cdot||'$이면 $||\cdot||' \sim ||\cdot||$이다.
iii. (전이성) $||\cdot|| \sim ||\cdot||'$이고 $||\cdot||' \sim ||\cdot||''$이면 $||\cdot|| \sim ||\cdot||''$이다.

Proof 반사성은 명백하므로 대칭성과 전이성만 보이자. 먼저 대칭성을 보이기 위해 $||\cdot|| \sim ||\cdot||'$라 하면 $C_1 \le C_2$인 $C_1, C_2 > 0$가 존재하여 임의의 $v \in \mathsf{V}$에 대해 $C_1||v|| \le ||v||' \le C_2||v||$이므로 $(1/C_2)||v||' \le ||v|| \le (1/C_1)||v||'$에서 $||\cdot||' \sim ||\cdot||$이 되어 대칭성이 성립한다. 다음으로 전이성을 보이기 위해 $||\cdot|| \sim ||\cdot||'$이고 $||\cdot||' \sim ||\cdot||''$라 하면 $C_1 \le C_2$이고 $C_3 \le C_4$인 $C_1, C_2, C_3, C_4 > 0$가 존재하여 임의의 $v \in \mathsf{V}$에 대해 $C_1||v|| \le ||v||' \le C_2||v||$이고 $C_3||v||' \le ||v||'' \le C_4||v||'$이므로 $C_1C_3||v|| \le C_3||v||' \le ||v||'' \le C_4||v||' \le C_2C_4||v||$에서 $||\cdot|| \sim ||\cdot||''$이 되어 전이성이 성립한다. □

Lemma 0.14 유한차원 노름공간 $(\mathsf{V}, ||\cdot||_1)$ 위의 임의의 노름 $||\cdot||$은 균등연속이다.

Proof 벡터공간 V의 기저를 $\beta = \{\beta_1, \cdots, \beta_n\}$라 하고 임의의 $\varepsilon > 0$을 택하면 $||v-w||_1 < \varepsilon / \max_{i=1}^n ||\beta_i||$인 임의의 $v, w \in \mathsf{V}$에 대해

$$\begin{aligned}
|||v|| - ||w||| &\le ||v-w|| \\
&\le ||(v_1, v_2, \cdots, v_n) - (w_1, v_2, \cdots, v_n)|| + ||(w_1, v_2, \cdots, v_n) - (w_1, w_2, \cdots, v_n)|| \\
&\qquad + \cdots + ||(w_1, w_2, \cdots, v_n) - (w_1, w_2, \cdots, w_n)|| \\
&= \sum_{i=1}^n |v_i - w_i| ||\beta_i|| \\
&\le \sum_{i=1}^n |v_i - w_i| \max_{i=1}^n ||\beta_i|| \\
&= (\max_{i=1}^n ||\beta_i||) ||v-w||_1 \\
&< \varepsilon
\end{aligned}$$

이므로 $||\cdot||$은 균등연속이다. □

Theorem 0.15 유한차원 벡터공간 V 위의 임의의 두 노름 $||\cdot||, ||\cdot||'$에 대해 $||\cdot|| \sim ||\cdot||'$이다.

PROOF 벡터공간 V에 1−노름을 장착하면 위의 보조정리로부터 V 위의 임의의 노름 $||\cdot||$이 연속이고, 집합 $S(1) \subseteq \mathsf{V}$이 compact하므로 $||\cdot||$은 $S(1)$ 위에서 최댓값 $M > 0$과 최솟값 $m > 0$을 가진다. 이제 0이 아닌 임의의 $v \in \mathsf{V}$에 대해 $\widetilde{v} = v/||v||_1$이라 하면 $\widetilde{v} \in S(1)$이므로 $m \leq ||\widetilde{v}|| \leq M$에서 $m||v||_1 \leq ||v|| \leq M||v||_1$이고, 이는 $v = 0$인 경우에도 성립하므로 $||\cdot|| \sim ||\cdot||_1$임을 안다. 여기서 $||\cdot||$이 임의의 노름이라는 점과 명제 0.13를 생각해보면 증명은 이로써 충분하다. □

이제 작용소 노름에 대해 알아보자. 작용소 노름은 선형사상에 대해 정의되는 일종의 함수인데, 앞서 소개한 노름들과 달리 ∞를 그 값으로 가질 수 있어서 이를 진짜 노름으로 만들기 위해서는 약간의 준비가 필요하다.

Definition 0.16 노름공간 $(\mathsf{V}, ||\cdot||), (\mathsf{W}, ||\cdot||')$에 대해 **작용소 노름 (operator norm)**을 $||\cdot||_{\text{op}} : \mathsf{L}(\mathsf{V}, \mathsf{W}) \to \overline{\mathbb{R}}_0^+$로 쓰고 $||\cdot||_{\text{op}} : T \mapsto \inf\{M \geq 0 : \text{임의의 } v \in \mathsf{V}\text{에 대해 } ||T(v)||' \leq M||v||\}$로 정의한다.

Proposition 0.17 노름공간 $(\mathsf{V}, ||\cdot||), (\mathsf{W}, ||\cdot||')$ 사이에서 정의된 선형사상 $T, T' : \mathsf{V} \to \mathsf{W}$에 대해 다음이 성립한다.

i. 임의의 $\lambda \in \mathbb{R}$에 대해 $||\lambda T||_{\text{op}} = |\lambda|||T||_{\text{op}}$이다.
ii. $||T + T'||_{\text{op}} \leq ||T||_{\text{op}} + ||T'||_{\text{op}}$.
iii. $||T||_{\text{op}} = 0$일 필요충분조건은 $T = 0$인 것이다.

PROOF iii. 만약 $T = 0$이면 $||T||_{\text{op}} = 0$임이 자명하다. 이제 $||T||_{\text{op}} = 0$인데 $T \neq 0$이라 하면 $T(v) \neq 0$인 0이 아닌 $v \in \mathsf{V}$를 적어도 하나 택할 수 있지만 가정으로부터 적당한 $M < ||T(v)||'/||v||$이 존재하여 $||T(v)||' \leq M||v|| < ||T(v)||'$의 모순이 발생하므로 $T = 0$이다.

i. 만약 $\lambda = 0$이면 iii으로부터 명제가 자명하므로 $\lambda \neq 0$이라 하자. 또한, 만약 $||T||_{\text{op}} = \infty$라면 임의의 $i \in \mathbb{N}$에 대해 적당한 $v_i \in \mathsf{V}$가 존재하여 $||T(v_i)||' > i||v_i||$이다. 이로부터 임의의 $j \in \mathbb{N}$에 대해 $i_j \geq j/|\lambda|$인 부분열 $\{i_j\}$를 잡으면 임의의 $j \in \mathbb{N}$에 대해 $||\lambda T(v_{i_j})||' = |\lambda|||T(v_{i_j})||' > i_j|\lambda|||v_{i_j}|| \geq j||v_{i_j}||$가 되어 $||\lambda T||_{\text{op}} = \infty$에서 이 경우에도 명제가 자명하므로 $||T||_{\text{op}} < \infty$라 하자.

이제 임의의 $v \in \mathsf{V}$에 대해 $||T(v)||' \leq M||v||$인 $M \geq 0$을 임의로 하나 택하면 임의의 $v \in \mathsf{V}$에 대해 $||\lambda T(v)||' = |\lambda|||T(v)||' \leq |\lambda|M||v||$이므로 $||\lambda T||_{\text{op}} \leq |\lambda|M$이고, 곧 $||\lambda T||_{\text{op}} \leq |\lambda|||T||_{\text{op}}$이다. 여기서 $\lambda \neq 0$가 임의였음을 상기한다면 $|\lambda|||T||_{\text{op}} \leq ||\lambda T||_{\text{op}}$도 성립하고, 곧 이상으로부터 $||\lambda T||_{\text{op}} = |\lambda|||T||_{\text{op}}$이다.

ii. 만약 $||T||_{\text{op}} = \infty$이거나 $||T'||_{\text{op}} = \infty$이면 명제가 자명하므로 $||T||_{\text{op}}, ||T'||_{\text{op}} < \infty$라 하자. 이제 임의의 $v \in \mathsf{V}$에 대해 $||T(v)||' \leq M||v||$이고 $||T'(v)||' \leq N||v||$인 $M, N \geq 0$을 임의로

하나씩 택하면 임의의 $v \in \mathsf{V}$에 대해 $||(T+T')(v)||' \leq ||T(v)||' + ||T'(v)||' \leq (M+N)||v||$이므로 $||T+T'||_{\text{op}} \leq M+N$이고, 곧 $||T+T'||_{\text{op}} \leq ||T||_{\text{op}} + ||T'||_{\text{op}}$이다. □

위의 명제를 보니 ∞의 작용소 노름을 가지는 선형사상들만 제거해주면 그 위에서 노름공간을 이룰 수 있을 듯하다.

Definition 0.18 노름공간 $(\mathsf{V}, ||\cdot||)$, $(\mathsf{W}, ||\cdot||')$ 사이에서 정의된 선형사상 $T : \mathsf{V} \to \mathsf{W}$에 대해 $||T||_{\text{op}} < \infty$이면 이때의 선형사상 T를 **유계 (bounded)**라 한다.

Theorem 0.19 노름공간 $(\mathsf{V}, ||\cdot||)$, $(\mathsf{W}, ||\cdot||')$ 사이에서 정의된 선형사상 $T : \mathsf{V} \to \mathsf{W}$에 대해 T가 유계일 필요충분조건은 이가 연속인 것이다.

Proof 먼저 충분조건임을 보이기 위해 T가 유계라 하면 $||T||_{\text{op}} < \infty$이므로 적당한 $M \geq 0$이 존재하여 임의의 $v \in \mathsf{V}$에 대해 $||T(v)||' \leq M||v||$이다. 이제 임의의 $\varepsilon > 0$과 임의의 $v \in \mathsf{V}$를 택하면 $||v-w|| < \varepsilon/(M+1)$인 임의의 $w \in \mathsf{V}$에 대해 $||T(v)-T(w)||' = ||T(v-w)||' \leq M||v-w|| < \varepsilon$이므로 T가 연속임을 안다. 다음으로 필요조건임을 보이기 위해 T가 연속이라 하면 임의의 $v \in \mathsf{V}$에 대해 적당한 $\delta > 0$가 존재하여 $||v|| < \delta$이면 $||T(v)||' = ||T(v)-T(0)||' < 1$이다. 따라서 0이 아닌 임의의 $v \in \mathsf{V}$에 대해 $\widetilde{v} = \delta v/2||v||$라 하면 $||\widetilde{v}|| = \delta/2 < \delta$이므로 $(\delta/2||v||)||T(v)||' = ||T(\widetilde{v})||' < 1$에서 $||T(v)||' < (2/\delta)||v||$이고, 이는 $v = 0$인 경우에도 성립하므로 곧 $||T||_{\text{op}} \leq 2/\delta$에서 T가 유계임을 안다. □

Proposition 0.20 노름공간 $(\mathsf{V}, ||\cdot||)$, $(\mathsf{W}, ||\cdot||')$ 사이에서 정의된 모든 유계인 선형사상의 집합을 $\mathsf{BL}(\mathsf{V}, \mathsf{W}) \subseteq \mathsf{L}(\mathsf{V}, \mathsf{W})$라 하면 $(\mathsf{BL}(\mathsf{V}, \mathsf{W}), ||\cdot||_{\text{op}})$는 노름공간을 이룬다.

Proof 집합 $\mathsf{BL}(\mathsf{V}, \mathsf{W})$이 벡터공간임을 보일 수만 있으면 명제 0. 17로부터 명제가 자명한데, 정리 0. 19로부터 $\mathsf{BL}(\mathsf{V}, \mathsf{W})$의 모든 원소는 연속함수이고, 연속함수의 합과 스칼라곱은 연속함수라는 점에서 다시 정리 0. 19로부터 그 결과가 $\mathsf{BL}(\mathsf{V}, \mathsf{W})$에 속하여 $\mathsf{BL}(\mathsf{V}, \mathsf{W})$은 덧셈과 스칼라곱에 대해 닫혀있음을 안다. 이제 $\mathsf{BL}(\mathsf{V}, \mathsf{W})$이 벡터공간이 되기 위해 만족해야할 나머지 조건들은 쉽게 확인해볼 수 있다. □

만약 V가 유한차원이라면 더 간단해진다.

Theorem 0.21 노름공간 $(\mathsf{V}, ||\cdot||)$, $(\mathsf{W}, ||\cdot||')$ 사이에서 정의된 선형사상 $T : \mathsf{V} \to \mathsf{W}$에 대해 $\dim \mathsf{V} < \infty$이면 T는 유계이고, 따라서 연속이다.

Proof 벡터공간 V의 기저를 $\beta = \{\beta_1, \cdots, \beta_n\}$이라 하고 $M = \max_{i=1}^n ||T(\beta_i)||'$라 하자. 그렇다면 임의의 $v \in \mathsf{V}$에 대해

$$||T(v)||' = \left|\left| T\left(\sum_{i=1}^n v_i \beta_i\right) \right|\right|'$$

$$\begin{aligned}
&= \left|\left| \sum_{i=1}^{n} v_i T(\beta_i) \right|\right|' \\
&\le \sum_{i=1}^{n} |v_i| ||T(\beta_i)||' \\
&\le M \sum_{i=1}^{n} |v_i| \\
&= M||v||_\infty
\end{aligned}$$

인데, 정리 0.15로부터 v와 무관한 $C > 0$가 존재하여 $||v||_\infty \le C||v||$이므로 $||T(v)||' \le CM||v||$이다. 이는 곧 $||T||_{\mathrm{op}} \le CM < \infty$임을 뜻하므로 T는 유계이고, 정리 0.19로부터 연속이다. □

Corollary 0.22 노름공간 $(\mathsf{V}, ||\cdot||), (\mathsf{W}, ||\cdot||')$에 대해 $\dim \mathsf{V} < \infty$이면 $(\mathsf{L}(\mathsf{V}, \mathsf{W}), ||\cdot||_{\mathrm{op}})$는 노름공간을 이룬다.

Proof 이는 위의 정리와 명제 0.20로부터 자명하다. □

작용소 노름의 중요한 성질 하나를 소개하고, 다음 주제로 넘어가자.

Theorem 0.23 노름공간 $(\mathsf{U}, ||\cdot||), (\mathsf{V}, ||\cdot||'), (\mathsf{W}, ||\cdot||'')$에 대해 선형사상 $T : \mathsf{U} \to \mathsf{V}, T' : \mathsf{V} \to \mathsf{W}$에 대해 다음이 성립한다.

i. 임의의 $v \in \mathsf{U}$에 대해 $||T(v)||' \le ||T||_{\mathrm{op}}||v||$이다.
ii. $||T' \circ T||_{\mathrm{op}} \le ||T||_{\mathrm{op}}||T'||_{\mathrm{op}}$.

Proof i. 우선 $v = 0$에 대해서는 정리가 자명하므로 $v \ne 0$이라 하자. 또한, $||T||_{\mathrm{op}} = \infty$인 경우에도 정리가 자명하므로 $||T||_{\mathrm{op}} < \infty$라 하자. 이제 임의의 $v \in \mathsf{U}$에 대해 $||T(v)||' \le M||v||$인 $M \ge 0$을 임의로 하나 택하면 0이 아닌 임의의 $v \in \mathsf{U}$에 대해 $||T(v)||'/||v|| \le M$이므로 곧 $||T(v)||'/||v|| \le ||T||_{\mathrm{op}}$에서 증명이 끝난다.

ii. 만약 $||T||_{\mathrm{op}} = 0$이거나 $||T'||_{\mathrm{op}} = 0$이면 $T' \circ T = 0$이 되어 정리가 자명하므로 $||T||_{\mathrm{op}}, ||T'||_{\mathrm{op}} > 0$이라 하자. 비슷하게, $||T||_{\mathrm{op}} = \infty$이거나 $||T'||_{\mathrm{op}} = \infty$인 경우에도 정리가 자명하므로 $||T||_{\mathrm{op}}, ||T'||_{\mathrm{op}} < \infty$라 하자. 이제 임의의 $v \in \mathsf{U}$와 임의의 $w \in \mathsf{V}$에 대해 $||T(v)||' \le M||v||$이고 $||T'(w)||'' \le N||w||$인 $M, N > 0$을 임의로 하나씩 택하면 임의의 $v \in \mathsf{U}$에 대해 $||(T' \circ T)(v)||'' \le N||T(v)||' \le MN||v||$이므로 $||T' \circ T||_{\mathrm{op}} \le MN$에서 $||T' \circ T||_{\mathrm{op}}/N \le M$이고, 곧 $||T' \circ T||_{\mathrm{op}}/N \le ||T||_{\mathrm{op}}$이다. 비슷하게, 이는 다시 $||T' \circ T||_{\mathrm{op}}/||T||_{\mathrm{op}} \le N$임을 의미하여 $||T' \circ T||_{\mathrm{op}}/||T||_{\mathrm{op}} \le ||T'||_{\mathrm{op}}$가 성립하고, 곧 $||T' \circ T||_{\mathrm{op}} \le ||T||_{\mathrm{op}}||T'||_{\mathrm{op}}$가 되어 증명이 끝난다. □

마지막으로 살펴볼 노름은 행렬 노름이다. 행렬에 노름을 주는 방식은 크게 두 가지가 있는데, 우선 $\mathsf{M}_{m,n}$을 $\mathsf{L}(\mathbb{R}^n, \mathbb{R}^m)$과 동일시하여 행렬을 선형사상으로 파악하고, 이에 방금 소개한 작용소 노름을 주는 방식을 소개한다.

Definition 0.24 행렬 $A \in \mathsf{M}_{m,n}$와 $p \geq 1$에 대해 이를 p−노름(혹은 최댓값 노름)이 장착된 $\mathbb{R}^n$과 $\mathbb{R}^m$ 사이에서 정의된 선형사상 $T_A : \mathbb{R}^n \to \mathbb{R}^m$로 생각하여 **$p$−노름(혹은 최댓값 노름)으로부터 유도된 작용소 노름(operator norm induced by p-norm(resp. maximum norm))**을 $||\cdot||_p : \mathsf{M}_{m,n} \to \mathbb{R}_0^+$로 쓰고 $||\cdot||_p : A \mapsto ||T_A||_{\text{op}}$로 정의한다.

Proposition 0.25 벡터공간 $\mathsf{M}_{m,n}$와 $p \geq 1$에 대해 p−노름(혹은 최댓값 노름)으로부터 유도된 작용소 노름은 노름이다. 따라서 $(\mathsf{M}_{m,n}, ||\cdot||_p)$(혹은 $(\mathsf{M}_{m,n}, ||\cdot||_p)$)는 노름공간을 이룬다.

Proof 이는 따름정리 0.22로부터 자명하다. □

p−노름(혹은 최댓값 노름)으로부터 유도된 작용소 노름은 본질적으로 작용소 노름이므로 정리 0.23가 그대로 성립한다.

Theorem 0.26 임의의 $A \in \mathsf{M}_{l,m}, B \in \mathsf{M}_{m,n}$와 $p \geq 1$에 대해 다음이 성립한다.

i. 임의의 $x \in \mathbb{R}^m$에 대해 $||Ax||_p \leq ||A||_p||x||_p$ (혹은 $||Ax||_\infty \leq ||A||_\infty||x||_\infty$)이다.
ii. $||AB||_p \leq ||A||_p||B||_p$. (혹은 $||AB||_\infty \leq ||A||_\infty||B||_\infty$.)

Proof 이는 p−노름(혹은 최댓값 노름)으로부터 유도된 작용소 노름의 정의와 정리 0.23로부터 자명하다. □

특별히, 1−노름으로부터 유도되거나 최댓값 노름으로부터 유도된 경우에는 그로부터 유도된 작용소 노름을 간단하게 구할 수 있다. 작용소 노름의 정의가 행렬의 성분과는 적어도 직접적으로는 관계가 없다는 점을 생각해보면 다음 정리는 조금 신기한 결과이다.

Theorem 0.27 임의의 $A \in \mathsf{M}_{m,n}$에 대해 다음이 성립한다.

i. $||A||_1 = \max_{j=1}^n \sum_{i=1}^m |a_{ij}|$.
ii. $||A||_\infty = \max_{i=1}^m \sum_{j=1}^n |a_{ij}|$.

Proof i. 임의의 $x \in \mathbb{R}^n$에 대해

$$\begin{aligned} ||T_A(x)||_1 &= ||Ax||_1 \\ &= \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij}x_j \right| \end{aligned}$$

$$
\begin{aligned}
&\le \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}x_j| \\
&= \sum_{j=1}^{n}\left(\sum_{i=1}^{m}|a_{ij}|\right)|x_j| \\
&\le \left(\max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|\right)\sum_{j=1}^{n}|x_j| \\
&= \left(\max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|\right)||x||_1
\end{aligned}
$$

이므로 $||A||_1 = ||T_A||_{\text{op}} \le \max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|$이다. 한편, 만약 임의의 $x \in \mathbb{R}^n$에 대해 $||T_A(x)||_1 \le M||x||_1$인 $M < \max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|$이 존재한다면 $j_0 = \operatorname{argmax}_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|$에 대해

$$
\begin{aligned}
\max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}| &= \sum_{i=1}^{n}|a_{ij_0}| \\
&= ||A\mathbf{e}_{j_0}||_1 \\
&= ||T_A(\mathbf{e}_{j_0})||_1 \\
&\le M||\mathbf{e}_{j_0}|| \\
&= M \\
&< \max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|
\end{aligned}
$$

의 모순이 발생하므로 이러한 M은 존재하지 않고, 곧 $||A||_1 = ||T_A||_{\text{op}} = \max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|$이다.

ii. 만약 $A = O$이면 $||A||_\infty = 0$이 되어 정리가 자명하므로 $A \ne O$이라 하자. 그렇다면 임의의 $x \in \mathbb{R}^n$에 대해

$$
\begin{aligned}
||T_A(x)||_\infty &= ||Ax||_\infty \\
&= \max_{i=1}^{m}\left|\sum_{j=1}^{n}a_{ij}x_j\right| \\
&\le \max_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}x_j| \\
&\le \left(\max_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|\right)\max_{j=1}^{n}|x_j| \\
&= \left(\max_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|\right)||x||_\infty
\end{aligned}
$$

이므로 $||A||_\infty = ||T_A||_{\text{op}} \le \max_{i=1}^m \sum_{j=1}^n |a_{ij}|$ 이다. 한편, 만약 임의의 $x \in \mathbb{R}^n$ 에 대해 $||T_A(x)||_\infty \le M||x||_\infty$ 인 $M < \max_{i=1}^m \sum_{j=1}^n |a_{ij}|$ 이 존재한다면 $i_0 = \operatorname{argmax}_{i=1}^m \sum_{j=1}^n |a_{ij}|$ 와 $x_0 = (\operatorname{sgn}(a_{i_0 j}))$ 에 대해 x_0 의 성분 중 적어도 하나는 0이 아니므로 $||x_0||_\infty = 1$ 이 되어

$$\begin{aligned}
\max_{i=1}^m \sum_{j=1}^n |a_{ij}| &= \max_{i=1}^m \left| \sum_{j=1}^n \operatorname{sgn}(a_{i_0 j}) a_{ij} \right| \\
&= ||Ax_0||_\infty \\
&= ||T_A(x_0)||_\infty \\
&\le M||x_0||_\infty \\
&= M \\
&< \max_{i=1}^m \sum_{j=1}^n |a_{ij}|
\end{aligned}$$

의 모순이 발생하므로 이러한 M 은 존재하지 않고, 곧 $||A||_\infty = ||T_A||_{\text{op}} = \max_{i=1}^m \sum_{j=1}^n |a_{ij}|$ 이다. 위의 식에서 첫번째 등호는 각 $i \le m$ 에 대해 $|\sum_{j=1}^n \operatorname{sgn}(a_{i_0 j}) a_{ij}| \le \sum_{j=1}^n |a_{ij}|$ 이고 $\max_{i=1}^m \sum_{j=1}^n |a_{ij}| = \sum_{j=1}^n |a_{i_0 j}| = \sum_{j=1}^n \operatorname{sgn}(a_{i_0 j}) a_{i_0 j}$ 이므로 성립한다. □

행렬에 노름을 주는 또다른 방식은 $\mathsf{M}_{m,n}$ 을 $\mathbb{R}^{mn}$ 과 동일시하여 행렬을 벡터로 파악하고, 이에 p−노름이나 최댓값 노름을 주는 방식이다. 사실 $\mathbb{R}^{mn}$ 에서 정의되는 어떤 노름이든 줄 수 있지만, 2−노름을 주는 경우가 일반적이며, 여기서도 이 경우에 대해서만 살펴본다.

Definition 0.28 행렬 $A \in \mathsf{M}_{m,n}$ 에 대해 이를 2−노름이 장착된 $(\mathbb{R}^{m+n}, ||\cdot||_2)$ 에 속하는 벡터 $\operatorname{vec} A \in \mathbb{R}^{m+n}$ 로 생각하여 **Frobenius 노름 (− norm)**을 $||\cdot||_{\mathrm{F}} : \mathsf{M}_{m,n} \to \mathbb{R}_0^+$ 로 쓰고 $||\cdot||_{\mathrm{F}} : A \mapsto ||\operatorname{vec} A||_2$ 로 정의한다.

신기하게도, Frobenius 노름은 선형사상으로서의 행렬과는 그다지 관련이 없음에도 불구하고 정리 0.23과 유사한 성질들을 가진다.

Theorem 0.29 임의의 $A \in \mathsf{M}_{l,m}$ 와 $B \in \mathsf{M}_{m,n}$ 에 대해 다음이 성립한다.

i. 임의의 $x \in \mathbb{R}^m$ 에 대해 $||Ax|| \le ||A||_{\mathrm{F}} ||x||$ 이다.
ii. $||AB||_{\mathrm{F}} \le ||A||_{\mathrm{F}} ||B||_{\mathrm{F}}$.

Proof i. Cauchy-Schwarz의 부등식으로부터

$$\begin{aligned}
||Ax|| &= \sqrt{\sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right)^2} \\
&\le \sqrt{\sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}^2 \right) \left(\sum_{j=1}^n x_j^2 \right)}
\end{aligned}$$

$$= ||A||_{\mathrm{F}}||x||$$

이므로 이는 자명하다.

ii. 행렬 B의 각 열을 $b_1, \cdots, b_n \in \mathbb{R}^m$이라 하면 $AB = [Ab_1, \cdots, Ab_n]$이고 곧 i로부터

$$\begin{aligned}||AB||_{\mathrm{F}} &= \sqrt{\sum_{i=1}^{n} ||Ab_i||^2} \\ &\le \sqrt{\sum_{i=1}^{n} ||A||_{\mathrm{F}}^2 ||b_i||^2} \\ &= ||A||_{\mathrm{F}} \sqrt{\sum_{i=1}^{n} ||b_i||^2} \\ &= ||A||_{\mathrm{F}}||B||_{\mathrm{F}}\end{aligned}$$

이다. □

0.1.3 Limits and Continuities

Definition 0.30 집합 $A \subseteq \mathbb{R}^m$에서 정의된 함수 $f : A \to \mathbb{R}^n$와 한 점 $x_0 \in \mathrm{acc}A$에 대해 $x_0 \in \mathrm{acc}\{x \in A : x \ge x_0\}$이고 극한 $\lim_{x\downarrow x_0} f(x)$가 존재하는 경우 그 값을 $f(x_0+)$로 쓴다. 여기서 $\lim_{x\downarrow x_0} f(x) = f(x_0+)$는 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 $0 < ||x - x_0|| < \delta$이고 $x \ge x_0$인 임의의 $x \in A$에 대해 $||f(x) - f(x_0+)|| < \varepsilon$이라는 의미이다. 이제 $f(x_0-)$도 비슷하게 정의한다.

Proposition 0.31 집합 $A \subseteq \mathbb{R}^m$에서 정의된 함수 $f : A \to \mathbb{R}^n$와 한 점 $x_0 \in \mathrm{acc}A$에 대해 $x_0 \in \mathrm{acc}\{x \in A : x \ge x_0\} \cap \mathrm{acc}\{x \in A : x \le x_0\}$이고 $x \to x_0$일 때 $f(x) \to L$이면 $f(x_0+) = f(x_0-) = L$이다. 특별히, $m = 1$인 경우 $f(x_0+) = f(x_0-) = L$이면 $x \to x_0$일 때 $f(x) \to L$이다.

PROOF 이는 거의 자명하다. □

Theorem 0.32 집합 $A \subseteq \mathbb{R}^n$에서 정의된 함수 $f : A \to \mathbb{R}$에 대해 이가 각 변수에 대해 증가(혹은 감소)한다고 하자. 그렇다면 임의의 $x_0 \in \mathrm{acc}A$에 대해 $x_0 \in \mathrm{acc}\{x \in A : x \ge x_0\}$이면 $f(x_0+)$가 존재한다. 비슷하게, $x_0 \in \mathrm{acc}\{x \in A : x \le x_0\}$이면 $f(x_0-)$가 존재한다.

PROOF 먼저 f가 각 변수에 대해 증가하는 경우를 생각하면 가정으로부터 각 $i \in \mathbb{N}$에 대해 $x_i > x_{i+1}$인 x_0로 수렴하는 수열 $\{x_i\}$가 존재한다. 그렇다면 수열 $\{f(x_i)\}$는 감소하는 수열

이 되어 MSP로부터 적당한 $L \in \mathbb{R}$로 수렴하며, 곧 임의의 $\varepsilon > 0$을 택하면 적당한 $i_0 \in \mathbb{N}$가 존재하여 $L \le f(x_{i_0}) < L + \varepsilon$이다. 따라서 $0 < ||x - x_0|| < ||x_{i_0}||_\infty$이고 $x \ge x_0$인 임의의 $x \in A$에 신해 $x < x_{i_0}$에서 $f(x) \le f(x_{i_0}) < L + \varepsilon$이고, 만약 $f(x) < L - \varepsilon$이면 충분히 큰 $i \in \mathbb{N}$에 대해 $x_i < x$에서 $L \le f(x_i) \le f(x) < L - \varepsilon$의 모순이 발생하므로 $f(x) > L - \varepsilon$이 되어 $|f(x) - L| < \varepsilon$임을 안다. 이상으로부터 $f(x_0+)$이 L로 존재함을 안다. 이제 $f(x_0-)$에 대해서도 비슷하게 하면 되고, f가 각 변수에 대해 감소하는 경우에는 $-f$를 대신 생각하면 된다. □

Definition 0.33 집합 $A \subseteq \mathbb{R}^m$에서 정의된 함수 $f : A \to \mathbb{R}^n$와 한 점 $x_0 \in A$를 생각하자. 만약 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 $||x_0 - x|| < \delta$이고 $x \ge x_0$인 모든 $x \in A$에 대해 $||f(x_0) - f(x)|| < \varepsilon$이 성립하면 f가 x_0 **에서 오른쪽 연속 (right-continuous at x_0)** 혹은 **continuous from above at** x_0라 한다. 나아가, 만약 f가 모든 $x \in A$에서 오른쪽 연속이면 이때의 f를 **오른쪽 연속 (right-continuous)** 혹은 **continuous from above**라 한다. 이제 x **에서 왼쪽 연속 (left-continuous at** x_0**)**과 **왼쪽 연속 (left-continuous)**도 비슷하게 정의한다.

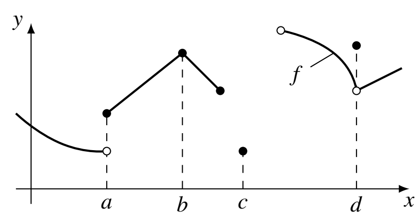


Figure 0.1 함수 f는 a에서는 오른쪽 연속이지만 연속이 아니고, b에서는 연속이면서 오른쪽 연속이며, c에서는 이가 고립점이므로 연속이면서 오른쪽 연속이고, d에서는 오른쪽 연속도, 연속도 아니다.

Proposition 0.34 집합 $A \subseteq \mathbb{R}^m$에서 정의된 함수 $f : A \to \mathbb{R}^n$와 한 점 $x_0 \in A$에 대해 f가 x_0에서 연속이면 이는 x_0에서 오른쪽 연속이고 왼쪽 연속이다. 특별히, $m = 1$인 경우 f가 x_0에서 오른쪽 연속인 동시에 왼쪽 연속이면 이는 x_0에서 연속이다.

Proof 이는 거의 자명하다. □

Proposition 0.35 집합 $A \subseteq \mathbb{R}^m$에서 정의된 함수 $f : A \to \mathbb{R}^n$와 한 점 $x_0 \in A$에 대해 $f(x_0+)$가 존재한다면 f가 x_0에서 오른쪽 연속일 필요충분조건은 $f(x_0+) = f(x_0)$인 것이다. 비슷하게, 만약 $f(x_0-)$가 존재한다면 f가 x_0에서 왼쪽 연속일 필요충분조건은 $f(x_0-) = f(x_0)$인 것이다.

Proof 이도 거의 자명하다. □

0.1.4 Convex Sets and Convex Functions

Definition 0.36 벡터공간 V의 부분집합 $C \subseteq \mathrm{V}$를 생각하자. 만약 임의의 $v, w \in C$와 임의의 $t \in [0,1]$에 대해 $tv + (1-t)w \in C$이면 이때의 집합 C를 **볼록집합 (convex set)**이라 한다.

Definition 0.37 벡터공간 V의 볼록 부분집합 $C \subseteq \mathrm{V}$에서 정의된 함수 $f : C \to \mathbb{R}$를 생각하자. 만약 임의의 $x, y \in C$와 임의의 $t \in [0,1]$에 대해 $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$가 성립하면 이때의 함수 f를 **볼록함수 (convex function)**라 한다. 나아가, 만약 서로다른 $x, y \in C$와 임의의 $t \in (0,1)$에 대해 $f(tx + (1-t)y) < tf(x) + (1-t)f(y)$가 성립하면 이때의 함수 f를 **순볼록함수 (strictly convex function)**라 한다. 한편, 만약 $-f$가 볼록하거나 순볼록하면 이때의 f를 각각 **오목함수 (concave function)**, **순오목함수 (strictly concave function)**라 한다.

Theorem 0.38 비어있지 않은 열린구간 $I \subseteq \mathbb{R}$에서 정의된 볼록함수 $f : I \to \mathbb{R}$와 임의의 $x_0 \in I$에 대해 x_0에서의 supporting line이 존재한다. 즉, 적당한 직선 $l : mx + n = g(x)$이 존재하여 $(x_0, f(x_0))$를 지나고 I에서 $g \leq f$이다.

Proof 먼저 x_0에서 f의 우미분계수가 존재함을 보이자. 이를 위해 함수 $h : I \setminus \{x_0\} \to \mathbb{R}$를 $h : x \mapsto [f(x_0) - f(x)]/(x_0 - x)$로 두고 $x \leq y$인 임의의 $x, y \in I \setminus \{x_0\}$를 택하자. 만약 $x_0 < x \leq y$이면 적당한 $t \in [0,1]$에 대해 $x = tx_0 + (1-t)y$이므로

$$\begin{aligned}
h(y) - h(x) &= \frac{f(x_0) - f(y)}{x_0 - y} - \frac{f(x_0) - f(x)}{x_0 - x} \\
&= \frac{f(x_0) - f(y)}{x_0 - y} - \frac{f(x_0) - f(tx_0 + (1-t)y)}{(1-t)(x_0 - y)} \\
&\geq \frac{f(x_0) - f(y)}{x_0 - y} - \frac{f(x_0) - tf(x_0) - (1-t)f(y)}{(1-t)(x_0 - y)} \\
&= 0
\end{aligned}$$

에서 $h(y) \geq h(x)$이다. 이제 $x < x_0 < y$이거나 $x \leq y < x_0$인 경우에도 비슷하게 $h(y) \geq h(x)$임을 보일 수 있으므로 h는 증가함수가 되어 $h(x_0+)$가 존재하고, 곧 x_0에서의 우미분계수가 존재한다.

표기의 편의를 위해 x_0에서의 우미분계수를 $f^{(+)}(x_0)$라 하고 함수 $g^+ : \mathbb{R} \to \mathbb{R}$를 $g : x \mapsto f^{(+)}(x_0)(x - x_0) + f(x_0)$로 두자. 그렇다면 임의의 $x \in I \setminus \{x_0\}$에 대해

$$\begin{aligned}
f(x) - g^+(x) &= f(x) - f^{(+)}(x_0)(x - x_0) - f(x_0) \\
&= (x - x_0)\left[\frac{f(x_0) - f(x)}{x_0 - x} - f^{(+)}(x_0)\right]
\end{aligned}$$

$$
\begin{aligned}
&= (x-x_0)[h(x)-h(x_0+)] \\
&\geq 0
\end{aligned}
$$

에서 $f(x) \geq g^+(x)$이고, 이는 $x = x_0$인 경우에도 성립하므로 I에서 $f \geq g^+$ 임을 안다.

위와 비슷하게 하면 f의 x_0에서의 좌미분계수 $f^{(-)}(x_0)$도 존재함을 알고, 함수 g^- : $\mathbb{R} \to \mathbb{R}$를 $g^- : x \mapsto f^{(-)}(x_0)(x-x_0)+f(x_0)$로 두면 I에서 $f \geq g^-$ 임도 안다. 이제 $m \in [f^{(-)}(x_0), f^{(+)}(x_0)]$을 적당히 택하여 직선 $l : m(x-x_0)+f(x_0) = g(x)$을 생각하면 이는 $(x_0, f(x_0))$을 지나며 임의의 $x \in I \setminus \{x_0\}$에 대해 만약 $x > x_0$이면

$$
\begin{aligned}
f(x)-g(x) &= f(x)-m(x-x_0)-f(x_0) \\
&= (x-x_0)\left[\frac{f(x_0)-f(x)}{x_0-x}-m\right] \\
&\geq (x-x_0)[h(x)-h(x_0+)] \\
&\geq 0
\end{aligned}
$$

이고, 비슷하게 $x < x_0$이라도 $f(x) \geq g(x)$이 되어 l이 우리가 찾는 supporting line이 된다. □

Definition 0.39 벡터공간 $\mathbb{V}$의 볼록 부분집합 $C \subseteq \mathbb{V}$에서 정의된 양함수 $f : C \to \mathbb{R}^+$를 생각하자. 만약 함수 $\log f : C \to \mathbb{R}$가 볼록하면 이때의 함수 f를 **log−볼록함수 (logarithmically convex function)**라 한다. 이제 **log−순볼록함수 (logarithmically strictly convex function)**, **log−오목함수 (logarithmically concave function)**, **log−순오목함수 (logarithmically strictly convex function)**도 비슷하게 정의한다.

Theorem 0.40 벡터공간 $\mathbb{V}$의 볼록 부분집합 $C \subseteq \mathbb{V}$에서 정의된 log−볼록한 (혹은 log−순볼록한, log−오목한, log−순오목한) 양함수 $f : C \to \mathbb{R}^+$는 볼록 (혹은 순볼록, 오목, 순오목) 하다.

PROOF 임의의 $v, w \in C$와 임의의 $t \in [0,1]$에 대해 지수함수가 순볼록하다는 사실로부터 $f(tv+(1-t)w) = \exp(\log f(tv+(1-t)w)) \leq \exp(t\log f(v)+(1-t)\log f(w)) \leq t\exp(\log f(v))+(1-t)\exp(\log f(w)) = tf(v)+(1-t)f(w)$가 되어 f가 볼록함을 안다. 한편, f가 순볼록, 오목, 순오목한 경우에도 이와 비슷하게 하면 된다. □

0.1.5 Power Series and Analytic Functions

Definition 0.41 수열 $\{a_i\}$와 $z \in \mathbb{C}$에 대해 $\sum_{i=0}^{\infty} a_i z^i$로 주어진 급수를 **멱급수 (power series)**라 한다. 나아가, 어떤 멱급수에 대해 $|z| < R$이면 이가 절대수렴하고 $|z| > R$이면 이가

수렴하지 않도록 하는 $R \geq 0$을 그 멱급수의 **수렴반경 (radius of convergence)**이라 하며, 집합 $B(R)$을 **수렴영역 (domain of convergence)**이라 한다.

Theorem 0.42 멱급수 $\sum_{i=0}^{\infty} a_i z^i, \sum_{i=0}^{\infty} b_i z^i$의 수렴반경을 각각 $R, S > 0$라 하여 함수 $f : B(R) \to \mathbb{C}, g : B(S) \to \mathbb{C}$를 각각 $f : z \mapsto \sum_{i=0}^{\infty} a_i z^i, g : z \mapsto \sum_{i=0}^{\infty} b_i z^i$로 두면 다음이 성립한다.

i. 임의의 $|z| < \min\{R, S\}$에 대해 $(f+g)(z) = \sum_{i=0}^{\infty}(a_i + b_i)z^i$이다.
ii. 임의의 $|z| < \min\{R, S\}$에 대해 $(fg)(z) = \sum_{i=0}^{\infty} \sum_{j=0}^{i} a_j b_{i-j} z^i$이다.
iii. 만약 $g(0) \neq 0$이면 적당한 $T > 0$와 적당한 수열 $\{c_i\}$가 존재하여 임의의 $|z| < T$에 대해 $(f/g)(z) = \sum_{i=0}^{\infty} c_i z^i$이다.

Proof i. 이는 자명하다.

ii. 임의의 $|z| < \min\{R, S\}$를 택하면 $\sum_{i=0}^{\infty} a_i z^i$와 $\sum_{i=0}^{\infty} b_i z^i$가 절대수렴하므로 적당한 $M \geq 0$이 존재하여 $\sum_{i=0}^{\infty} |a_i z^i|, \sum_{i=0}^{\infty} |b_i z^i| \leq M$이다. 그렇다면 임의의 $k \in \mathbb{N}$에 대해 $\sum_{i=0}^{k} |\sum_{j=0}^{i} a_j b_{i-j} z^i| \leq \sum_{i=0}^{k} \sum_{j=0}^{i} |a_j b_{i-j} z^i| \leq (\sum_{i=0}^{\infty} |a_i z^i|)(\sum_{i=0}^{\infty} |b_i z^i|) \leq M^2$이므로 MSP에서 $\sum_{i=0}^{\infty} \sum_{j=0}^{i} a_j b_{i-j} z^i$는 절대수렴한다. 이제 절대수렴하는 급수는 그 합의 순서를 바꿀 수 있으므로

$$\begin{aligned}
\sum_{i=0}^{\infty} \sum_{j=0}^{i} a_j b_{i-j} z^i &= \sum_{i=0}^{\infty} \sum_{j=0}^{i} a_j x^j b_{i-j} z^{i-j} \\
&= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_i z^i b_j z^j \\
&= \left(\sum_{i=0}^{\infty} a_i z^i \right) \left(\sum_{i=0}^{\infty} b_i z^i \right) \\
&= (fg)(z)
\end{aligned}$$

이다.

iii. 이는 적당한 $T > 0$와 적당한 수열 $\{c_i\}$가 존재하여 임의의 $|z| < T$에 대해 $(1/g)(z) = \sum_{i=0}^{\infty} c_i z^i$임을 보이는 것으로 충분하다. 먼저 g가 연속이고 $g(0) \neq 0$이므로 $b_0 \neq 0$이고 적당한 0의 근방 $U \subseteq B(S)$에서 g가 0이 되지 않아 $1/g$가 well-define된다. 이제 수열 $\{c_i\}$를 점화식

$$c_i = \begin{cases} \dfrac{1}{b_0} & i = 0\text{인 경우} \\ -\dfrac{1}{b_0} \displaystyle\sum_{j=0}^{i-1} b_{i-j} c_j & i > 0\text{인 경우} \end{cases}$$

를 통해 정의하고 임의의 $|z| < S$를 택하면 $\sum_{i=0}^{\infty} c_i z^i$가 수렴한다는 가정 하에 $g(z)(\sum_{i=0}^{\infty} c_i z^i) = (\sum_{i=0}^{\infty} b_i z^i)(\sum_{i=0}^{\infty} c_i z^i) = \sum_{i=0}^{\infty}\sum_{j=0}^{i} b_j c_{j-i} z^i = 1$이 되어 $(1/g)(z) = \sum_{i=0}^{\infty} c_i z^i$임을 안다. 따라서 적당한 $T > 0$가 존재하여 임의의 $|z| < T$에 대해 $\sum_{i=0}^{\infty} c_i z^i$가 수렴한다는 것만 보이면 증명은 끝난다. 이를 위해 함수 $h : B(S) \to \mathbb{R}$를 $h : z \mapsto \sum_{i=0}^{\infty} |b_i z^i|$로 두면 이는 연속이고 $h(0) = |b_0|$이므로 적당한 $\delta > 0$가 존재하여 $|z| \le |b_0|\delta$이면 $\sum_{i=1}^{\infty} |b_i z^i| = |h(z) - |b_0|| < |b_0|$이다. 이제 수학적 귀납법을 사용하여 모든 $i \in \mathbb{N}$에 대해 $|c_i| \le (|b_0|\delta)^{-i}$임을 보이자. 먼저 $i = 0$인 경우에는 자명하므로 귀납가정으로서 모든 $j \le i \in \mathbb{N}$에 대해 $|c_j| \le (|b_0|\delta)^{-j}$라 하면

$$
\begin{aligned}
|c_{i+1}| &= \left| \frac{1}{b_0} \sum_{j=0}^{i} b_{i+1-j} c_j \right| \\
&\le \frac{1}{|b_0|} \sum_{j=0}^{i} |b_{i+1-j}||c_j| \\
&\le \frac{1}{|b_0|} \sum_{j=0}^{i} \frac{|b_{i+1-j}|}{(|b_0|\delta)^j} \\
&= \frac{1}{|b_0|(|b_0|\delta)^{i+1}} \sum_{j=0}^{i} |b_{i+1-j}|(|b_0|\delta)^{i+1-j} \\
&= \frac{1}{|b_0|(|b_0|\delta)^{i+1}} \sum_{j=1}^{i+1} |b_j|(|b_0|\delta)^j \\
&\le \frac{1}{|b_0|(|b_0|\delta)^{i+1}} \sum_{j=1}^{\infty} |b_j|(|b_0|\delta)^j \\
&< \frac{1}{(|b_0|\delta)^{i+1}}
\end{aligned}
$$

에서 $|c_{i+1}| \le (|b_0|\delta)^{-(i+1)}$이고, 곧 모든 $i \in \mathbb{N}$에 대해 $|c_i| \le (|b_0|\delta)^{-i}$임을 안다. 이로부터 $\limsup_{i\to\infty} |c_i|^{1/i} \le 1/|b_0|\delta$이므로 $\sum_{i=0}^{\infty} c_i z^i$의 수렴반경은 $|b_0|\delta$ 이상이고, $T = \min\{S, |b_0|\delta\}$로 두면 증명이 끝난다. □

Definition 0.43 열린집합 $U \subseteq \mathbb{R}$에서 정의된 $\mathscr{C}^\infty$급 함수 $f : U \to \mathbb{R}$와 한 점 $x_0 \in U$를 생각하자. 만약 적당한 x_0의 근방 $V \subseteq U$가 존재하여 임의의 $x \in V$에 대해 급수 $\sum_{i=0}^{\infty} f^{(i)}(x_0)(x - x_0)^i/i!$가 $f(x)$로 수렴하면 함수 f가 x_0에서 **(실) 해석적 ((real) analytic)**라 한다. 나아가, 만약 f가 모든 $x \in U$에서 해석적이면 이때의 f를 **(실) 해석함수 ((real) analytic function)**이라 한다.

Theorem 0.44 열린집합 $U \subseteq \mathbb{R}$에서 정의된 함수 $f : U \to \mathbb{R}$와 한 점 $x_0 \in U$를 생각하자. 만약 적당한 수열 $\{a_i\}$와 임의의 $x \in U$에 대해 $f(x) = \sum_{i=0}^{\infty} a_i (x - x_0)^i$가 성립하면 f는 x_0에서 해석적이고 각 $i \in \mathbb{N}$에 대해 $a_i = f^{(i)}(x_0)/i!$이다.

Proof 멱급수 $\sum_{i=0}^{\infty} a_i(x-x_0)^i$가 임의의 $x \in U$에 대해 수렴하므로 함수 f는 $\mathscr{C}^\infty$급이고 각 $i \in \mathbb{N}$에 대해 $f(x) = \sum_{j=0}^{\infty} a_j(x-x_0)^j$의 양변을 i번 미분하면 $f^{(i)}(x) = \sum_{j=i}^{\infty} j^{\underline{i}} a_j(x-x_0)^{j-i}$에서 $f^{(i)}(x_0)/i! = a_i$이다. 이로부터 f가 x_0에서 해석적임은 자명하므로 증명이 끝난다. □

Theorem 0.45 열린집합 $U \subseteq \mathbb{R}$에서 정의된 $\mathscr{C}^\infty$급 함수 $f, g : U \to \mathbb{R}$에 대해 이가 모두 $x_0 \in U$에서 해석적이라 하면 함수 $f \pm g, fg, f/g$는 모두 x_0에서 해석적이다. 단, 이때 f/g의 경우 $g(x_0) \neq 0$이라 하자.

Proof 가정으로부터 적당한 x_0의 근방 $V \subseteq U$가 존재하여 임의의 $x \in V$에 대해 $f(x) = \sum_{i=0}^{\infty} f^{(i)}(x_0)(x-x_0)^i/i!$이고 $g(x) = \sum_{i=0}^{\infty} g^{(i)}(x_0)(x-x_0)^i/i!$이다. 이는 곧 두 멱급수의 수렴반경이 양수임을 뜻하므로 이를 각각 $R, S > 0$라 하면 $V \subseteq (x_0-R, x_0+R) \cap (x_0-S, x_0+S)$이다. 이제 함수 $\widetilde{f} : (x_0-R, x_0+R) \to \mathbb{R}, \widetilde{g} : (x_0-S, x_0+S) \to \mathbb{R}$를 각각 $\widetilde{f} : x \mapsto \sum_{i=0}^{\infty} f^{(i)}(x_0)(x-x_0)^i/i!, \widetilde{g} : x \mapsto \sum_{i=0}^{\infty} g^{(i)}(x_0)(x-x_0)^i/i!$로 두면 정리 0.42로부터 적당한 $T > 0$와 적당한 수열 $\{a_i\}$가 존재하여 임의의 $x \in (x_0-T, x_0+T)$에 대해

$$
\begin{aligned}
(\widetilde{f}+\widetilde{g})(x) &= \sum_{i=0}^{\infty} [f^{(i)}(x_0)+g^{(i)}(x_0)]\frac{(x-x_0)^i}{i!} \\
(\widetilde{f}\widetilde{g})(x) &= \sum_{i=0}^{\infty}\sum_{j=0}^{i} f^{(j)}(x_0)g^{(i-j)}(x_0)\frac{(x-x_0)^i}{i!} \\
(\widetilde{f}\widetilde{g})(x) &= \sum_{i=0}^{\infty} a_i \frac{(x-x_0)^i}{i!}
\end{aligned}
$$

이다. 이제 $\widetilde{f}|_V = f|_V, \widetilde{g}|_V = g|_V$와 정리 0.44로부터 증명이 끝난다. □

0.1.6 Multivariate Power Series

Definition 0.46 수열 $\{a_\alpha\}$와 $z \in \mathbb{C}^n$에 대해 $\sum_{\alpha \in \mathbb{N}_0^n} a_\alpha z^\alpha$로 주어진 급수를 **(다변수) 멱급수 ((multivariate) power series)**라 한다. 나아가, 어떤 멱급수를 절대수렴하도록 하는 모든 $z \in \mathbb{C}^n$의 집합을 그 멱급수의 **수렴영역 (domain of convergence)**이라 한다.

Definition 0.47 열린집합 $U \subseteq \mathbb{C}^n$가 임의의 $z \in U$와 임의의 $\theta \in \mathbb{R}^n$에 대해 $(z_i e^{\mathbf{i}\theta_i}) \in U$를 만족하면 이때의 집합 U를 **Reinhardt domain**이라 한다. 특별히, Reinhardt domain $D \subseteq \mathbb{C}^n$가 임의의 $z \in D$에 대해 $\{w \in \mathbb{C}^n : |w_1| \le |z_1|, \cdots, |w_n| \le |z_n|\} \subseteq D$를 만족하면 이때의 Reinhardt domain D를 **complete**하다고 한다. 또한, Reinhardt domain $D \subseteq \mathbb{C}^n$가 임의의 $z, w \in D$와 임의의 $t \in [0, 1]$에 대해 $(z_i^t w_i^{1-t}) \in D$를 만족하면 이때의 Reinhardt domain D를 **log−볼록 (− convex)**하다고 한다.

Theorem 0.48 멱급수 $\sum_{\alpha\in\mathbb{N}_0^n} a_\alpha z^\alpha$ 에 대해 이의 수렴영역은 log−볼록한 complete Reinhardt domain이다.

Proof 만약 $n=1$ 이면 정리가 자명하므로 $n>1$ 이라 하고, 주어진 멱급수의 수렴영역을 $D\subseteq\mathbb{C}^n$ 라 하자. 먼저 D가 열려있음을 보이기 위해 임의의 $(z,w)\in\mathbb{C}^{n-1}\times\mathbb{C}$에 대해 $\sum_{\alpha\in\mathbb{N}_0^n} a_\alpha(z,w)^\alpha$ 가 절대수렴할 필요충분조건은 각 $i\in\mathbb{N}_0$에 대해 $\sum_{\beta\in\mathbb{N}_0^{n-1}} a_{\beta i}z^\beta$ 가 절대수렴하고 $\sum_{i=0}^\infty\sum_{\beta\in\mathbb{N}_0^{n-1}} a_{\beta i}z^\beta w^i$ 도 절대수렴하는 것임을 상기하자. 이로부터 각 $i\in\mathbb{N}_0$에 대해 $\sum_{\beta\in\mathbb{N}_0^{n-1}} a_{\beta i}z^\beta$ 의 수렴영역을 $D_i\subseteq\mathbb{C}^{n-1}$ 라 하고 $A=\bigcap_{i=0}^\infty D_i$ 라 하여 각 $z\in A$에 대한 $\sum_{i=0}^\infty(\sum_{\beta\in\mathbb{N}_0^{n-1}} a_{\beta i}z^\beta)w^i$ 의 수렴반경을 $R_z\geq 0$ 라 하면 $D=\bigcup_{z\in A}B(R_z)$ 에서 D가 열려있음을 안다.

다음으로, 수렴영역 D가 log−볼록한 complete Reinhardt domain임을 보이자. 먼저 임의의 $z\in D$와 임의의 $\theta\in\mathbb{R}^n$에 대해 $\sum_{\alpha\in\mathbb{N}_0^n}|a_\alpha(z_ie^{\mathbf{i}\theta_i})^\alpha|=\sum_{\alpha\in\mathbb{N}_0^n}|a_\alpha z^\alpha|<\infty$ 이므로 $(z_ie^{\mathbf{i}\theta_i})\in D$ 이다. 나아가, 임의의 $z,w\in D$와 임의의 $t\in[0,1]$에 대해 Hölder의 부등식으로부터 $\sum_{\alpha\in\mathbb{N}_0^n}|a_\alpha(z_i^tw_i^{1-t})^\alpha|=\sum_{\alpha\in\mathbb{N}_0^n}|a_\alpha z^\alpha|^t|a_\alpha w^\alpha|^{1-t}\leq(\sum_{\alpha\in\mathbb{N}_0^n}|a_\alpha z^\alpha|)(\sum_{\alpha\in\mathbb{N}_0^n}|a_\alpha w^\alpha|)<\infty$ 에서 $(z_i^tw_i^{1-t})\in D$ 이다. 마지막으로 임의의 $z\in D$를 택하면 각 $i\leq n$에 대해 $|w_i|\leq|z_i|$ 인 $w\in\mathbb{C}^n$에 대해 $\sum_{\alpha\in\mathbb{N}_0^n}|a_\alpha w^\alpha|\leq\sum_{\alpha\in\mathbb{N}_0^n}|a_\alpha z^\alpha|<\infty$ 에서 $\{w\in\mathbb{C}^n:|w_1|\leq|z_1|,\cdots,|w_n|\leq|z_n|\}\subseteq D$ 이고, 이상으로부터 수렴영역이 log−볼록한 complete Reinhardt domain임을 안다. □

Theorem 0.49 멱급수 $\sum_{\alpha\in\mathbb{N}_0^n} a_\alpha z^\alpha$ 의 수렴영역 $D\subseteq\mathbb{C}^n$ 에 대해 함수 $f:D\cap\mathbb{R}^n\to\mathbb{R}$ 를 $f:x\mapsto\sum_{\alpha\in\mathbb{N}_0^n} a_\alpha x^\alpha$ 로 두면 $D\cap\mathbb{R}^n$ 은 $\mathbb{R}^n$ 에 대해 열려있고, f 는 $\mathscr{C}^\infty$ 급이다. 특별히, 임의의 $\beta\in\mathbb{N}_0^n$ 와 임의의 $x\in D\cap\mathbb{R}^n$ 에 대해 멱급수 $\sum_{\alpha\geq\beta} a_\alpha\alpha^{\underline{\beta}}x^{\alpha-\beta}$ 가 절대수렴하고 $\partial^\beta f(x)=\sum_{\alpha\geq\beta} a_\alpha\alpha^{\underline{\beta}}x^{\alpha-\beta}$ 이다.

Proof 만약 $n=1$ 이면 정리가 자명하므로 $n>1$ 이라 하자. 우선 $D\cap\mathbb{R}^n$ 가 $\mathbb{R}^n$ 에 대해 열려있음을 보이기 위해 임의의 $x\in D\cap\mathbb{R}^n$ 를 택하자. 그렇다면 D가 열려있으므로 적당한 $r>0$ 이 존재하여 $\prod_{i=1}^n B(x_i,r)\subseteq D$ 이고, 곧 $x\in(x-r\mathbf{1},x+r\mathbf{1})\subseteq D\cap\mathbb{R}^n$ 에서 $D\cap\mathbb{R}^n$ 가 $\mathbb{R}^n$ 에 대해 열려있음을 안다.

다음으로, 정리의 남은 부분을 보이기 위해 임의의 $\beta\in\mathbb{N}_0^n$ 를 택하되, 만약 $\beta=0$ 이면 더 이상 보일 것이 없으므로 $\beta\neq 0$ 이라 하고, WLOG, $\beta_n\neq 0$ 이라 하자. 한편, 임의의 $x\in D\cap\mathbb{R}^n$ 에 대해 주어진 멱급수가 절대수렴하므로 합의 순서를 바꾸어 $f:x\mapsto\sum_{i=0}^\infty\sum_{\gamma\in\mathbb{N}_0^{n-1}} a_{\gamma i}(x_1,\cdots,x_{n-1})^\gamma x_n^i$ 로 쓸 수 있다. 따라서 임의의 $x\in D\cap\mathbb{R}^n$ 을 택하여 그 중 처음 $n-1$ 개의 성분을 $y\in\mathbb{R}^{n-1}$ 라 하면 $f(x)=\sum_{i=0}^\infty\sum_{\gamma\in\mathbb{N}_0^{n-1}} a_{\gamma i}y^\gamma x_n^i$ 이고 $\sum_{i=1}^\infty\sum_{\gamma\in\mathbb{N}_0^{n-1}} a_{\gamma i}iy^\gamma x_n^{i-1}$ 가 절대수렴한다. 이는 곧 f가 x에서 x_n에 대해 편미분가능함을 뜻하며 합의 순서를 다시 바꾸어 $(\partial/\partial x_n)f(x)=\sum_{i=1}^\infty\sum_{\gamma\in\mathbb{N}_0^{n-1}} a_{\gamma i}iy^\gamma x_n^{i-1}=\sum_{\alpha\geq\mathbf{e}_n} a_\alpha\alpha^{\underline{\mathbf{e}_n}}x^{\alpha-\mathbf{e}_n}$ 로 쓸 수 있고, 이를 $|\beta|-1$ 번 반복하면 원하는 결과를 얻는다. 한편, 이로부터 f가 $\mathscr{C}^\infty$ 급임은 자명하다. □

0.2 Special Functions

이번 절에서는 이 책에서 사용될 특수함수에 대해 살펴본다. 이들 중 대부분은 통계학의 다양한 분포를 공부할 때 유용하게 사용될 것이다. 비록 특수함수라는 이름이 그다지 살갑게 느껴지지 않는 것은 사실이지만, 그렇다고 어렵게 생각할 필요는 없다. 따지고 보면 우리가 일상적으로 사용하는 삼각함수나 로그함수와 하나도 다를 바가 없다. 다만 이들을 우리가 고등학교 시절에 먼저 배워 익숙하게 느낄 뿐이다. 생각해보면, 다항함수나 유리함수와는 달리 삼각함수나 로그함수의 경우 특정한 점에서의 함숫값을 수치해석적인 방법을 사용하지 않고 손으로 계산하는 것은 거의 불가능하다. 다만, $\sin\pi/3 = \sqrt{3}/2$와 같이 몇몇 특별한 함숫값과 $\sin^2\theta + \cos^2\theta = 1$과 같은 공식들을 외워 사용할 따름이다. 특수함수도 이와 마찬가지이다. 비록 그 함숫값을 직접 계산하지 못하겠지만, 몇몇 공식만 외워 사용하면 어느덧 삼각함수 못지 않게 편안한 함수가 되어 있을 것이다.

0.2.1 Special Functions Related to e^{-x^2}

첫 번째로 살펴볼 특수함수들은 e^{-x^2}와 관련된 특수함들이다. 눈치챘겠지만 이들 대부분 정규분포를 공부할 때 필요한 것들이다. 이후에 살펴볼 다른 특수함수들에 비해 이론적으로 쉬우니 편안한 마음으로 시작해보자.

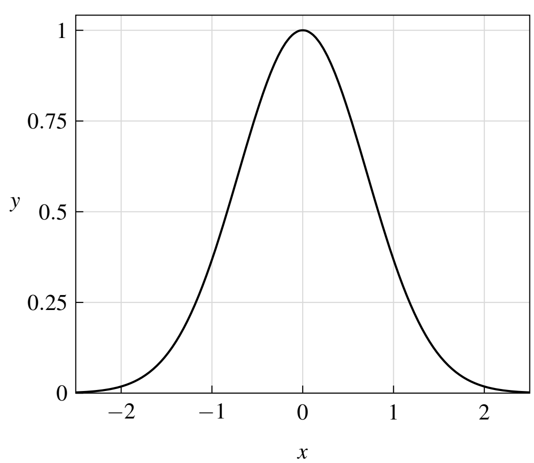


Figure 0.2 함수 $x \mapsto e^{-x^2}$ 의 그래프. Computed by Wolfram MATHEMATICA.

Definition 0.50 자연수 $k \in \mathbb{N}_0$에 대해 다음과 같이 정의되는 함수 $He_k : \mathbb{R} \to \mathbb{R}$을 k**次 Hermite 다항식** **(**k**th − polynomial)**이라 한다.

$$He_k : x \mapsto (-1)^k e^{x^2} \frac{d^k}{dx^k} e^{-x^2}$$

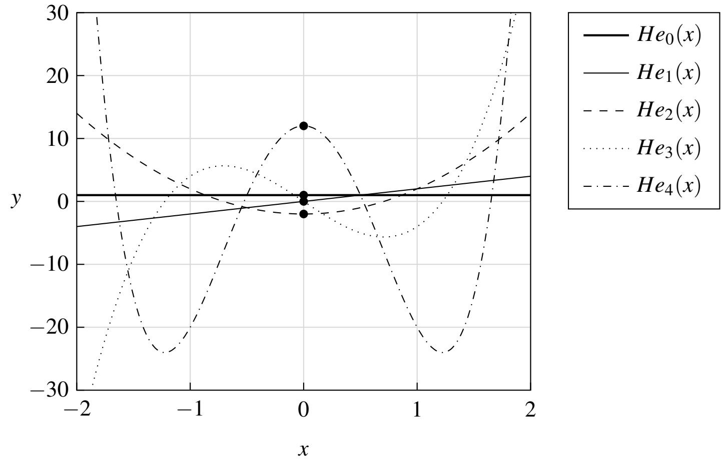


Figure 0.3 Hermite 다항식의 그래프. Computed by Wolfram MATHEMATICA.

Table 0.1 Hermite 다항식

차수 (k)	$He_k(x)$
0	1
1	$2x$
2	$4x^2 - 2$
3	$8x^3 - 12x$
4	$16x^4 - 48x^2 + 12$
5	$32x^5 - 160x^3 + 120x$
6	$64x^6 - 480x^4 + 720x^2 - 120$
7	$128x^7 - 1344x^5 + 3360x^3 - 1680x$
8	$256x^8 - 3584x^6 + 13440x^4 - 13440x^2 + 1680$
9	$512x^9 - 9216x^7 + 48384x^5 - 80640x^3 + 30240x$
10	$1024x^{10} - 23040x^8 + 161280x^6 - 403200x^4 + 302400x^2 - 30240$

함수 $x \mapsto e^{-x^2}$을 미분하게 되면 그 결과는 e^{-x^2}에 어떤 다항식이 곱해진 형태로 얻어지게 되는데, k차 Hermite 다항식은 $x \mapsto e^{-x^2}$을 k번 미분했을 때 그 앞에 곱해진 다항식이다.

Hermite 다항식은 물리학과 통계학에서 주로 사용되는데, 각각의 분야에서 조금 다르게 정의되며 위에서 소개한 정의는 물리학의 것으로 physicists' Hermite 다항식이라 하기도 한다. 그렇다고 크게 다르지는 않고, 단지 scaling과 정규화를 어떻게 하는가의 차이가 있을 뿐이라 우리는 그냥 위에서 정의된 Hermite 다항식을 사용하기로 한다. 한편, Hermite 다항식를 계산하기 위해서는 $x \mapsto e^{-x^2}$ 을 정말 k번 미분해봐도 되지만, 다음 점화식을 사용하면 편하다.

Theorem 0.51 임의의 $k \in \mathbb{N}_0$와 임의의 $x \in \mathbb{R}$에 대해

$$He_k(x) = \begin{cases} 2xHe_{k-1}(x) - He'_{k-1}(x) & k > 0\text{인 경우} \\ 1 & k = 0\text{인 경우} \end{cases}$$

이다.

Proof 우선 $He_0 = 1$임은 정의로부터 분명하고, 임의의 $k > 0$에 대해서는

$$\begin{aligned} He_k(x) &= (-1)^k e^{x^2} \frac{d^k}{dx^k} e^{-x^2} \\ &= (-1)^k e^{x^2} \frac{d}{dx} (-1)^{k-1} e^{-x^2} He_{k-1}(x) \\ &= -e^{x^2} [-2xHe_{k-1}(x) + He'_{k-1}(x)] e^{-x^2} \\ &= 2xHe_{k-1}(x) - He'_{k-1}(x) \end{aligned}$$

가 되어 정리가 성립한다. □

복잡하지만 Hermite 다항식을 다음과 같이 닫힌 형태로 정확히 얻어낼 수도 있다. 물론, 그 복잡함 때문에 크게 쓸모는 없다.

Theorem 0.52 임의의 $k \in \mathbb{N}_0$와 임의의 $x \in \mathbb{R}$에 대해

$$He_k(x) = k! \sum_{i=0}^{\lfloor k/2 \rfloor} \frac{(-1)^i}{i!(k-2i)!} (2x)^{k-2i}$$

이다.

Proof 수학적 귀납법을 사용하자. 우선 $He_0(x) = 1$이고 $He_1(x) = 2x$이므로 $k = 0, 1$에 대해서는 정리가 자명하다. 이제 귀납가정으로서 짝수 $k \in \mathbb{N}_0$에 대해 정리가 성립한다 하면 정리 0.51로부터

$$He_{k+1}(x) = 2xHe_k(x) - He'_k(x)$$

$$
\begin{aligned}
&= 2xk!\sum_{i=0}^{k/2}\frac{(-1)^i}{i!(k-2i)!}(2x)^{k-2i} - k!\sum_{i=0}^{k/2-1}2(k-2i)\frac{(-1)^i}{i!(k-2i)!}(2x)^{k-2i-1}\\
&= k!\left[\sum_{i=0}^{k/2}\frac{(-1)^i}{i!(k-2i)!}(2x)^{k-2i+1} - \sum_{i=0}^{k/2-1}2\frac{(-1)^i}{i!(k-2i-1)!}(2x)^{k-2i-1}\right]\\
&= k!\left[\sum_{i=1}^{k/2}(k-2i+1+2i)\frac{(-1)^i}{i!(k-2i+1)!}(2x)^{k-2i+1} + \frac{(2x)^{k+1}}{k!}\right]\\
&= k!\sum_{i=1}^{k/2}(k+1)\frac{(-1)^i}{i!(k-2i+1)!}(2x)^{k-2i+1} + (2x)^{k+1}\\
&= (k+1)!\sum_{i=0}^{\lfloor (k+1)/2\rfloor}\frac{(-1)^i}{i!(k+1-2i)!}(2x)^{k+1-2i}
\end{aligned}
$$

에서 정리가 성립하고, 홀수 $k \in \mathbb{N}_0$에 대해서도 비슷하게 하면 같은 결론을 얻으므로 증명이 끝난다. □

이어서 Hermite 다항식의 기본적인 성질들을 보자.

Theorem 0.53 임의의 $k \in \mathbb{N}_0$에 대해 다음이 성립한다.

i. k차 Hermite 다항식은 k차 다항식이다.
ii. k차 Hermite 다항식은 k개의 서로다른 영점을 가진다.
iii. $He_k(x) = (-1)^k He_k(-x)$.
iv. 만약 $k > 0$이면 $\lim_{x\to-\infty}(-1)^k He_k(x) = \infty$이다.
v. 만약 $k > 0$이면 $\lim_{x\to\infty} He_k(x) = \infty$이다.
vi. k차 Hermite 다항식은 $\mathscr{C}^\infty$급이며 $n \le k$인 $n \in \mathbb{N}_0$에 대해 $He_k^{(n)} = 2^n k^{\underline{n}} He_{k-n}$이다.
vii. k차 Hermite 다항식은 k가 짝수이면 짝수차수 항만을 가지고, k가 홀수이면 홀수차수 항만을 가진다.

Proof i, vii. 이는 정리 0.51와 수학적 귀납법으로부터 자명하다.

ii. 수학적 귀납법을 사용하자. 먼저 $k = 0, 1$인 경우에는 정리가 자명하므로 귀납가정으로서 $k \in \mathbb{N}$에 대해 He_k가 k개의 서로다른 영점으로 $x_1 < \cdots < x_k$를 가진다고 하고, 표기의 편의를 위해 함수 $f : \mathbb{R} \to \mathbb{R}$를 $f : x \mapsto e^{-x^2}$으로 두자. 그렇다면 정의로부터 $f^{(k)}$도 $x_1, \cdots, x_k$를 영점으로 가지고, Rolle의 정리로부터 서로다른 $y_1, \cdots, y_{k-1} \in (x_1, x_k)$가 존재하여 $f^{(k+1)}$의 영점이 된다. 한편, 명백히 $\lim_{x\to\pm\infty} f^{(k)}(x) = 0$이므로 $f^{(k+1)}$는 $(-\infty, x_1)$과 (x_k, ∞)에서도 적어도 하나씩의 영점을 가지고,[1] 이를 각각 $y_0 < x_1, y_k > x_k$라 하자. 그렇다면 이상으로부터 $f^{(k+1)}$는 서로다른 $y_0, \cdots, y_k$를 영점으로 가지고, 곧 He_{k+1}도 $k+1$개의 서로다른 영점을 가지므로 증명이 끝난다.

iii. 이는 vii로부터 자명하다.

iv, v. 정리 0.52로부터 He_k의 최고차항의 계수는 항상 양수임을 알 수 있고, 이로부터 iv와 v는 자명하다.

vi. 임의의 $k \in \mathbb{N}$에 대해 $He_k' = 2kHe_{k-1}$이 성립함을 증명하는 것으로 충분한데, 이는

$$\begin{aligned}
He_k'(x) &= \frac{d}{dx}(-1)^k e^{x^2}\frac{d^k}{dx^k}e^{-x^2} \\
&= (-1)^k 2xe^{x^2}\frac{d^k}{dx^k}e^{-x^2} + (-1)^k e^{x^2}\frac{d^{k+1}}{dx^{k+1}}e^{-x^2} \\
&= (-1)^k 2xe^{x^2}\frac{d^k}{dx^k}e^{-x^2} + (-1)^{k+1} e^{x^2}\frac{d^k}{dx^k}(2xe^{-x^2}) \\
&= (-1)^k 2xe^{x^2}\frac{d^k}{dx^k}e^{-x^2} + (-1)^{k+1} e^{x^2}\left(2x\frac{d^k}{dx^k}e^{-x^2} + 2k\frac{d^{k-1}}{dx^{k-1}}e^{-x^2}\right) \\
&= 2k(-1)^{k-1}e^{x^2}\frac{d^{k-1}}{dx^{k-1}}e^{-x^2} \\
&= 2kHe_{k-1}(x)
\end{aligned}$$

에서 자명하다. □

Hermite 다항식의 0에서의 함숫값은 특별한 이름이 붙어있다. 이를 살펴보는 것을 끝으로 다음 특수함수로 넘어가자.

Definition 0.54 k차 Hermite 다항식의 0에서의 함숫값 $He_k(0)$을 k**번째 Hermite 수 (**k**th − number)**라 하고 He_k로 쓴다.

Table 0.2 Hermite 수

k	0	1	2	3	4	5	6	7	8	9	10
He_k	1	0	−2	0	12	0	−120	0	1680	0	−30240

Theorem 0.55 임의의 $k \in \mathbb{N}_0$에 대해

$$He_k = \begin{cases} (-2)^{k/2}(k-1)!! & k\text{가 짝수인 경우} \\ 0 & k\text{가 홀수인 경우} \end{cases}$$

이다.

Proof 정리 0.53의 ii로부터 k가 홀수인 경우는 자명하다. 한편, k가 짝수인 경우에는 정리 0.52로부터 $He_k = k!(-1)^{k/2}/(k/2)! = (-2)^{k/2}k(k-1)\cdots 2\cdot 1/k(k-2)\cdots 4\cdot 2 = (-2)^{k/2}(k-1)!!$이다. □

고등학교 시절 문제를 많이 풀어봤다면 $\int e^{-x^2}\,dx$와 같은 적분을 본 적이 있을 것이고, 어린 마음에 이를 구하기 위해 온갖 시도를 해 보다 실패한 경험이 있을 것이다. 하지만 단언하건대, $x \mapsto e^{-x^2}$의 부정적분은 초등함수로 표현되지 않으니 너무 실망할 필요 없다. (이는 이미 증명된 사실이니 괜한 오기는 품지 말자.[2]) 다음 특수함수는 바로 이 적분에 관한 것이다.

Definition 0.56 다음과 같이 정의되는 함수 $\mathrm{erf} : \mathbb{R} \to \mathbb{R}$를 **오차함수 (error function)**라 한다.

$$\mathrm{erf} : x \mapsto \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\,dt$$

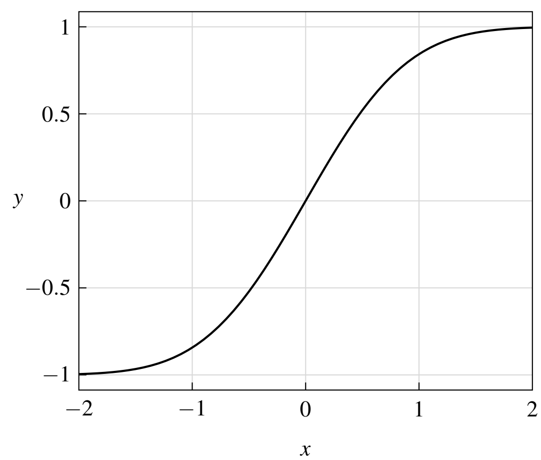


Figure 0.4 오차함수의 그래프. Computed by Wolfram MATHEMATICA.

누가 보더라도 정규분포의 CDF를 위해 도입한 것이 분명하니 이를 굳이 숨길 필요는 없을 것이다. 이는 나중에 정규분포와 그로부터 유도되는 몇몇 특별한 분포의 CDF를 표현하는데 사용된다. 그렇다면 오차함수의 기본적인 성질을 알아보자.

Proposition 0.57 (Gauss integral)

$$\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}$$

Proof 구하고자 하는 적분을 I라 하면 극좌표 변환을 통해

$$I^2 = \int_{\mathbb{R}^2} e^{-(x^2+y^2)}\,dx\,dy$$

$$
\begin{aligned}
&= \int_0^{\pi/2} \int_0^{\infty} e^{-r^2} r\,dr\,d\theta \\
&= \int_0^{\pi/2} \left[-\frac{1}{2} e^{-r^2} \right]_0^{\infty} d\theta \\
&= \int_0^{\pi/2} \frac{1}{2}\,d\theta \\
&= \frac{\pi}{4}
\end{aligned}
$$

임을 안다. □

Theorem 0.58 다음이 성립한다.

i. $-1 < \mathrm{erf} < 1$.
ii. 오차함수는 기함수이다.
iii. 오차함수는 $x = 0$에서 유일한 영점을 가진다.
iv. 오차함수는 순증가한다.
v. 오차함수는 $\mathbb{R}^-$에서 순볼록하고 $\mathbb{R}^+$에서 순오목하며 $x = 0$에서 변곡점을 가진다.
vi. $\lim_{x\to-\infty} \mathrm{erf}(x) = -1$.
vii. $\lim_{x\to\infty} \mathrm{erf}(x) = 1$.
viii. 오차함수는 $\mathscr{C}^\infty$급이며 $n \in \mathbb{N}$에 대해

$$\mathrm{erf}^{(n)} : x \mapsto (-1)^{n-1} \frac{2}{\sqrt{\pi}} He_{n-1}(x) e^{-x^2}$$

이다.

PROOF iv와 v를 제외한 나머지는 Hermite 다항식과 오차함수의 정의, Gauss 적분으로부터 자명하고, iv와 v의 경우에도 $\mathrm{erf}' : x \mapsto e^{-x^2} > 0, \mathrm{erf}'' : x \mapsto -2xe^{-x^2}$에서 자명하다. □

마지막으로 e^{-x^2}와 관련하여 알아볼 특수함수는 Gaussian 함수이다. 이건 그냥 표준정규분포의 PDF이다.

Definition 0.59 다음과 같이 정의되는 함수 $\phi : \mathbb{R} \to \mathbb{R}$를 **Gaussian 함수 (− function)** 혹은 간단히 **Gaussian**이라 한다.

$$\phi : x \mapsto \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

Theorem 0.60 다음이 성립한다.

i. $0 < \phi \le 1/\sqrt{2\pi}$.

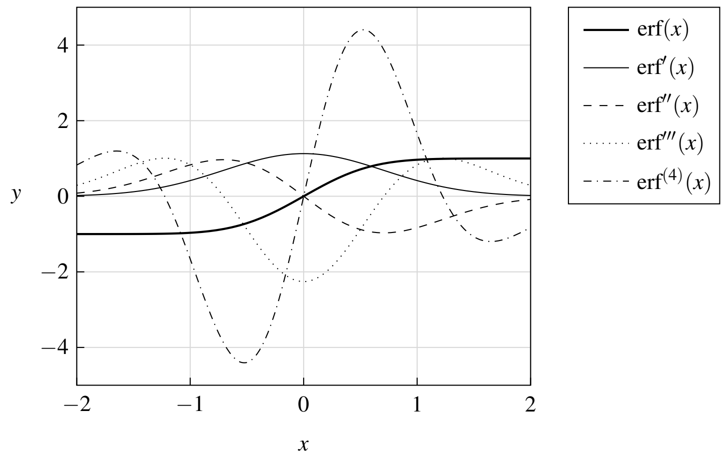


Figure 0.5 오차함수와 그 도함수들의 그래프. Computed by Wolfram MATHEMATICA.

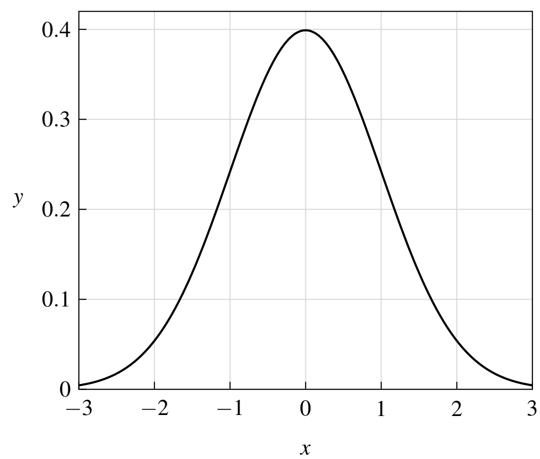


Figure 0.6 Gaussian의 그래프. Computed by Wolfram MATHEMATICA.

ii. Gaussian 함수는 $\mathbb{R}^-$에서 순증가하고 $\mathbb{R}^+$에서 순감소하며 $x=0$에서 최댓값 $1/\sqrt{2\pi}$를 가진다.
iii. Gaussian 함수는 $(-\infty,-1)$와 $(1,\infty)$에서 순볼록하고 $(-1,1)$에서 순오목하며 $x=\pm 1$에서 변곡점을 가진다.
iv. $\lim_{x\to-\infty}\phi(x)=\lim_{x\to\infty}\phi(x)=0$.
v. Gaussian 함수는 $\mathscr{C}^\infty$급이며 $n\in\mathbb{N}_0$에 대해

$$\phi^{(n)}: x\mapsto \frac{(-1)^n}{\sqrt{\pi}2^{(n+1)/2}}He_n\left(\frac{x}{\sqrt{2}}\right)e^{-x^2/2}$$

이다.

Proof ii와 iii을 제외한 나머지는 Hermite 다항식과 Gaussian 함수의 정의로부터 자명하고, ii와 iii도 $\phi' : x \mapsto -xe^{-x^2/2}/\sqrt{2\pi} = -x\phi(x)$, $\phi'' : x \mapsto (x^2-1)e^{-x^2/2}/\sqrt{2\pi}$에서 자명하다. □

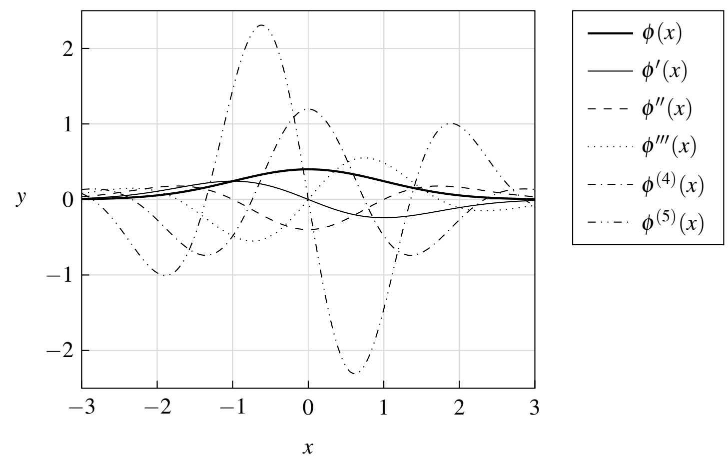


Figure 0.7 Gaussian과 그 도함수들의 그래프. Computed by Wolfram MATHEMATICA.

앞서 Hermite 다항식을 소개하며 물리학과 통계학에서 조금씩 다르게 정의한다고 했는데, 통계학에서 주로 사용하는 k차 Hermite 다항식은 $2^{-k/2}He_k(x/\sqrt{2})$로 정의되며 $\widetilde{He}_k(x)$로 쓴다. 이를 이용하면 위의 정리의 v를 $\phi^{(n)} : x \mapsto (-1)^n\widetilde{He}_k(x)e^{-x^2/2}/\sqrt{2\pi}$로 조금 더 간단하게 쓸 수 있다. 한편, 다음 따름정리는 기억해 둘 가치가 있다.

Corollary 0.61 임의의 $x \in \mathbb{R}$에 대해 $\phi'(x) = -x\phi(x)$이다.

Proof 이는 위의 정리로부터 자명하다. □

0.2.2 Special Functions Related to Gamma Function

두 번째로 살펴볼 특수함수는 감마함수라는 특수함수와 이로부터 유도되는 특수함수들이다. 감마함수가 대체 무엇이길래 이로부터 유도되는 특수함수까지 살펴봐야 하는지는 차차 알게 될 것이다. 한편, 지금까지 살펴봤던 특수함수들은 고등학교 수준의 미적분 지식으로도 충분히 관련 성질을 보일 수 있었다면 지금부터는 해석학의 지식이 조금씩 필요하다.

Definition 0.62 다음과 같이 정의되는 함수 $\Gamma:\mathbb{R}^+\to\mathbb{R}$를 **감마함수 (gamma function)**라 한다.

$$\Gamma : x \mapsto \int_0^\infty t^{x-1}e^{-t}\,dt$$

오차함수와는 달리 감마함수는 그 피적분함수가 $\mathbb{R}^+$에서 적분가능한지를 따져 well-define 됨을 확인할 필요가 있다.

Proposition 0.63 임의의 $p\in\mathbb{R}$에 대해 $I(p)=\int_0^\infty x^{p-1}e^{-x}\,dx$라 하면 $p>0$일 때 $I(p)<\infty$이고, 그렇지 않으면 $I(p)=\infty$이다.

Proof 먼저 $p>0$인 경우를 생각하면 적당한 $M>0$이 존재하여 $x\geq M$이면 $x^{p-1}<e^{x/2}$이므로 $\int_1^\infty x^{p-1}e^{-x}\,dx\leq\int_1^M x^{p-1}e^{-x}\,dx+\int_M^\infty e^{-x/2}\,dx=\int_1^M x^{p-1}e^{-x}\,dx+2e^{-M/2}<\infty$이다. 따라서 $\int_0^1 x^{p-1}e^{-x}\,dx<\infty$임을 보이면 이 경우에 대해서는 충분한데, 여기서 $p\geq1$이면 피적분함수 $x^{p-1}e^{-x}$가 $[0,1]$에서 유계이므로 $\int_0^1 x^{p-1}e^{-x}\,dx<\infty$이고, $0<p<1$이라도 $\int_0^1 x^{p-1}e^{-x}\,dx\leq\int_0^1 x^{p-1}\,dx=[x^p/p]_0^1=1/p<\infty$이므로 곧 $p>0$이면 $I(p)<\infty$이다. 한편, $p\leq0$인 경우에는 모든 $x\in\mathbb{R}$에 대해 $e^{-x}\geq1-x$이므로 $I(p)\geq\int_0^1 x^{p-1}e^{-x}\,dx\geq\int_0^1 e^{-x}/x\,dx\geq\int_0^1(1-x)/x\,dx=[\log x-x]_0^1=\infty$이다. □

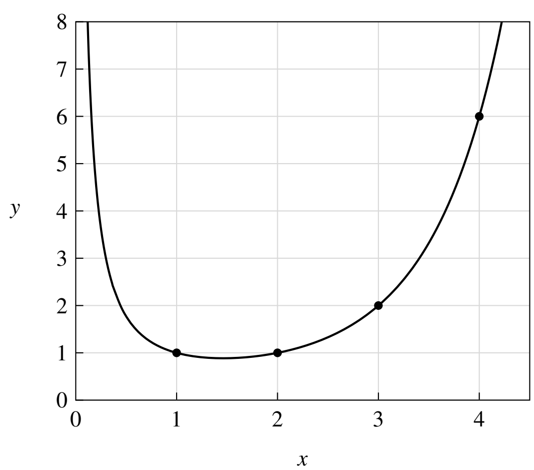


Figure 0.8 감마함수의 그래프. Computed by Wolfram MATHEMATICA.

다음 순서는 감마함수의 기본적인 성질들이다. 여기서 ii의 증명에 사용된 Hölder의 부등식과 Leibniz의 법칙은 1장의 측도론에서 각각 정리 ? ? 와 따름정리 ? ? 로 증명되어 있으니 잠깐 넘어가 결과 정도만 참조하는 것을 추천한다.

Theorem 0.64 다음이 성립한다.

i. 감마함수는 양함수이다.
ii. 감마함수는 적당한 $x_* \in (1,2)$에 대해 $(0, x_*)$에서 순감소하고 (x_*, ∞)에서 순증가한다.
iii. 감마함수는 log-순볼록하다.
iv. $\lim_{x\downarrow 0}\Gamma(x) = \lim_{x\to\infty}\Gamma(x) = \infty$.
v. 감마함수는 $\mathscr{C}^\infty$급이며 $n \in \mathbb{N}_0$에 대해

$$\Gamma^{(n)} : x \mapsto \int_0^\infty (\log t)^n t^{x-1} e^{-t}\,dt$$

이다.

Proof i. 이는 자명하다.

iii. 임의의 서로다른 $x, y > 0$와 임의의 $t \in (0,1)$에 대해 Hölder의 부등식으로부터

$$\begin{aligned}\Gamma(tx+(1-t)y) &= \int_0^\infty s^{tx+(1-t)y-1}e^{-s}\,ds \\ &= \int_0^\infty (s^{x-1}e^{-s})^t (s^{y-1}e^{-s})^{1-t}\,ds \\ &\le \left(\int_0^\infty s^{x-1}e^{-s}\,ds\right)^t \left(\int_0^\infty s^{y-1}e^{-s}\,ds\right)^{1-t} \\ &= [\Gamma(x)]^t[\Gamma(y)]^{1-t}\end{aligned}$$

이다. 만약 여기서 등호가 성립한다면 곧 모두 0은 아닌 적당한 $a, b \ge 0$가 존재하여 거의 대부분의 $s > 0$에 대해 $as^{x-1}e^{-s} + bs^{y-1}e^{-s} = (as^{x-y}+b)s^{y-1}e^{-s} = 0$인데, 이는 불가능하므로 위에서 등호는 성립하지 않는다. 따라서 감마함수는 log-순볼록하다.

iv. 임의의 $x > 0$에 대해 $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt \ge \int_0^1 t^{x-1}e^{-t}\,dt \ge \int_0^1 t^{x-1}(1-t)\,dt = 1/x - 1/(x+1)$이므로 $x \downarrow 0$이면 $\Gamma(x) \to \infty$이다. 여기서 두 번째 부등호는 모든 $t \in \mathbb{R}$에 대해 $e^{-t} \ge 1-t$이므로 성립한다. 비슷하게, 임의의 $x > 0$에 대해 $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt \ge \int_2^\infty t^{x-1}e^{-t}\,dt \ge 2^{x-1}\int_2^\infty e^{-t}\,dt = 2^{x-1}/e^2$이므로 $x \to \infty$이면 $\Gamma(x) \to \infty$이다.

v. 증명은 n에 대한 수학적 귀납법을 사용한다. 먼저 $n = 0$인 경우에는 정리가 자명하므로 귀납가정으로서 $n \in \mathbb{N}_0$에 대해 정리가 성립하다고 하자. 또한, 임의의 $x_0 > 0$를 고정하고 $x_0 \in (a,b) =: I$인 적당한 $a, b > 0$를 택하여 함수 $f : \mathbb{R}^+ \times I \to \mathbb{R}$를 $f : (t,x) \mapsto (\log t)^n t^{x-1} e^{-t}$로 정의하자. 그렇다면 f는 Leibniz의 법칙의 조건 i, ii, iii을 만족함이 분명하고 임의의 $(t,x) \in \mathbb{R}^+ \times I$에 대해

$$\begin{aligned}\left|\frac{\partial}{\partial x}f(t,x)\right| &= |\log t|^{n+1}t^{x-1}e^{-t} \\ &\le t^{x+n}e^{-t}\mathbf{1}_{[1,\infty)}(t) + |\log t|^{n+1}t^{x-1}\mathbf{1}_{(0,1)}(t)\end{aligned}$$

$$\begin{aligned}&\le t^{b+n}e^{-t}\mathbf{1}_{[1,\infty)}(t)+|\log t|^{n+1}t^{a-1}\mathbf{1}_{(0,1)}(t)\\ &\le t^{b+n}e^{-t}+(-\log t)^{n+1}t^{a-1}\mathbf{1}_{(0,1)}(t)\end{aligned}$$

이다. 여기서 첫 번째 부등호는 모든 $t\in\mathbb{R}^+$에 대해 $\log t\le t$이므로 성립한다. 한편, $\int_0^\infty t^{b+n}e^{-t}\,dt=\Gamma(n+b+1)$이고

$$\begin{aligned}\int_0^1(-\log t)^{n+1}t^{a-1}\,dt&=\left[(-\log t)^{n+1}\frac{t^a}{a}\right]_0^1+\frac{n+1}{a}\int_0^1(-\log t)^n t^{a-1}\,dt\\ &=\cdots\\ &=(n+1)!\left[\frac{t^a}{a^{n+2}}\right]_0^1\\ &=\frac{(n+1)!}{a^{n+2}}\end{aligned}$$

에서 $\mathbb{R}^+$에서 정의된 함수 $t\mapsto t^{b+n}e^{-t}+(-\log t)^{n+1}t^{a-1}\mathbf{1}_{(0,1)}(t)$가 적분가능하여 f가 Leibniz의 법칙의 조건 iv도 만족시킴을 알고, 곧 이로부터 I에서 정의된 함수 $x\mapsto\int_0^\infty(\log t)^n t^{x-1}e^{-t}\,dt$는 x_0에서 미분가능하며 그 미분계수는 $\int_0^\infty(\log t)^{n+1}t^{x_0-1}e^{-t}\,dt$로 주어진다. 증명은 이로써 충분하다.

ii. 간단한 계산을 통해 $\Gamma(1)=\Gamma(2)=1$임을 알 수 있으므로, MVT에서 $\Gamma'(x_*)=0$인 $x_*\in(1,2)$가 존재함을 안다. 한편, iii으로부터 감마함수가 순볼록하므로 Γ'가 순증가하여 정리가 성립한다. □

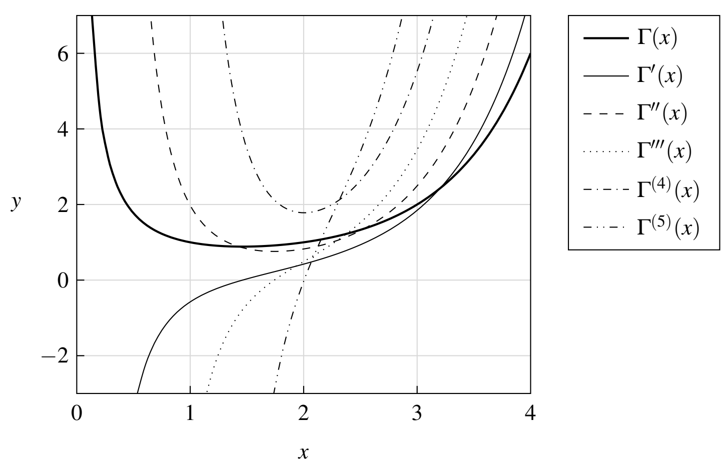


Figure 0.9 감마함수와 그 도함수들의 그래프. Computed by Wolfram MATHEMATICA.

수치해석적인 방법을 사용하면 감마함수는 $x_* \approx 1.46163$에서 최솟값 $\Gamma(x_*) \approx 0.885603$을 가짐을 알 수 있다. 한편, 다음 결과는 감마함수를 그토록 유명하게 만든 결과이다. 그리고 사실상 감마함수의 본질이기도 하다.

Theorem 0.65 임의의 $x > 0$에 대해 $\Gamma(x+1) = x\Gamma(x)$이다.

Proof 이는 $\Gamma(x+1) = \int_0^\infty t^x e^{-t}\,dt = [-t^x e^{-t}]_0^\infty + x\int_0^\infty t^{x-1}e^{-t}\,dt = x\Gamma(x)$로부터 자명하다. □

Corollary 0.66 임의의 $n \in \mathbb{N}$에 대해 다음이 성립한다.

i. $\Gamma(n+1) = n!$.
ii. $\Gamma(n+1/2) = \sqrt{\pi}(2n-1)!!/2^n$. 단, 여기서 $(-1)!! = 1$로 생각한다.

Proof i과 ii모두 정리 0.65로부터 거의 자명하다. i의 경우 $\Gamma(1) = \int_0^\infty e^{-t}\,dt = 1$에서 $\Gamma(n+1) = n\Gamma(n) = \cdots = n!\Gamma(1) = n!$임을 알 수 있고, ii의 경우에도 비슷하게 Gauss 적분으로부터 $\Gamma(1/2) = \int_0^\infty t^{-1/2}e^{-t}\,dt = 2\int_0^\infty e^{-s^2}\,ds = \sqrt{\pi}$이므로 $\Gamma(n+1/2) = (n-1/2)\Gamma(n-1/2) = \cdots = (n-1/2)^{\underline{n}}\Gamma(1/2) = \sqrt{\pi}(2n-1)!!/2^n$임을 안다. □

곧 감마함수는 계승의 확장으로 문헌에 따라 계승을 나타내는 !를 그대로 사용하여 $\Gamma(x+1) = x!$로 쓰는 경우도 있다. 각종 수학 교양도서에서 이미 하도 많이 소개한 까닭에 큰 감동으로 다가오지 않을 수도 있겠지만, 감마함수가 계승의 확장이라는 단 하나의 사실이 감마함수를 그토록 유용하고 유명하게 만든 것이다. 게다가 단순한 확장이 아니라 무려 $\mathscr{C}^\infty$급 함수로의 확장이니 엄청난 결과임에는 틀림이 없다. 사족으로, 위의 따름정리의 증명에서 $\Gamma(1/2) = \sqrt{\pi}$라는 결과는 기억해두면 좋다.

또한, 감마함수하면 빠뜨릴 수 없는 근사 공식이 있다. 이 공식은 감마함수가 뒤섞인 극한식을 계산할 때 유용하게 사용된다.

Theorem 0.67 (Stirling's formula) 임의의 $x > 0$에 대해 다음이 성립한다.

$$\left|\frac{\Gamma(x)}{\sqrt{2\pi/x}(x/e)^x} - 1\right| \le \sqrt{\frac{2}{\pi x}}$$

Proof 임의의 $x > 0$를 고정하면 $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt = \int_{-1}^\infty [x(s+1)]^{x-1}e^{-x(s+1)}x\,ds = x^x e^{-x}\int_{-1}^\infty (s+1)^{x-1}e^{-sx}\,ds$이다. 여기서 두 번째 등호는 $s = t/x - 1$의 변수변환으로부터 성립한다. 이제 $I = \int_{-1}^\infty (t+1)^{x-1}e^{-tx}\,dt$로 두면 $|I - \sqrt{2\pi/x}| \le 2/x$임을 보이는 것으로 증명은 충분하다. 이를 위해 함수 $f : (-1,\infty) \to \mathbb{R}$를

$$f : t \mapsto \begin{cases} \sqrt{t - \log(t+1)} & t \ge 0\text{인 경우} \\ -\sqrt{t - \log(t+1)} & \text{ow.} \end{cases}$$

로 두면 이는 순증가하는 전단사함수이며 $[f(t)]^2 = t - \log(t+1)$ 임을 쉽게 알 수 있다. 또한, 표기의 편의를 위해 $g = f^2$ 이라 하면 g 는 $\mathscr{C}^\infty$ 급 함수가 되어 Taylor의 정리로부터 임의의 $t > -1$ 에 대해 0과 t 사이의 적당한 $s_t \in \mathbb{R}$ 가 존재하여 $g(t) = g(0) + g'(0)t + g''(s_t)t^2/2 = t^2/2(s_t+1)^2$ 이므로 $f(t)/t = 1/\sqrt{2}(s_t+1) > 0$ 이고, 곧 $|f(t)/t - 2^{-1/2}| = |(s_t+1)^{-1} - 1|/\sqrt{2} = |s_t|/\sqrt{2}(s_t+1) = |s_t|f(t)/t \le |f(t)|$ 이다. 이로부터

$$\begin{aligned}
I &= \int_{-1}^{\infty} (t+1)^{x-1} e^{-tx}\, dt \\
&= \int_{-1}^{\infty} \exp(-x(t - \log(t+1))) \frac{1}{t+1}\, dt \\
&= \int_{-\infty}^{\infty} e^{-xs^2} \frac{2s}{f^{-1}(s)}\, ds && \because s = f(t)
\end{aligned}$$

에서

$$\begin{aligned}
\left| I - \sqrt{2}\int_{-\infty}^{\infty} e^{-xs^2}\, ds \right| &= 2\left| \int_{-\infty}^{\infty} e^{-xs^2} \left[\frac{s}{f^{-1}(s)} - \frac{1}{\sqrt{2}} \right] ds \right| \\
&\le 2\int_{-\infty}^{\infty} e^{-xs^2} \left| \frac{f(f^{-1}(s))}{f^{-1}(s)} - \frac{1}{\sqrt{2}} \right| ds \\
&\le 2\int_{-\infty}^{\infty} e^{-xs^2} |s|\, ds \\
&= 4\int_{0}^{\infty} s e^{-xs^2}\, ds
\end{aligned}$$

이고, 여기서 Gauss 적분으로부터 $\int_{-\infty}^{\infty} e^{-xs^2}\, ds = 2\int_0^\infty e^{-xs^2}\, ds = 2\int_0^\infty e^{-u^2}/\sqrt{x}\, du = \sqrt{\pi/x}$ 이고 $\int_0^\infty s e^{-xs^2}\, ds = -[e^{-xs^2}]_0^\infty/2x = 1/2x$ 이므로 이상을 종합하면 $|I - \sqrt{2\pi/x}| \le 2/x$ 임을 안다. □

Corollary 0.68

$$\lim_{x\to\infty} \frac{\Gamma(x)}{\sqrt{2\pi/x}(x/e)^x} = 1$$

Proof 이는 Stirling의 공식으로부터 자명하다. □

Stirling의 공식을 통해 우리는 감마함수의 함숫값을 적당히 근사할 수 있고, 오차한계까지 정확히 구할 수 있다. 그리고 대부분의 경우 Stirling의 공식을 이용한 근사는 실용적인 기준에서 만족스러운 결과를 제공한다. 다만, Stirling의 공식은 충분히 큰 $x > 0$ 에 대해 감마함수의 참값 $\Gamma(x)$ 와 근삿값 $\sqrt{2\pi/x}(x/e)^x$ 의 비, 즉 상대오차가 1로 가까워짐을 함의하는 것이지 참값과 근삿값의 차, 즉 절대오차가 0으로 가까워짐을 함의하는 것은 아니라는 점에 주의해야 한다.

지금까지 감마함수의 성질에 대해서는 충분히 봤으니, 우리는 이제 클라이맥스로 향한다.

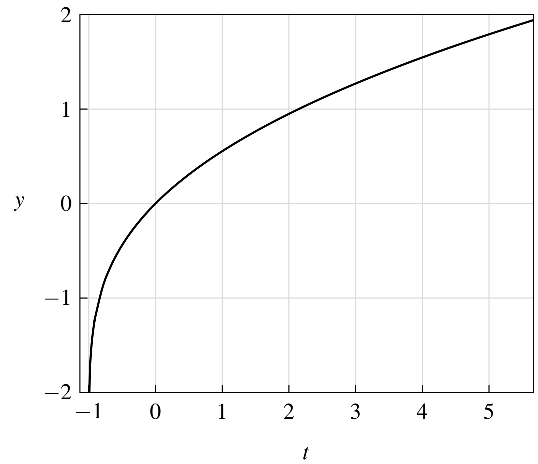


Figure 0.10 Stirling의 공식의 증명에서의 함수 $f(t)$의 그래프. Computed by Wolfram MATHEMATICA.

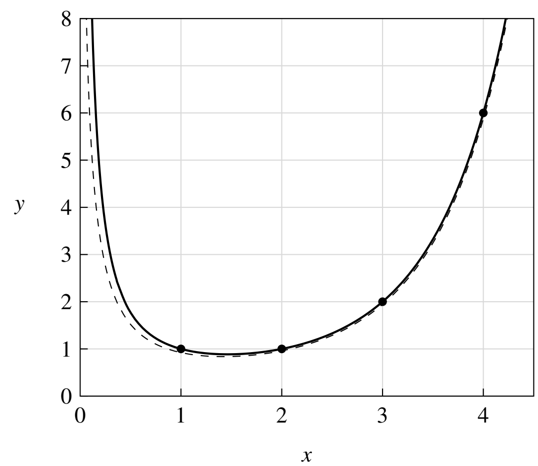


Figure 0.11 감마함수(실선)의 Stirling 근사(점선). Computed by Wolfram MATHEMATICA.

Definition 0.69 수열 $\{H_i\}$를

$$H_i = \sum_{j=1}^{i} \frac{1}{j}$$

라 하자. 이때 $\{H_i - \log i\}$의 극한값을 **Euler-Mascheroni 상수 (− constant)** 혹은 간단히 **Euler 상수 (− constant)**라 하고 γ로 쓴다. 또한, 이때의 수열 $\{H_i\}$를 **조화수열 (harmonic number)**이라 한다.

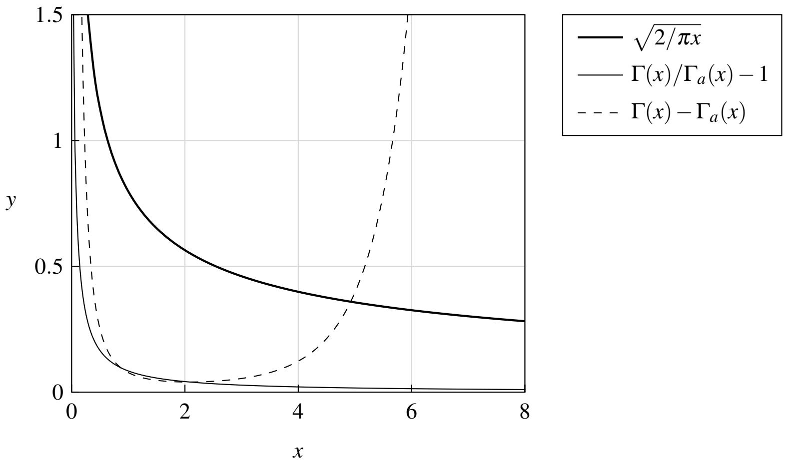


Figure 0.12 감마함수 $\Gamma(x)$의 Stirling 근사 $\Gamma_a(x)$의 상대오차(실선)와 절대오차(점선). 상대오차는 이론적으로 주어지는 오차한계(굵은선)보다 작으며 x가 커짐에 따라 빠르게 0으로 수렴하지만 절대오차는 오히려 빠르게 증가한다. Computed by Wolfram MATHEMATICA.

Table 0.3 조화수열

i	1	2	3	4	5	6	7	8
H_i	1	$3/2$	$11/6$	$25/12$	$137/60$	$49/20$	$363/140$	$761/280$
근삿값	1	1.5	1.83333	2.08333	2.28333	2.45	2.59286	2.71786

Proposition 0.70 조화수열 $\{H_i\}$에 대해 $\{H_i - \log i\}$는 수렴한다. 따라서 Euler 상수는 well-define된다. 한편, 조화수열 자체는 발산한다.

Proof 표기의 편의를 위해 수열 $\{a_i\}$를 $a_i = H_i - \log i$로 두면 각 $i \in \mathbb{N}$에 대해 $a_{i+1} - a_i = 1/(i+1) - \log(i+1) + \log i = 1/(i+1) + \log(1 - 1/(i+1))$인데, $(-\infty, 1)$에서 정의된 함수 $x \mapsto x + \log(1-x)$가 양이 아님을 쉽게 알 수 있으므로 곧 $\{a_i\}$는 단조감소한다. 한편, 각 $i \in \mathbb{N}$에 대해 $\log(i+1) = \int_1^{i+1} 1/x\,dx \le H_i$이므로 $a_i \ge \log(i+1) - \log i = \log(1+1/i) \ge 0$이 되어 $\{a_i\}$는 0에 의해 아래로 유계이다. 그렇다면 MSP로부터 $\{a_i\}$가 수렴하고, 곧 증명이 끝난다.

한편, 조화수열이 발산함은 잘 알려져 있는데, 간단한 증명을 하나 소개하도록 하도록 하겠다. 임의의 $M \in \mathbb{R}$에 대해 $M < k/2$를 만족하는 충분히 큰 $k \in \mathbb{N}$를 택하면

$$
\begin{aligned}
H_{2^k} &= \sum_{i=1}^{2^k} \frac{1}{i} \\
&= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \cdots + \left(\frac{1}{2^{k-1}+1} + \cdots + \frac{1}{2^k}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\geq 1+\frac{1}{2}+\left(\frac{1}{4}+\frac{1}{4}\right)+\cdots+\left(\underbrace{\frac{1}{2^k}+\cdots+\frac{1}{2^k}}_{2^{k-1}\text{개}}\right)\\
&=1+\frac{k}{2}\\
&>M
\end{aligned}
$$

이고 $\{H_i\}$가 증가함은 자명하므로 $H_i \to \infty$임을 안다. □

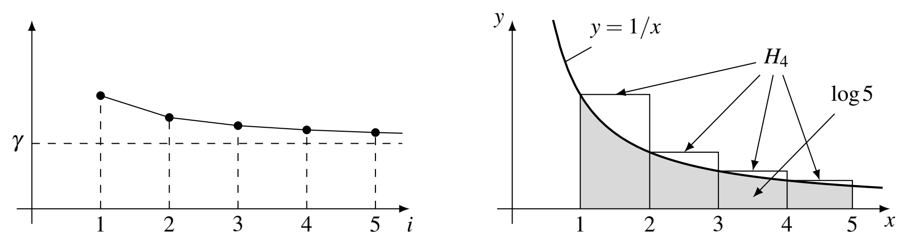


Figure 0.13 명제 0. 70에서의 수열 $\{H_i - \log i\}$의 그래프(왼쪽)와 이의 증명에서의 부등식 $\log(i+1) \leq H_i$의 $i = 4$인 경우(오른쪽).

Euler 상수에 대해서는 아직까지도 많은 부분이 베일에 싸여있다. 그 값은 대략 0.57721 정도로 알려져 있는데, 1734년 즈음에 처음 등장했지만 아직 이가 초월수인지 대수적 수인지 심지어 유리수인지 무리수인지도 밝혀지지 않았다. 이런 미스터리한 면모에도 불구하고 Euler 상수는 수학과 공학의 여러 분야에서 자주 등장하는 중요한 상수 중 하나이다.

Lemma 0.71 임의의 $p \geq 0$에 대해 $\lim_{x\to\infty} \Gamma(x+p)/x^p\Gamma(x) = 1$이다.

PROOF Stirling의 공식으로부터

$$
\begin{aligned}
\lim_{x\to\infty}\frac{\Gamma(x+p)}{x^p\Gamma(x)} &= \lim_{x\to\infty}\sqrt{\frac{2\pi}{x+p}}\left(\frac{x+p}{e}\right)^{x+p}\Bigg/ x^p\sqrt{\frac{2\pi}{x}}\left(\frac{x}{e}\right)^{x}\\
&= e^{-p}\lim_{x\to\infty}\left(1+\frac{p}{x}\right)^{x+p-1/2}\\
&= 1
\end{aligned}
$$

이므로 보조정리는 자명하다. □

Theorem 0.72 (Weierstrass) 임의의 $x > 0$에 대해서 다음의 성립한다.

$$
\Gamma(x) = \frac{e^{-\gamma x}}{x}\prod_{i=1}^{\infty}\left(1+\frac{x}{i}\right)^{-1}e^{x/i}
$$

Proof 임의의 $x>0$를 고정하고 수열 $\{H_i\}$를 $H_i=\sum_{j=1}^{i}1/j$라 하자. 이제 각 $i\in\mathbb{N}$에 대해 수열 $\{P_i\}$를 $P_i:=x\prod_{j=1}^{i-1}(1+x/j)e^{-x/j}=x^{\bar{i}}e^{(i^{-1}-H_i)x}/\Gamma(i)$로 두면 $\Gamma(x)=\Gamma(x+i)/x^{\bar{i}}=\Gamma(x+i)/P_i\Gamma(i)e^{(H_i-i^{-1})x}=[\Gamma(x+i)/i^x\Gamma(i)]e^{x/i}/P_ie^{(H_i-\log i)x}$인데 위의 보조정리와 Euler 상수의 정의로부터 $i\to\infty$이면 $\Gamma(x+i)/i^x\Gamma(i)\to 1$이고 $H_i-\log i\to\gamma$이므로 $P_i\to e^{-\gamma x}/\Gamma(x)$가 되고, 증명은 이로써 충분하다. □

이 결과를 두고 흔히 감마함수의 **Weierstrass의 정의** 혹은 **무한곱 표현 (infinite product representation)**이라 하고, 이 책에서 우리가 사용하고 있는 감마함수의 정의를 감마함수의 **적분 표현 (integral representation)**이라 한다. 그리고 이런 류의 결과가 하나 더 있다.

Theorem 0.73 (Euler) 임의의 $x>0$에 대해서 다음의 성립한다.

$$\Gamma(x)=\lim_{k\to\infty}\frac{k!k^x}{x^{\overline{k+1}}}$$

Proof 감마함수의 Weierstrass의 정의로부터

$$\begin{aligned}\Gamma(x)&=\frac{e^{-\gamma x}}{x}\prod_{i=1}^{\infty}\left(1+\frac{x}{i}\right)^{-1}e^{x/i}\\&=\frac{1}{x}\lim_{k\to\infty}\exp\left(\left(\log k-\sum_{i=1}^{k}\frac{1}{i}\right)x\right)\prod_{i=1}^{k}\frac{i}{x+i}e^{x/i}\\&=\lim_{k\to\infty}\frac{k!e^{x\log k}}{x^{\overline{k+1}}}\\&=\lim_{k\to\infty}\frac{k!k^x}{x^{\overline{k+1}}}\end{aligned}$$

이므로 정리는 자명하다. □

이는 감마함수의 **Euler의 정의** 혹은 **극한식 표현 (limit representation)**이라 한다. 지금까지 소개된 감마함수의 세 가지 '정의'는 주의깊게 보지 않으면 같다는 것을 알기 힘들 정도로 서로 다른 표현으로 이루어져 있지만 완전히 동등하기에 필요한 상황에 따라 맞는 것을 골라 쓸 수 있어 편리하다. 우리의 클라이맥스를 장식할 다음 정리를 증명하는 과정에서는 Euler의 정의를 사용한다.

Theorem 0.74 (Bohr-Mollerup) 다음 조건을 만족하는 양함수 $f:\mathbb{R}^+\to\mathbb{R}^+$는 감마함수로 유일하다.

i. $f(1)=1$.
ii. 임의의 $x>0$에 대해 $f(x+1)=xf(x)$이다.
iii. 함수 f는 log−볼록하다.

Proof 우선 감마함수가 주어진 세 조건을 만족함은 정리 0.64의 i, iii, 0.65로부터 자명하다. 이제 양함수 $f:\mathbb{R}^+\to\mathbb{R}^+$가 주어진 조건을 만족한다고 하면 임의의 $x>1$에 대해 적당한 $n\in\mathbb{N}$이 존재하여 $f(x)=(x-1)\cdot\cdots\cdot(x-n)f(x-n)$이고 $0<x-n\le 1$이므로 $(0,1]$에서 $f=\Gamma$임을 보이는 것으로 증명은 충분하다. 이를 위해 서로다른 임의의 $x,y\in(0,1]$에 대해 $S(x,y)$를 두 점 $(x,\log f(x))$와 $(y,\log f(y))$를 잇는 선분의 기울기라 하면 함수 $S:\{(x,y)\in(0,1]^2 : x\neq y\}\to\mathbb{R}$는 각 변수에 대해 증가한다. 이로부터 임의의 $x\in(0,1]$와 충분히 큰 임의의 $k\in\mathbb{N}$를 고정하면

$$\begin{aligned}
\log(k-1) &= \log(k-1)!-\log(k-2)!\\
&= \log f(k)-\log f(k-1)\\
&= S(k-1,k)\\
&\le S(k,x+k)\\
&= \frac{\log f(x+k)-\log f(k)}{x}\\
&= \frac{\log x^{\overline{k}}f(x)-\log(k-1)!}{x}
\end{aligned}$$

에서 $\log(k-1)!(k-1)^x=\log(k-1)!+x\log(k-1)\le\log x^{\overline{k}}f(x)$인데, 로그함수는 증가함수이므로 이는 곧 $(k-1)!(k-1)^x\le x^{\overline{k}}f(x)$를 뜻하여 $(k-1)!(k-1)^x/x^{\overline{k}}\le f(x)$임을 안다. 한편, $S(k,x+k)\le S(k,k+1)$임을 이용하여 비슷하게 하면 $f(x)\le k^x k!(x+k)/x^{\overline{k+1}}k$도 성립함을 알고, 이상을 종합하면

$$\frac{(k-1)!(k-1)^x}{x^{\overline{k}}}\le f(x)\le\frac{k!k^x}{x^{\overline{k+1}}}\cdot\frac{x+k}{k}$$

이다. 그렇다면 조임정리와 감마함수의 Euler의 정의로부터 $f(x)=\Gamma(x)$이고, x가 $(0,1]$의 임의의 원소였음을 상기한다면 $(0,1]$에서 $f=\Gamma$가 되어 증명이 끝난다. □

따라서 감마함수는 단순히 계승의 적당히 좋은 확장 정도가 아니라 사실상 유일한 확장이며, 전술했다시피 정리 0.65의 결과는 감마함수의 본질이 된다. 이로써 우리는 감마함수의 클라이맥스를 뒤로 하고 다음 특수함수로 넘어간다. 그 전에, 감마함수의 불완전한 형태가 있는데, 이 책에서는 크게 중요하지 않으니 이름 정도만 익혀두도록 하자.

Definition 0.75 다음과 같이 정의되는 함수 $\Gamma,\gamma:\mathbb{R}^+\times\mathbb{R}_0^+\to\mathbb{R}$를 각각 **상부 불완전 감마함수 (upper incomplete gamma function)**, **하부 불완전 감마함수 (lower incomplete gamma function)**라 한다.

$$\Gamma:(x,y)\mapsto\int_y^\infty t^{x-1}e^{-t}\,dt,\qquad \gamma:(x,y)\mapsto\int_0^y t^{x-1}e^{-t}\,dt$$

나아가, 다음과 같이 정의되는 함수 $Q, P : \mathbb{R}^+ \times \mathbb{R}_0^+ \to \mathbb{R}$를 각각 **정규화된 상부 불완전 감마함수(regularized upper incomplete gamma function)**, **정규화된 하부 불완전 감마함수(regularized lower incomplete gamma function)**라 한다.

$$Q : (x, y) \mapsto \frac{\Gamma(x, y)}{\Gamma(x)}, \qquad P : (x, y) \mapsto \frac{\gamma(x, y)}{\Gamma(x)}$$

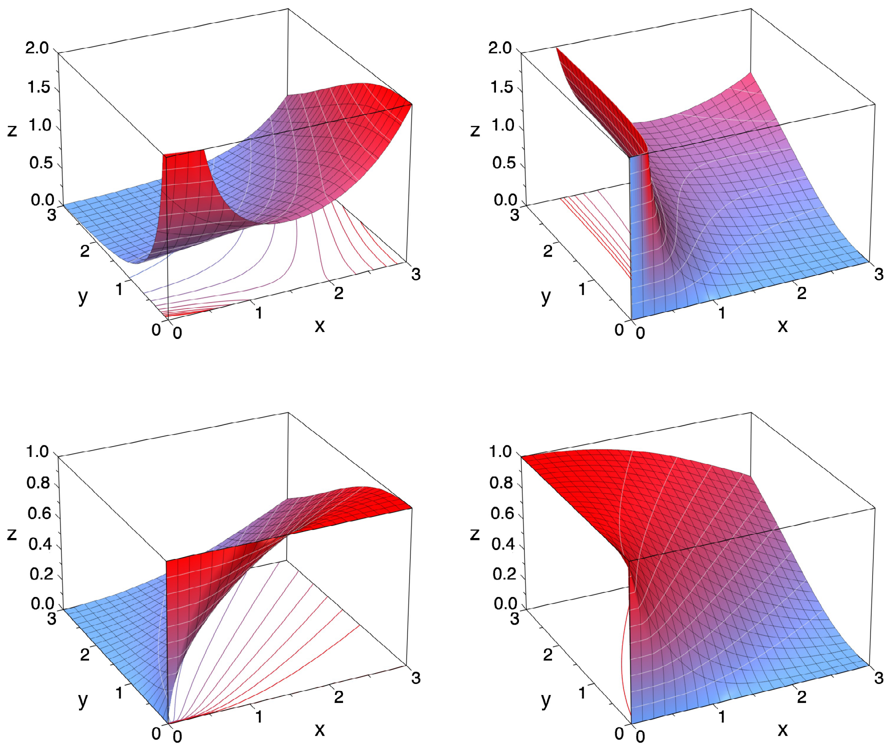


Figure 0.14 상부 불완전 감마함수(왼쪽 위), 하부 불완전 감마함수(오른쪽 위), 정규화된 상부 불완전 감마함수(왼쪽 아래), 정규화된 하부 불완전 감마함수(오른쪽 아래)의 그래프. Rendered by MATLAB Symbolic Math Toolbox MuPAD.

이제부터 우리가 살펴볼 특수함수는 감마함수에서 파생된 특수함수들로, 다음 특수함수는 $\log \Gamma$의 도함수로 얻어지는 함수이다. 감마함수는 $\mathscr{C}^\infty$급 양함수이므로 well-definedness는 자명할 것이다.

Definition 0.76 자연수 $k \in \mathbb{N}_0$에 대해 함수 $\log\Gamma : \mathbb{R}^+ \to \mathbb{R}$의 $k+1$차 도함수를 k**차 polygamma 함수** (**kth − function**)라 하고 $\psi_k : \mathbb{R}^+ \to \mathbb{R}$로 쓴다. 특별히, ψ_0을 **digamma 함수** (**− function**) 혹은 **프사이함수** (**psi function**)라 하고 간단히 $\psi : \mathbb{R}^+ \to \mathbb{R}$로 쓰며 ψ_1은 **trigamma 함수** (**− function**)라 하기도 한다.

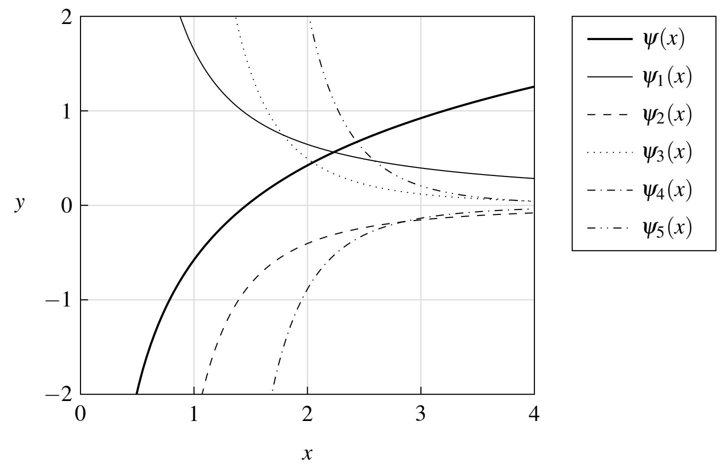


Figure 0.15 polygamma 함수의 그래프. Computed by Wolfram MATHEMATICA.

위의 정의에서 ψ가 ψ_1의 약식 표현이 아닌 ψ_0의 약식 표현이라는 점에 주의해야 한다. 한편, 감마함수의 Weierstrass의 정의를 사용하면 polygamma 함수를 무한급수의 형태로 전개할 수 있다.

Theorem 0.77 임의의 $k \in \mathbb{N}_0$과 임의의 $x > 0$에 대해

$$\psi_k(x) = \begin{cases} -\gamma - \dfrac{1}{x} + \displaystyle\sum_{i=1}^{\infty} \frac{x}{i(x+i)} & k = 0\text{인 경우} \\ \displaystyle\sum_{i=0}^{\infty} \frac{(-1)^{k+1}k!}{(x+i)^{k+1}} & \text{ow.} \end{cases}$$

이다.

Proof 임의의 $x_0 > 0$를 고정하자. 먼저 $k = 0$인 경우를 증명하기 위해 감마함수의 Weierstrass의 정의를 생각하면 임의의 $x > 0$에 대해 $\log\Gamma(x) = -\gamma x - \log x + \sum_{i=1}^{\infty}[x/i - \log(1+x/i)]$인데, 여기서 급수의 각 항이 미분가능하고 각 항의 도함수로 이루어진 급수 $\sum_{i=1}^{\infty}[1/i - x/(x+i)] = \sum_{i=1}^{\infty} x/i(x+i)$가 x_0의 근방에서 균등수렴하므로[3]위 식의 양변을 x_0에서 미분하면 $\psi(x_0) = -\gamma - 1/x_0 + \sum_{i=1}^{\infty} x_0/i(x_0+i)$를 얻는다.

다음으로, $k>0$인 경우를 보이기 위해 수학적 귀납법을 이용하자. 우선 $k=1$인 경우에는 급수 $\sum_{i=1}^{\infty} x/i(x+i)$의 각 항이 미분가능하고 각 항의 도함수로 이루어진 급수 $\sum_{i=1}^{\infty} i/(x+i)^2$이 x_0의 근방에서 균등수렴하므로[4] $k=0$인 경우의 결과를 x_0에서 미분하면 $\psi_1(x_0)=\sum_{i=0}^{\infty} i/(x_0+i)^2$을 얻어 증명이 끝난다. 이어서, 귀납가정으로서 $k\in\mathbb{N}$에 대해 보조정리가 성립한다고 하면 이번에도 급수 $\sum_{i=0}^{\infty}(-1)^{k+1}k!/(x+i)^{k+1}$의 각 항이 미분가능하고 각 항의 도함수로 이루어진 급수 $\sum_{i=1}^{\infty}(-1)^{k+2}(k+1)!/(x+i)^{k+2}$가 x_0의 근방에서 균등수렴하므로(각주 4 참조.) $\psi_{k+1}(x_0)=\sum_{i=1}^{\infty}(-1)^{k+2}(k+1)!/(x_0+i)^{k+2}$가 되어 $k+1$에 대해서도 보조정리가 성립함을 알고, 곧 증명이 끝난다. □

Theorem 0.78 임의의 $k\in\mathbb{N}_0$에 대해 다음이 성립한다.

i. 만약 $k>0$이면 $(-1)^{k+1}\psi_k>0$이다.
ii. 프사이함수는 $x_*\in(1,2)$에서 유일한 영점을 가진다.
iii. 함수 $(-1)^{k+1}\psi_k$는 순감소한다.
iv. 함수 $(-1)^{k+1}\psi_k$는 순볼록하다.
v. $\lim_{x\downarrow 0}(-1)^{k+1}\psi_k(x)=\infty$.
vi. 만약 $k>0$이면 $\lim_{x\to\infty}\psi_k(x)=0$이다.
vii. $\lim_{x\to\infty}\psi(x)=\infty$.
viii. k차 polygamma 함수는 $\mathscr{C}^\infty$급이며 $n\in\mathbb{N}_0$에 대해 $\psi_k^{(n)}=\psi_{k+n}$이다.

Proof i, iii, iv, viii. 이는 정리 0.77와 polygamma 함수의 정의로부터 자명하다.

ii. 감마함수가 적당한 $x_*\in(1,2)$에서 극솟값을 가지므로 $\psi(x_*)=(\log\Gamma)'(x_*)=\Gamma'(x_*)/\Gamma(x_*)=0$이다. 한편, iii으로부터 프사이함수가 순증가하므로 x_*는 프사이함수의 유일한 영점이다.

v. 만약 $k>0$이면 정리 0.77로부터

$$\begin{aligned}\lim_{x\downarrow 0}|\psi_k(x)|&=\lim_{x\downarrow 0}\sum_{i=0}^{\infty}\frac{k!}{(x+i)^{k+1}}\\&\geq\lim_{x\downarrow 0}\frac{1}{x^{k+1}}\\&=\infty\end{aligned}$$

이므로 i에서 정리가 자명하다. 한편, $k=0$인 경우에도 $0\leq\lim_{x\downarrow 0}\sum_{i=1}^{\infty}x/i(x+i)\leq\lim_{x\downarrow 0}x\sum_{i=1}^{\infty}1/i^2=\lim_{x\downarrow 0}x\pi^2/6=0$이므로 정리 0.77로부터 $\lim_{x\downarrow 0}\psi(x)=-\gamma-\lim_{x\downarrow 0}1/x+\lim_{x\downarrow 0}\sum_{i=1}^{\infty}x/i(x+i)=-\infty$이다.

vi. 가정으로부터 $\lim_{x\to\infty}\sum_{i=1}^{\infty}1/(x+i)^{k+1}\leq\lim_{x\to\infty}\sum_{i=1}^{\infty}1/i^{3/2}x^{k-1/2}$인데 $p-$급수 판정법에 따라 $\sum_{i=1}^{\infty}1/i^{3/2}$가 적당한 $L>0$로 수렴하므로 $0\leq\lim_{x\to\infty}\sum_{i=1}^{\infty}1/(x+i)^{k+1}\leq$

$\lim_{x\to\infty} L/x^{k-1/2} = 0$이고, 곧 정리 0.77로부터 $\lim_{x\to\infty}|\psi_k(x)| = \lim_{x\to\infty}\sum_{i=0}^{\infty} k!/(x+i)^{k+1} = 0$이다.

vii. 각 $j \in \mathbb{N}$에 대해 $\lim_{x\to\infty}\sum_{i=1}^{\infty} x/i(x+i) \geq \lim_{x\to\infty}\sum_{i=1}^{j} x/i(x+i) = \sum_{i=1}^{j} 1/i = H_j$이므로 $\lim_{x\to\infty}\sum_{i=1}^{\infty} x/i(x+i) \geq \lim_{j\to\infty} H_j = \infty$이고, 곧 정리 0.77로부터 $\lim_{x\to\infty}\psi(x) = -\gamma - \lim_{x\to\infty} 1/x + \lim_{x\to\infty}\sum_{i=1}^{\infty} x/i(x+i) = \infty$이다. □

polygamma 함수가 감마함수에서 파생되었으니 감마함수가 가지는 성질 $\Gamma(x+1) = x\Gamma(x)$를 어느 정도 물려받게 되는데, 다음 정리가 바로 이 성질에 관한 것이다.

Theorem 0.79 임의의 $x > 0$와 임의의 $k \in \mathbb{N}_0$에 대해 $\psi_k(x+1) = \psi_k(x) + (-1)^k k!/x^{k+1}$이다.

Proof 이는 감마함수의 성질로부터 거의 자명하다. 수학적 귀납법을 사용하자. 우선 $k = 0$인 경우에는 $\log\Gamma(x+1) = \log x\Gamma(x) = \log x + \log\Gamma(x)$가 임의의 $x > 0$에 대해 성립하므로 양변을 x에 대해 미분하면 $\psi(x+1) = \psi(x) + 1/x$가 되어 $k = 0$인 경우에 정리가 성립함을 안다. 이제 귀납가정으로서 $k \in \mathbb{N}_0$에 대해 정리가 성립한다고 가정하고 양변을 x에 대해 한 번 더 미분하면 $\psi_{k+1}(x+1) = \psi_{k+1}(x) + (-1)^{k+1}(k+1)!/x^{k+2}$가 되어 $k+1$에 대해서도 정리가 성립함을 알고, 곧 증명이 끝난다. □

나중에 사용할 명제를 하나 증명하는 것을 끝으로 polygamma 함수에 대한 논의를 끝마치고, 다음 특수함수로 넘어간다.

Proposition 0.80 임의의 $x > 0$에 대해 다음이 성립한다.

i. $\psi(x) = \Gamma'(x)/\Gamma(x)$.
ii. $\psi_1(x) = \Gamma''(x)/\Gamma(x) - [\psi(x)]^2$.

Proof 이는 정의로부터 거의 자명하다. i의 경우 $\psi(x) = (\log\Gamma)'(x) = \Gamma'(x)/\Gamma(x)$이고, ii의 경우 i로부터 $\psi_1(x) = \psi'(x) = \{\Gamma''(x)\Gamma(x) - [\Gamma'(x)]^2\}/[\Gamma(x)]^2 = \Gamma''(x)/\Gamma(x) - [\psi(x)]^2$이다. □

감마함수와 관련해서 마지막으로 알아볼 특수함수는 베타함수로, 감마함수처럼 적분으로 정의된다. 그리고 이번에도 피적분함수가 $(0, 1)$에서 적분가능한지를 보여 이가 well-define 됨을 보장해 주어야 한다.

Definition 0.81 다음과 같이 정의되는 함수 $\mathrm{B} : (\mathbb{R}^+)^2 \to \mathbb{R}$를 **베타함수 (beta function)**라 한다.

$$\mathrm{B} : (x, y) \mapsto \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$$

Proposition 0.82 임의의 $\alpha, \beta \in \mathbb{R}$에 대해 $I(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx$라 하면 $\alpha, \beta > 0$일때 $I(\alpha, \beta) < \infty$이고 그렇지 않으면 $I(\alpha, \beta) = \infty$이다.

Proof 먼저 $\alpha \leq 0$인 경우를 생각하면 충분히 큰 $n \in \mathbb{N}$에 대해

$$\begin{aligned} I(\alpha, \beta) &\geq \int_0^1 \frac{(1-x)^\beta}{x}\,dx \\ &\geq \int_0^1 \frac{(1-x)^n}{x}\,dx \\ &= \int_0^1 \sum_{k=0}^n \binom{n}{k}(-1)^k x^{k-1}\,dx \\ &= \sum_{k=1}^n \binom{n}{k}(-1)^k \int_0^1 x^{k-1}\,dx + \int_0^1 \frac{1}{x}\,dx \\ &= \sum_{k=1}^n \binom{n}{k}\frac{(-1)^k}{k} + [\log x]_0^1 \\ &= \infty \end{aligned}$$

이고, $\beta \leq 0$인 경우에도 비슷하게 같은 결론을 얻는다. 한편, $\alpha, \beta > 0$이면 간단한 치환적분으로부터

$$\begin{aligned} \Gamma(\alpha)\Gamma(\beta) &= \int_0^\infty t^{\alpha-1}(1-t)^{\beta-1}\,dx \int_0^\infty s^{\alpha-1}e^{-s}\,ds \\ &= \int_{(\mathbb{R}^+)^2} t^{\alpha-1}s^{\beta-1}e^{-(t+s)}\,dt\,ds \\ &= \int_0^1\int_0^\infty (uv)^{\alpha-1}[u(1-v)]^{\beta-1}e^{-u}u\,du\,dv \qquad \because (t,s) = (uv, u(1-v)) \\ &= \left[\int_0^\infty u^{\alpha+\beta-1}e^{-u}\,du\right]\left[\int_0^1 v^{\alpha-1}(1-v)^{\beta-1}\,dv\right] \\ &= \Gamma(\alpha+\beta)I(\alpha, \beta) \end{aligned}$$

가 성립하고, 곧 $I(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta) < \infty$이다. □

앞서 살펴본 polygamma 함수와는 달리 베타함수는 그 정의만 봐서는 감마함수와 어떤 관련이 있는지 잘 보이지 않는데, 다음 따름정리로부터 그 관련성이 확실해진다.

Corollary 0.83 임의의 $x, y > 0$에 대해 $\mathrm{B}(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$이다.

Proof 증명은 명제 0.82의 증명으로부터 자명하다. □

Corollary 0.84 임의의 $x, y > 0$에 대해 $\mathrm{B}(x, y) = \mathrm{B}(y, x)$이다.

Proof 이 또한 보조정리 0.83로부터 자명하다. □

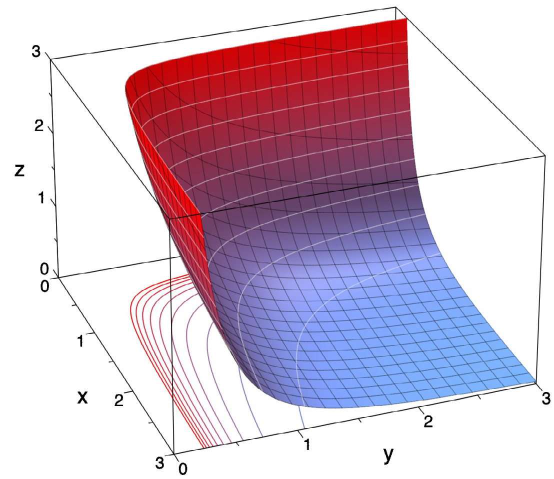


Figure 0.16 베타함수의 그래프. Computed and rendered by MATLAB Symbolic Math Toolbox MuPAD.

여기서 감마함수가 계승의 일반화였다는 점을 떠올린다면 따름정리 0.83에서의 $\Gamma(x)\Gamma(y)/\Gamma(x+y)$라는 식에서 이항계수의 정의가 언뜻 보일 것이다. 실제로 $m \geq n$인 $m, n \in \mathbb{N}_0$에 대해

$$\begin{aligned}\binom{m}{n} &= \frac{m!}{n!(m-n)!} \\ &= \frac{\Gamma(m+1)}{\Gamma(n+1)\Gamma(m-n+1)} \\ &= \frac{\Gamma(m+2)}{(m+1)\Gamma(n+1)\Gamma(m-n+1)} \\ &= \frac{1}{(m+1)\mathrm{B}(n, m-n)}\end{aligned}$$

이 되어 베타함수는 이항계수의 일반화로 생각할 수 있다. 이런 맥락에서 따름정리 0.84의 결과는 이항계수의 성질

$$\binom{m}{n} = \binom{m}{m-n}$$

에 대응하는 것으로 파악할 수 있겠다.

다음 순서는 베타함수의 기본적인 성질들을 알아보는 것인데, 이번에도 Hölder의 부등식과 Leibniz의 법칙이 사용되니 잠시 1장으로 넘어갔다 오는 것을 추천한다.

Theorem 0.85 다음이 성립한다.

i. 베타함수는 양함수이다.
ii. 베타함수는 각 변수에 대해 순감소한다.
iii. 베타함수는 log−순볼록하다.
iv. $\lim_{(x,y)\downarrow 0} \mathrm{B}(x, y) = \lim_{x\downarrow 0} \mathrm{B}(x, y) = \lim_{y\downarrow 0} \mathrm{B}(x, y) = \infty$.
v. $\lim_{(x,y)\to\infty} \mathrm{B}(x, y) = \lim_{x\to\infty} \mathrm{B}(x, y) = \lim_{y\to\infty} \mathrm{B}(x, y) = 0$.
vi. 베타함수는 $\mathscr{C}^\infty$ 급이며 $m, n \in \mathbb{N}_0$ 에 대해

$$\frac{\partial^{m+n}}{\partial x^m \partial y^n}\mathrm{B}(x, y) = \int_0^1 (\log t)^m t^{x-1}[\log(1-t)]^n (1-t)^{y-1}\, dt$$

이다. 단, 이때 편미분의 순서는 중요하지 않다.

Proof i. 이는 따름정리 0. 83로부터 자명하다.

ii. 프사이함수가 순증가하므로 임의의 $x, y > 0$에 대해

$$\begin{aligned}
\frac{\partial}{\partial x}\mathrm{B}(x, y) &= \frac{\partial}{\partial x}\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \\
&= \frac{\Gamma'(x)\Gamma(x+y) - \Gamma(x)\Gamma'(x+y)}{[\Gamma(x+y)]^2}\Gamma(y) \\
&= \frac{\Gamma'(x)\Gamma(y)}{\Gamma(x+y)} - \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \cdot \frac{\Gamma'(x+y)}{\Gamma(x+y)} \\
&= \mathrm{B}(x, y)\psi(x) - \mathrm{B}(x, y)\psi(x+y) \\
&= [\psi(x) - \psi(x+y)]\mathrm{B}(x, y) \\
&< 0
\end{aligned}$$

이고, 따라서 베타함수는 첫 번째 변수에 대해 순감소한다. 이제 두 번째 변수에 대해서도 비슷하게 하면 같은 결론을 얻는다.

iii. 서로다른 임의의 $(x, y), (z, w) \in (\mathbb{R}^+)^2$ 와 임의의 $t \in (0, 1)$ 에 대해 Hölder의 부등식으로부터

$$\begin{aligned}
\mathrm{B}(t(x, y) + (1-t)(z, w)) &= \int_0^1 s^{tx+(1-t)z-1}(1-s)^{ty+(1-t)w-1}\, ds \\
&= \int_0^1 [s^{x-1}(1-s)^{y-1}]^t [s^{z-1}(1-s)^{w-1}]^{1-t}\, ds \\
&\le \left[\int_0^1 s^{x-1}(1-s)^{y-1}\, ds\right]^t \left[\int_0^1 s^{z-1}(1-s)^{w-1}\, ds\right]^{1-t}
\end{aligned}$$

$$= [\mathrm{B}(x,y)]^t[\mathrm{B}(z,w)]^{1-t}$$

이다. 만약 여기서 등호가 성립한다면 곧 모두 0은 아닌 적당한 $a, b \geq 0$가 존재하여 거의 대부분의 $s \in (0,1)$에 대해 $as^{x-1}(1-s)^{y-1} + bs^{z-1}(1-s)^{w-1} = [as^{x-z}(1-s)^{y-w} + b]s^{z-1}(1-s)^{w-1} = 0$인데, 이는 불가능하므로 위에서 등호는 성립하지 않는다. 따라서 베타함수는 log−순볼록하다.

iv. 우선 감마함수의 성질로부터

$$\begin{aligned}
\lim_{(x,y)\downarrow 0} \mathrm{B}(x,y) &= \lim_{(x,y)\downarrow 0} \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \\
&= \lim_{(x,y)\downarrow 0} \frac{[\Gamma(x+1)/x][\Gamma(y+1)/y]}{\Gamma(x+y+1)/(x+y)} \\
&= \lim_{(x,y)\downarrow 0} \frac{\Gamma(x+1)\Gamma(y+1)}{\Gamma(x+y+1)} \cdot \frac{x+y+1}{xy} \\
&= \infty
\end{aligned}$$

임을 안다. 한편, 고정된 $y > 0$에 대해서는 $\lim_{x\downarrow 0} \mathrm{B}(x,y) = \lim_{x\downarrow 0} \Gamma(x)\Gamma(y)/\Gamma(x+y) = \infty$임이 분명하고, 고정된 $x > 0$에 대해서도 비슷하게 원하는 결과를 얻을 수 있다.

v. 우선 Stirling의 공식으로부터

$$\begin{aligned}
\lim_{x\to\infty} \mathrm{B}(x,x) &= \lim_{x\to\infty} \frac{[\Gamma(x)]^2}{\Gamma(2x)} \\
&= \lim_{x\to\infty} \left[\sqrt{\frac{2\pi}{x}}\left(\frac{x}{e}\right)^x\right]^2 \Big/ \sqrt{\frac{\pi}{x}}\left(\frac{2x}{e}\right)^{2x} \\
&= \lim_{x\to\infty} \frac{1}{\sqrt{x}2^{2x-1/2}} \\
&= 0
\end{aligned}$$

이므로 임의의 $\varepsilon > 0$에 대해 적당한 $M > 0$이 존재하여 $\mathrm{B}(M,M) < \varepsilon$이다. 따라서 ii로부터 임의의 $x, y > 0$에 대해 $x, y \geq M$이면 $\mathrm{B}(x,y) \leq \mathrm{B}(M,M) < \varepsilon$이 되어 $\lim_{(x,y)\to\infty} \mathrm{B}(x,y) = 0$임을 안다. 한편, 고정된 $y > 0$에 대해 이번에도 Stirling의 공식으로부터

$$\begin{aligned}
\lim_{x\to\infty} \mathrm{B}(x,y) &= \lim_{x\to\infty} \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \\
&= \lim_{x\to\infty} \left[\sqrt{\frac{2\pi}{x}}\left(\frac{x}{e}\right)^x\right]\left[\sqrt{\frac{2\pi}{y}}\left(\frac{y}{e}\right)^y\right] \Big/ \sqrt{\frac{2\pi}{x+y}}\left(\frac{x+y}{e}\right)^{x+y} \\
&= \lim_{x\to\infty} \frac{x^{x-1/2}y^{y-1/2}}{(x+y)^{x+y-1/2}}
\end{aligned}$$

$$
\begin{aligned}
&= y^{y-1/2}\lim_{x\to\infty}\left(1+\frac{y}{x}\right)^{-x+1/2}\frac{1}{(x+y)^y}\\
&= 0
\end{aligned}
$$

이다. 여기서 마지막 등호는 $x\to\infty$이면 $(1+y/x)^{-x+1/2}\to e^{-y}$이고 $1/(x+y)^y\to 0$이므로 성립한다. 이제 고정된 $x>0$에 대해서도 비슷하게 원하는 결과를 얻을 수 있다.

vi. 우선 따름정리 0.83로부터 베타함수가 $\mathscr{C}^\infty$급임이 자명하므로 편미분의 순서가 문제되지 않는다. 본격적인 증명은 m, n에 대한 수학적 귀납법을 중첩하여 사용한다. 먼저 $m=0$으로 고정하고 n에 대한 수학적 귀납법을 사용하자. 그런데 $n=0$인 경우에는 정리가 자명하므로 귀납귀정으로서 $n\in\mathbb{N}_0$에 대해 정리가 성립한다고 하여 임의의 $x_0, y_0>0$를 고정하고 $y_0\in(a,b)=:I$인 적당한 $a,b>0$를 택하여 함수 $f:(0,1)\times I\to\mathbb{R}$를 $f:(t,y)\mapsto t^{x_0-1}[\log(1-t)]^n(1-t)^{y-1}$로 정의하자. 그렇다면 f는 Leibniz의 법칙의 조건 i, ii, iii을 만족함이 분명하고 $\mathbb{R}^+$에서 정의된 함수 $t\mapsto t^{x_0-1}$이 $[1-1/e,1]$에서 최댓값 $M>0$을 가지므로 임의의 $(t,y)\in(0,1)\times I$에 대해

$$
\begin{aligned}
\left|\frac{\partial}{\partial y}f(t,y)\right| &= t^{x_0-1}|\log(1-t)|^{n+1}(1-t)^{y-1}\\
&\le t^{x_0-1}(1-t)^{y-1}\mathbf{1}_{(0,1-1/e]}(t)+M|\log(1-t)|^{n+1}(1-t)^{y-1}\mathbf{1}_{(1-1/e,1)}(t)\\
&\le t^{x_0-1}(1-t)^{a-1}\mathbf{1}_{(0,1-1/e]}(t)+M|\log(1-t)|^{n+1}(1-t)^{a-1}\mathbf{1}_{(1-1/e,1)}(t)\\
&\le t^{x_0-1}(1-t)^{a-1}+M|\log(1-t)|^{n+1}(1-t)^{a-1}
\end{aligned}
$$

인데, 여기서 $\int_0^1 t^{x_0-1}(1-t)^{a-1}\,dt=\mathrm{B}(x_0,a)$이고 정리 0.64의 v의 증명에서와 같이 하면 $(0,1)$에서 정의된 함수 $t\mapsto M|\log(1-t)|^{n+1}(1-t)^{a-1}$가 적분가능함을 쉽게 알 수 있으므로 f가 Leibniz의 법칙의 조건 iv도 만족시킴을 안다. 이로부터 I에서 정의된 함수 $y\mapsto\int_0^1 t^{x_0-1}[\log(1-t)]^n(1-t)^{y-1}\,dt$는 y_0에서 미분가능하며 그 미분계수는 $\int_0^1 t^{x_0-1}[\log(1-t)]^{n+1}(1-t)^{y-1}\,dt$로 주어지고, 곧 $m=0$이면 모든 $n\in\mathbb{N}_0$에 대해 정리가 성립한다.

이제 다시 귀납가정으로서 $m\in\mathbb{N}_0$과 임의의 $n\in\mathbb{N}_0$에 대해 정리가 성립한다고 가정하면 $m+1$과 임의의 $n\in\mathbb{N}_0$에 대해서도 정리가 성립함을 방금과 비슷하게 보일 수 있으므로 증명이 끝난다. □

베타함수의 미분은 감마함수와 polygamma 함수를 이용하여 다음과 같이 쓸 수도 있다.

Proposition 0.86 임의의 $x, y>0$에 대해 다음이 성립한다.

$$
\begin{aligned}
&\nabla\mathrm{B}(x,y)=(\psi(x)-\psi(x+y),\ \psi(y)-\psi(x+y))\mathrm{B}(x,y)\\
&\mathbf{H}_{(x,y)}\mathrm{B}
\end{aligned}
$$

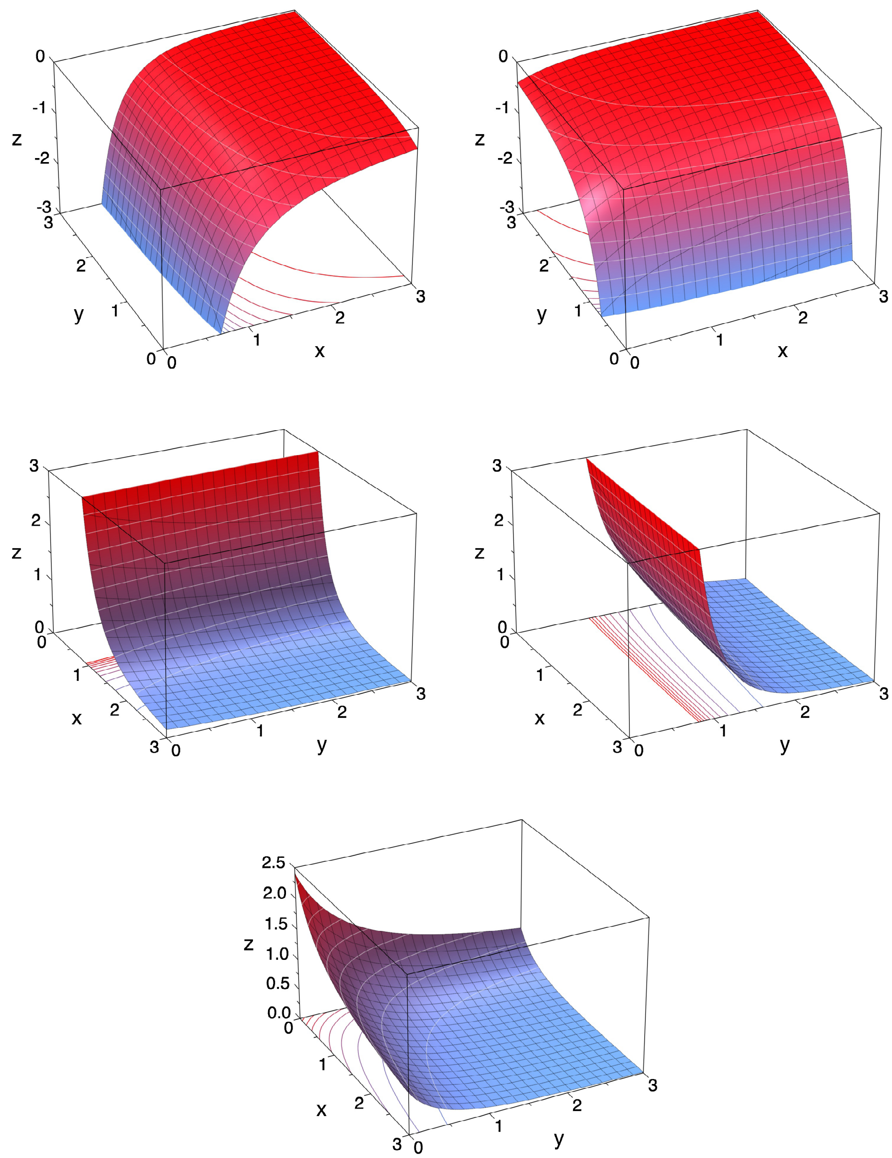


Figure 0.17 베타함수의 도함수들의 그래프. x에 대한 일계편도함수(왼쪽 위), y에 대한 일계편도함수(오늘쪽 위), x에 대한 이계편도함수(왼쪽 중간), y에 대한 이계편도함수(오른쪽 중간), x, y에 대한 이계편도함수(아래). Computed and rendered by MATLAB Symbolic Math Toolbox MuPAD.

$$= \begin{bmatrix} [\psi(x)-\psi(x+y)]^2+\psi_1(x)-\psi_1(x+y) & [\psi(x)-\psi(x+y)][\psi(y)-\psi(x+y)]-\psi_1(x+y) \\ \text{sym.} & [\psi(y)-\psi(x+y)]^2+\psi_1(y)-\psi_1(x+y) \end{bmatrix} \mathrm{B}(x,y)$$

Proof 이는 따름정리 0.83로부터 거의 자명하다. 우선 0.85의 ii의 증명에서와 같이 하면 $\nabla\mathrm{B}(x,y)$에 대한 표현을 얻고, 이를 한 번 더 편미분하면 $\mathbf{H}_{(x,y)}\mathrm{B}$에 대한 표현을 얻는다. □

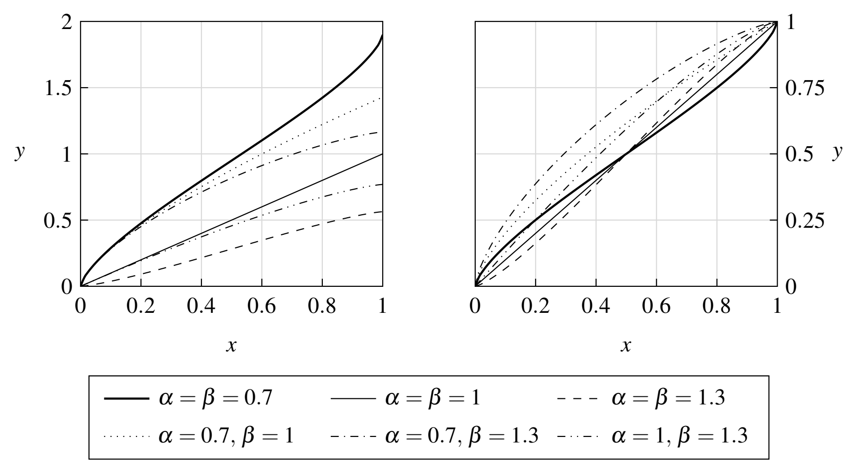
베타함수도 감마함수와 비슷하게 불완전한 형태가 있고, 이와 더불어 원래는 이변수함수인 베타함수를 $n \geq 3$변수함수로 확장한 것도 있는데 모두 이름 정도만 익혀두도록 하자. 다변수 베타함수는 나중에 Dirichlet 분포를 공부할 때 잠시 등장할 것이다.

Definition 0.87 임의의 $\alpha, \beta > 0$에 대해 다음과 같이 정의되는 함수 $\mathrm{B}(\cdot;\alpha,\beta):[0,1]\to\mathbb{R}$를 **불완전 베타함수 (incomplete beta function)**라 한다.

$$\mathrm{B}(\cdot;\alpha,\beta): x \mapsto \int_0^x t^{\alpha-1}(1-t)^{\beta-1}\,dt$$

나아가, 다음과 같이 정의되는 함수 $I_x(\alpha,\beta):[0,1]\to\mathbb{R}$를 **정규화된 불완전 베타함수 (regularized incomplete beta function)**라 한다.

$$I_x(\alpha,\beta) = \frac{\mathrm{B}(x;\alpha,\beta)}{\mathrm{B}(\alpha,\beta)}$$



Definition 0.88 1보다 큰 $n \in \mathbb{N}$에 대해 다음과 같이 정의되는 함수 $\mathrm{B}:(\mathbb{R}^+)^n\to\mathbb{R}$를 **다변수 베타함수 (multivariate beta function)**라 한다.

$$\mathrm{B}: x \mapsto \frac{\prod_{i=1}^n \Gamma(x_i)}{\Gamma(\sum_{i=1}^n x_i)}$$

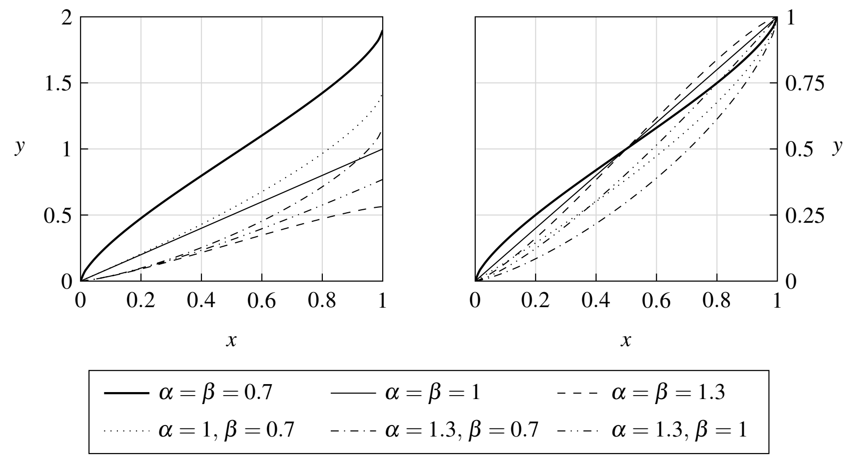


Figure 0.18 불완전 베타함수(위)와 정규화된 불완전 베타함수(아래)의 그래프. Computed by Wolfram MATHEMATICA.

0.2.3 Special Functions Related to Riemann Zeta Function

감마함수라는 큰 주제가 일단락되었으니 이번에는 조금 가벼운 주제를 다루어보자. 정확히 말하면 감마함수보다 훨씬 더 무거운 주제이지만 여기서는 가볍게 다루어보자. 이번에 다룰 특수함수는 Riemann 가설로 그 인기를 톡톡히 누리고 있는 제타함수이다.

Definition 0.89 다음과 같이 정의되는 함수 $\zeta : (1, \infty) \to \mathbb{R}$를 **(Riemann) 제타함수(– zeta function)**라 한다.

$$\zeta : x \mapsto \sum_{i=1}^{\infty} \frac{1}{i^x}$$

바로 이 함수가 Riemann 가설의 주인공으로, 여기서 Riemann 가설에 대해서도 간단히 다룰 수 있으면 좋겠지만, 이는 절 하나로 가볍게 훑고 지나갈 수 있는 정도의 무게가 아니다. **Riemann 가설(RH; – Hypothesis)**은 방금 정의한 제타함수의 $\mathbb{C} \setminus \{1\}$로의 해석적 확장의 비자명한 영점의 실수부가 모두 $1/2$라는 가설로, 1859년 Riemann이 그의 논문 *Über die Anzahl der Primzahlen unter einer gegenenen Größe*(영어: *On the Number of Primes Less Then a Given Magnitude*)에서 처음 제시하였다. 그는 고작 8쪽인 이 논문에서 실로 어마어마한 일들을 해냈는데 위에서 정의된 제타함수를 $\mathbb{C} \setminus \{1\}$에서도 잘 정의되도록 확장하고, 이의 자명한 영점과 비자명한 영점을 탐구하였으며 소수계량함수 $\pi : \mathbb{R} \to \mathbb{N}_0$를 제타함수의 비자명한 영점들의 합으로 표현하여 RH가 참이라는 전제 하에 $\lim_{x\to\infty} \pi(x) \log x / x = 1$이라는 소수정리를 증명했다. 많은 천재들이 그러하듯이 그는 이런

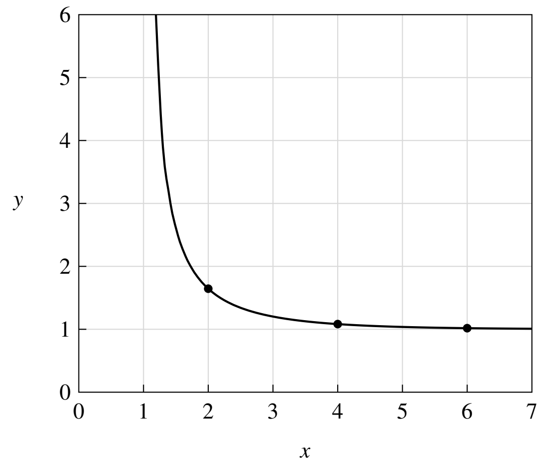


Figure 0.19 제타함수의 그래프. Computed by Wolfram MATHEMATICA.

엄청난 결과의 중간과정을 거의 다 생략했으며 특히 컴퓨터가 없던 당시 천문학적인 노가다를 통해서만 계산할 수 있었던 제타함수의 비자명한 영점을 몇 개 예시로 툭 던져두는 바람에 이 논문은 출간된 당시 굉장한 파장을 가져왔다.

하지만 Riemann이 이 논문을 쓸 당시에 그의 주된 관심사는 RH가 아니라 소수정리의 증명이었다. 논문의 제목도 결론도 모두 소수정리에 초점이 맞추어져 있었으며 RH는 논의의 과정에서 주요한 도구 중 하나로 조명될 뿐이었다. 실제로 그는 RH에 대한 증명은 별로 어렵지 않을 것이라 생각했고, 논문에서도 시간 낭비를 하지 않기 위해 당연해 보이는 사실에 대한 증명은 건너뛴다는 식으로 적어두어 처음에 RH는 별 대수롭지 않은 연습문제 정도로 인식되었다. (선대군의 '독자들이 심심해 하는 것 같다. ㅋㅋㅋ' 정도의 느낌이랄까.) 그러나 RH는 생각만큼 만만한 상대가 아니었고 정작 이를 통해 증명하려고 했던 소수정리는 1896년 Jacques Hadamard와 Vallée-Poussin에 의해 RH를 요령껏 우회하는 방식으로 증명되었다. 결국엔 배보다 배꼽이 더 컸던 격이다. 이렇게 RH는 세기의 난제로 남게 되었으며 미국의 Clay 수학연구소가 2000년에 이를 밀레니엄 난제 중 하나로 선정하며 그 인기가 폭발하여 오늘날에 이르게 되었다.

그러나 RH가 아직 증명되지 못했다고 해서 Riemann의 논문이 가치가 없는 것은 아니다. 오히려, 이렇게 오래도록 풀리지 않을 난제를 제시했다는 것 자체로 엄청난 가치를 지닌 논문으로 평가되며, 소수정리라는 정수론의 주제를 제타함수라는 해석학의 주제로 접근한 시도로도 높이 평가된다. 실제로 이 논문으로부터 해석적 정수론이라는 분야가 탄생하게 되었으며 Riemann의 공로를 기리는 의미에서 제타함수의 변수는 그가 논문에서 사용한 표기를 그대로 따라 s로 쓰는 것이 오랜 관례로 굳어져 있다. (이 책에서는 그냥 x로 썼다.) 사족으로, RH는 세기의 난제이지만 다른 난제와는 달리 일반인들도 약간의 설명으로 문제

자체는 '적당히' 이해할 수 있는 까닭에 매년 이 난제에 도전하는 사람도, 그들로부터 나오는 엉터리 증명도 넘쳐난다. 유명 수학 커뮤니티나 arXiv 같은 사이트에서는 지금 이 순간에도 RH를 증명했다고 주장하는 사람들과 이들을 반박하는 사람들로 키배가 끊이지 않는다. 심지어 이를 증명했다고 주장하는 사람들 중에는 진짜 수학자들도 있어서 정말 혼란 그 자체인데, 최근에는 가환대수학의 명저 *Introduction to Commutative Algebra*의 저자인 Michael Atiyah경이 RH를 증명했다고 주장하는 안타까운 해프닝이 있기도 했다.

아무튼, RH에 대해 여기서 말할 수 있는 건 이런 교양 수준의 지식 뿐이다. 일단 당장에 제타함수의 $\mathbb{C}\setminus\{1\}$로의 해석적 확장을 설명하려면 복소해석학이 필요하다. 그러니 이 정도로 만족하고 우리는 제타함수의 기본적인 성질을 살펴보자.

Theorem 0.90 다음이 성립한다.

i. $\zeta > 1$.
ii. 제타함수는 순감소한다.
iii. 제타함수는 순볼록하다.
iv. $\lim_{x\downarrow 1}\zeta(x) = \infty$.
v. $\lim_{x\to\infty}\zeta(x) = 1$.
vi. 제타함수는 $\mathscr{C}^\infty$급이며 $n\in\mathbb{N}_0$에 대해

$$\zeta^{(n)}(x) = \sum_{i=1}^{\infty}\frac{(-\log i)^n}{i^x}$$

이다.

Proof i, ii. 이는 제타함수의 정의로부터 자명하다.

vi. 임의의 $x_0 > 1$을 고정하고 수학적 귀납법을 사용하자. 우선 $n = 0$인 경우에는 정리가 자명하므로 귀납가정으로서 $n\in\mathbb{N}_0$에 대해 정리가 성립한다고 가정하면 급수 $\sum_{i=1}^{\infty}(-1)^n(\log i)^n/i^x$의 각 항이 미분가능하고 그 도함수로 이루어진 급수 $\sum_{i=1}^{\infty}(-\log i)^{n+1}/i^x$가 x_0의 근방에서 균등수렴하므로[5] $\zeta^{(n+1)}(x_0) = \sum_{i=1}^{\infty}(-\log i)^{n+1}/i^{x_0}$이 되어 증명이 끝난다.

iii. vi에서 $\zeta'' > 0$이므로 이는 자명하다.

iv. 각 $j\in\mathbb{N}$에 대해 $\lim_{x\downarrow 1}\zeta(x) \geq \lim_{x\downarrow 1}\sum_{i=1}^{j}1/i^x = \sum_{k=1}^{j}1/i = H_j$이므로 $\lim_{x\downarrow 1}\zeta(x) \geq \lim_{j\to\infty}H_j = \infty$이다.

v. 급수 $\sum_{i=1}^{\infty}1/i^x$가 $[2,\infty)$에서 균등수렴하므로[6] $\lim_{x\to\infty}\zeta(x) = \sum_{i=1}^{\infty}\lim_{x\to\infty}1/i^x = 1$이다. □

이대로 제타함수라는 주제를 끝내기에는 뭔가 아쉬우니 RH 대신 지금은 풀린 제타함수와 관련된 다른 난제를 하나 풀어보도록 하자. 이 문제는 스위스의 Basel 대학에 교수로 재직하던 Jakob Bernoulli와 Johann Bernoulli 형제에 의해 제시된 문제로 자연수의 거듭제곱의

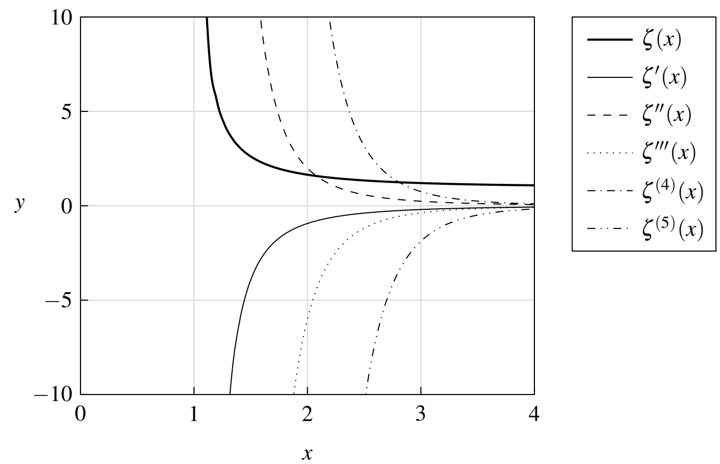


Figure 0.20 제타함수와 그 도함수들의 그래프. Computed by Wolfram MATHEMATICA.

역수의 합, 즉 $\zeta(2) = \sum_{i=1}^{\infty} 1/i^2$ 를 구하는 문제이다. p−급수 판정법에 의해 이 급수가 수렴함은 자명하므로 수렴값만 구하면 되는, RH에 비하면 간단한 문제였음에도 불구하고 120년간 내로라하는 수학자들의 도전에도 정복되지 않았기에 Basel 문제라는 이름으로 유럽 전역에 널리 알려지게 되었다. Basel 문제는 결국 Euler에 의해 1735년 $\pi^2/6 \approx 1.644934$로 수렴한다는 것이 증명되었는데, 이를 통해 Euler는 세계적인 수학자라는 타이틀을 획득하며 수학자로의 성공적인 데뷔무대를 치르게 된다. 이때 Euler가 사용한 핵심적인 아이디어는 $\sin x/x$을 '인수분해' 하는 것으로, 오늘날 여러 수학 교양도서에서 소개되고 있는 방법이 바로 이 방법이다. 그러나 이는 이론적으로 미묘한 문제가 있었기에 (오류는 아니다.) 6년이 지난 1741년, Euler가 수학적으로 엄밀한 다른 풀이를 제시하며 Basel 문제는 완벽하게 정복되었다. 한편, Euler가 1735년 제시한 첫 번째 풀이의 미묘한 부분은 시간이 한참 흐른 뒤 19세기에 Weierstrass의 곱 정리로 비로소 해결되게 된다. (잠시 후에 우리는 이 미묘한 부분도 초등적인 방법으로 해결할 것이다.)

이번 절에서는 여기서 한 술 더 떠서 자연수의 짝수 거듭제곱의 역수의 합, 즉 $\zeta(2n) = \sum_{i=1}^{\infty} 1/i^{2n}$ 의 값을, 복소해석학과 같은 현대수학의 문물을 사용하지 않고 초등적인 방법으로 구해보려고 한다. 초등적인 방법인 만큼 준비가 조금 필요하지만, 이는 제타함수라는 주제를 마무리하는 훌륭한 피날레가 될 것이다.

Definition 0.91 함수 $f : \mathbb{R} \to \mathbb{R}$를 다음과 같이 정의하자.

$$f : x \mapsto \begin{cases} x/(e^x - 1) & x \neq 0\text{인 경우} \\ 1 & \text{ow.} \end{cases}$$

이때, 이의 0에서의 Taylor 전개의 k차항의 계수를 k**번째 Bernoulli 수 (**k**th – number)** 라 하고 B_k로 쓴다. 즉, 0 근방의 $x \in \mathbb{R}$에 대해 $f(x) = \sum_{i=0}^{\infty} B_i x^i / i!$이며, 이때의 함수 f를 Bernoulli 수의 **생성함수 (generating function)**라 한다.

Bernoulli 수의 well-definedness를 위해서는 그 생성함수가 0에서 해석적임을 보여야 한다.

Proposition 0.92 Bernoulli 수의 생성함수 $f : \mathbb{R} \to \mathbb{R}$는 해석함수이다.

Proof 함수 $g : \mathbb{R} \to \mathbb{R}$를

$$g : x \mapsto \begin{cases} (e^x - 1)/x & x \neq 0 \text{인 경우} \\ 1 & \text{ow.} \end{cases}$$

로 두면 이가 $x \neq 0$에서 해석적임은 자명하므로 $x = 0$에서도 해석적임을 보이면 충분한데, 이는 정리 0. 44와 임의의 $x \in \mathbb{R}$에 대해 $g(x) = \sum_{i=1}^{\infty} x^{i-1}/i!$이라는 점에서 자명하다. □

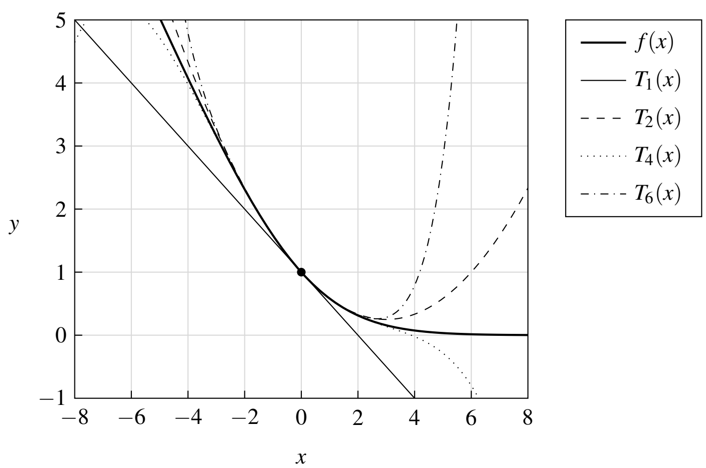


Figure 0.21 Bernoulli 수의 생성함수와 이의 0에서의 Taylor 근사식의 그래프. Computed by Wolfram MATHEMATICA.

수열 $\{B_i\}$의 짝수 항은 양수와 음수를 오가며 굉장히 불규칙하게 행동하는 반면 홀수 항은 B_1을 제외하고는 모두 0이다.

Theorem 0.93 임의의 $k \in \mathbb{N}$에 대해

Table 0.4 Bernoulli 수

k	0	1	2	3	4	5	6	7	8
B_k	1	$-1/2$	$1/6$	0	$-1/30$	0	$1/42$	0	$-1/30$
근사치	1	-0.5	0.16667	0	-0.03333	0	0.02381	0	-0.03333

$$B_{2k-1} = \begin{cases} -1/2 & k = 1\text{인 경우} \\ 0 & \text{ow.} \end{cases}$$

이다.

Proof 함수 $f : \mathbb{R} \to \mathbb{R}$를 Bernoulli 수의 생성함수라 하면

$$\begin{aligned} B_1 &= f'(0) \\ &= \lim_{h\to 0} \frac{f(h)-1}{h} \\ &= \lim_{h\to 0} \frac{h-e^h+1}{h(e^h-1)} \\ &= \lim_{h\to 0} \frac{1-e^h}{(h+1)e^h-1} \\ &= -\lim_{h\to 0} \frac{1}{h+2} \\ &= -\frac{1}{2} \end{aligned}$$

이다. 여기서 네 번째와 다섯 번째 등호는 L'Hospital의 법칙으로부터 성립한다. 한편, 함수 $g : \mathbb{R} \to \mathbb{R}$를 $g : x \mapsto f(x) + x/2$로 두면 이는 우함수가 되어 임의의 $n \in \mathbb{N}$에 대해 $g^{(2n-1)}(0) = 0$이고, 이로써 증명은 충분하다. □

다음으로 준비할 것은 Bernoulli 다항식으로, Bernoulli 수와 비슷하게 생성함수로써 정의된다.

Definition 0.94 함수 $f_x : \mathbb{R} \to \mathbb{R}$를 다음과 같이 정의하자.

$$f_x : t \mapsto \begin{cases} te^{xt}/(e^t-1) & t \neq 0\text{인 경우} \\ 1 & \text{ow.} \end{cases}$$

이때, 이의 0에서의 Taylor 전개의 k차항의 계수를 k**번째 Bernoulli 다항식 (**k**th − polynomial)**이라 하고 $B_k(x)$로 쓴다. 즉, 0 근방의 $t \in \mathbb{R}$에 대해 $f_x(t) = \sum_{i=0}^{\infty} B_i(x)t^i/i!$이며, 이때의 함수 f_x를 Bernoulli 다항식의 **생성함수 (generating function)**라 한다.

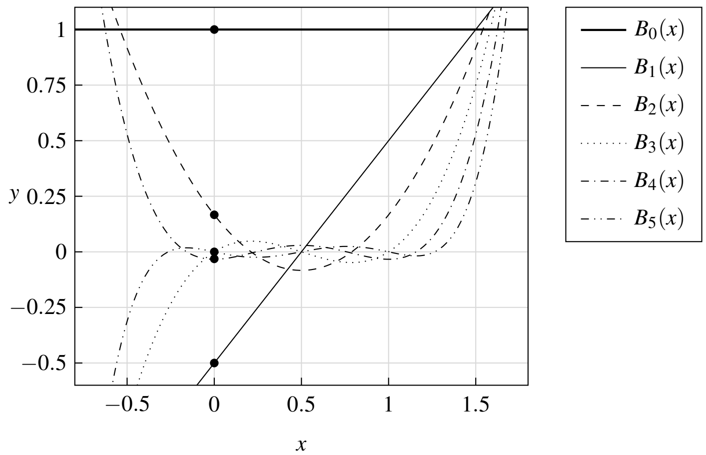


Figure 0.22 Bernoulli 다항식의 그래프. Computed by Wolfram MATHEMATICA.

Table 0.5 Bernoulli 다항식

차수 (k)	$B_k(x)$
0	1
1	$x-1/2$
2	$x^2-x+1/6$
3	$x^3-3x^2/2+x/2$
4	$x^4-2x^3+x^2-1/30$
5	$x^5-5x^4/2+5x^3/3-x/6$
6	$x^6-3x^5+5x^4/2-x^2/2+1/42$
7	$x^7-7x^6/2+7x^5/2-7x^3/6+x/6$
8	$x^8-4x^7+14x^6/3-7x^4/3+2x^2/3-1/30$
9	$x^9-9x^8/2+6x^7-21x^5/5+2x^3-3x/10$
10	$x^{10}-5x^9+15x^8/2-7x^6+5x^4-3x^2/2+5/66$

이번에도 Bernoulli 다항식의 well-definedness를 위해서는 그 생성함수가 0에서 해석적임을 보여야 하지만 이는 Bernoulli 수의 생성함수가 해석적이라는 점에서 자명하다. 한편, 그 이름에서 알 수 있듯이 Bernoulli 다항식은 Bernoulli 수와 밀접한 관련이 있다.

Theorem 0.95 임의의 $k \in \mathbb{N}_0$과 임의의 $x \in \mathbb{R}$에 대해

$$B_k(x) = \sum_{i=0}^{k} \binom{k}{i} B_i x^{k-i}$$

이다.

Proof 임의의 $x \in \mathbb{R}$에 대해 함수 f_x를 Bernoulli 다항식의 생성함수라 하면 f_0이 바로 Bernoulli 수의 생성함수이므로 0 근방의 $t \in \mathbb{R}$에 대해

$$\begin{aligned}\sum_{i=0}^{\infty} B_i(x)\frac{t^i}{i!} &= f_x(t) \\ &= f_0(t)e^{xt} \\ &= \left(\sum_{i=0}^{\infty} B_i \frac{t^i}{i!}\right)\left[\sum_{i=0}^{\infty} \frac{(xt)^i}{i!}\right] \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{i} B_j x^{i-j}\frac{t^i}{j!(i-j)!} \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{i} \binom{i}{j} B_j x^{i-j}\frac{t^i}{i!}\end{aligned}$$

이므로 양변에서 $t^k/k!$의 계수를 비교하면 증명이 끝난다. □

Theorem 0.96 임의의 $k \in \mathbb{N}_0$에 대해 다음이 성립한다.

i. k차 Bernoulli 다항식은 k차 다항식이다.
ii. $B_k(1-x) = (-1)^k B_k(x)$.
iii. $B_k(0) = B_k$.
iv. $2k+1$차 Bernoulli 다항식은 $1/2$을 영점으로 가진다.
v. $\lim_{x\to-\infty}(-1)^k B_k(x) = \infty$.
vi. $\lim_{x\to\infty} B_k(x) = \infty$.
vii. k차 Bernoulli 다항식은 $\mathscr{C}^\infty$급이며 $n \le k$인 $n \in \mathbb{N}_0$에 대해 $B_k^{(n)} = k^{\underline{n}} B_{k-n}$이다.

Proof i, iii, v, vi, vii. 이는 위의 보조정리로부터 자명하다.

ii. 임의의 $x \in \mathbb{R}$에 대해 함수 f_x를 Bernoulli 다항식의 생성함수라 하면 0 근방의 $t \in \mathbb{R}$에 대해 $\sum_{i=0}^{\infty} B_i(1-x)t^i/i! = f_{1-x}(t) = f_x(-t) = \sum_{i=0}^{\infty}(-1)^i B_i(x)t^i/i!$이므로 임의의 $k \in \mathbb{N}$에 대해 양변의 $t^k/k!$의 계수를 비교하면 $B_k(1-x) = (-1)^k B_k(x)$가 성립함을 안다.

iv. ii에서 $B_{2k+1}(1/2) = -B_{2k+1}(1/2)$이므로 $B_{2k+1}(1/2) = 0$이다.

viii. 임의의 $k \in \mathbb{N}$에 대해 $B_k' = kB_{k-1}$이 성립함을 보이는 것으로 충분한데, 이는

$$\begin{aligned}B_k'(x) &= \sum_{i=0}^{k-1}(k-i)\binom{k}{i}B_i x^{k-i-1} \\ &= \sum_{i=0}^{k-1}\frac{k!}{(k-i-1)!i!}B_i x^{k-i-1} \\ &= k\sum_{i=0}^{k-1}\binom{k-1}{i}B_i x^{k-i-1}\end{aligned}$$

$$= kB_{k-1}(x)$$

에서 자명하다. □

이제 준비는 끝났다. 우리는 4개의 보조정리를 통해 고등학생도 이해할 수 있을 만한 방법으로 $\zeta(2n)$의 값을 구해낼 것이다.

Lemma 0.97 (Newton's power sum formula) k차 다항식 $P(x) = \sum_{i=0}^{k} a_i x^i$에 대해 방정식 $P(x) = 0$의 k개의 근을 $z_1, \cdots, z_k \in \mathbb{C}$라 하고 임의의 $i \in \mathbb{N}$에 대해 $S_i = \sum_{j=1}^{k} z_j^i$라 하면 임의의 $l \in \mathbb{N}$에 대해

$$\begin{cases} la_{k-l} + \displaystyle\sum_{i=1}^{l} a_{k-l+i} S_i = 0 & l \leq k\text{인 경우} \\ \displaystyle\sum_{i=l-k}^{k} a_{k-l+i} S_i = 0 & \text{ow.} \end{cases}$$

이다.

Proof 먼저 $z_1, \cdots, z_k$가 모두 0이 아닌 경우를 생각하자. 이때 $a_0 = \prod_{i=1}^{k} \alpha_i \neq 0$이므로 다항식 $Q(x) = \sum_{i=0}^{k} a_{k-i} x^i = x^k P(1/x)$를 생각하면 이는 k차 다항식이고 방정식 $Q(x) = 0$는 $1/z_1, \cdots, 1/z_k$를 근으로 갖는다. 한편, $Q(x) = 0$는 0을 근으로 갖지 않으므로 0의 적당한 근방 $I \subseteq \mathbb{R}$에서 Q는 0이 아니고, 곧 함수 $f : I \to \mathbb{R}$를 $f : x \mapsto \log|Q(x)|$로 정의하면 이는 well-define되며 $\mathscr{C}^\infty$급이다. 이로부터 임의의 $n \in \mathbb{N}$에 대해

$$\begin{aligned} f^{(n)}(x) &= \frac{d^n}{dx^n} \log\left(|a_0| \prod_{i=1}^{k} \left| x - \frac{1}{z_i} \right| \right) \\ &= \sum_{i=1}^{k} \frac{d^{n-1}}{dx^{n-1}} \frac{\alpha_i}{\alpha_i x - 1} \\ &= (-1)^{n-1}(n-1)! \sum_{i=1}^{k} \left(\frac{z_i}{z_i x - 1} \right)^n \end{aligned}$$

이므로 $f^{(n)}(0) = -(n-1)!S_n$이고, 곧 임의의 $l \in \mathbb{N}$에 대해

$$\begin{aligned} \frac{Q^{(l)}(0)}{l!} &= \frac{(f'Q)^{(l-1)}(0)}{l!} \\ &= \frac{1}{l!} \sum_{i=0}^{l-1} \binom{l-1}{i} f^{(i+1)}(0) Q^{(l-i-1)}(0) \\ &= -\frac{1}{l!} \sum_{i=0}^{l-1} \binom{l-1}{i} i! S_{i+1} Q^{(l-i-1)}(0) \end{aligned}$$

$$= -\frac{1}{l}\sum_{i=1}^{l}\frac{Q^{(l-i)}(0)}{(l-i)!}S_i$$

인데 만약 $l \le k$이면 이는 $ka_{k-l} + \sum_{i=1}^{l} a_{k-l+i}S_i = 0$을 함의하고, 반대로 $l > k$이면 $\sum_{i=l-k}^{k} a_{k-l+i}S_i = 0$을 함의하여 증명이 끝난다.

이제 일반적인 경우에 대해, $z_1, \cdots, z_n$ 중에서 0이 아닌 것의 개수를 $m \in \mathbb{N}_0$이라 하자. 만약 $m = 0$이면, 즉 $z_1 = \cdots = z_k = 0$이면 보조정리가 자명하므로 $m > 0$이라 하고, WLOG, 필요하다면 relabel하여 $z_1, \cdots, z_m$이 0이 아니라 하자. 그렇다면 m차 다항식 $\widetilde{P}(x) = \sum_{i=0}^{m} a_{i+k-m}x^i = P(x)/x^{k-m}$에 대해 방정식 $\widetilde{P}(x) = 0$은 모두 0이 아닌 $z_1, \cdots, z_m$을 근으로 가지므로 앞선 결론으로부터 임의의 $l \in \mathbb{N}$에 대해

$$\begin{cases} la_{k-l} + \sum_{i=1}^{l} a_{k-l+i}S_i = 0 & l \le m\text{인 경우} \\ \sum_{i=l-m}^{k} a_{k-l+i}S_i = 0 & \text{ow.} \end{cases}$$

이고, 가정으로부터 $a_0 = \cdots = a_{k-m-1} = 0$이므로 보조정리가 성립한다. □

Lemma 0.98 임의의 $k \in \mathbb{N}$에 대해 다음이 성립한다.

$$\sum_{i=0}^{k}\frac{2^{2i}B_{2i}}{(2i)!(2k+1-2i)!} = \frac{1}{(2k)!}$$

Proof 정리 0.93, 0.95, 0.96의 iv로부터

$$\begin{aligned} 0 &= B_{2k+1}\left(\frac{1}{2}\right) \\ &= \sum_{i=0}^{2k+1}\binom{2k+1}{i}\frac{B_i}{2^{2k+1-i}} \\ &= \sum_{i=0}^{k}\binom{2k+1}{2i}\frac{B_{2i}}{2^{2k+1-2i}} - \frac{2k+1}{2^{2k+1}} \\ &= \frac{(2k+1)!}{2^{2k+1}}\left[\sum_{i=0}^{k}\frac{2^{2i}B_{2i}}{(2i)!(2k+1-2i)!} - \frac{1}{(2k)!}\right] \end{aligned}$$

이 되어 보조정리가 성립함을 안다. □

Lemma 0.99 (Papadimitriou) 임의의 $k \in \mathbb{N}$에 대해 방정식

$$\sum_{i=0}^{k}(-1)^i\binom{2k+1}{2i+1}x^{k-i} = 0$$

은 $\cot^2 \pi/(2k+1), \cdots, \cot^2 k\pi/(2k+1)$ 을 서로다른 k개의 근으로 갖는다.

Proof Euler의 공식으로부터 임의의 $\theta \in \mathbb{R}$에 대해

$$\begin{aligned}
&\cos(2k+1)\theta + \mathbf{i}\sin(2k+1)\theta \\
&\quad = e^{\mathbf{i}(2k+1)\theta} \\
&\quad = (\cos\theta + \mathbf{i}\sin\theta)^{2k+1} \\
&\quad = \sum_{i=0}^{2k+1} \binom{2k+1}{i} \cos^{2k+1-i}\theta \sin^i \theta \mathbf{i}^i \\
&\quad = \sum_{i=0}^{k} (-1)^i \binom{2k+1}{2i} \cos^{2k+1-2i}\theta \sin^{2i}\theta + \mathbf{i}\sum_{i=0}^{k} (-1)^i \binom{2k+1}{2i+1} \cos^{2k-2i}\theta \sin^{2i+1}\theta
\end{aligned}$$

이므로 양변의 허수부를 비교하면 π의 정수배가 아닌 θ에 대해

$$\begin{aligned}
\frac{\sin(2k+1)\theta}{\sin^{2k+1}\theta} &= \frac{1}{\sin^{2k+1}\theta} \sum_{i=0}^{k} (-1)^i \binom{2k+1}{2i+1} \cos^{2k-2i}\theta \sin^{2i+1}\theta \\
&= \sum_{i=0}^{k} (-1)^i \binom{2k+1}{2i+1} \cot^{2(k-i)}\theta
\end{aligned}$$

가 되어 $\cot^2 \pi/(2k+1), \cdots, \cot^2 k\pi/(2k+1)$ 이 주어진 방정식의 근임을 알고, 이들이 서로 다르다는 점은 자명하다. □

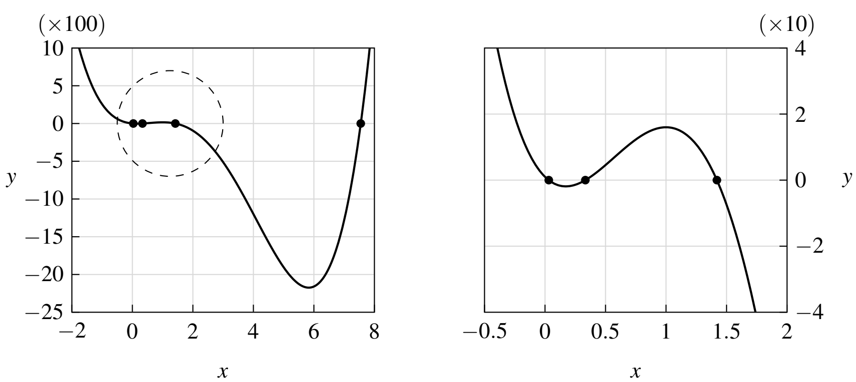


Figure 0.23 Papadimitriou의 보조정리에서 $k = 4$인 경우의 다항식 $9x^4 - 84x^3 + 126x^2 - 36x + 1$의 그래프 (왼쪽) 와 $[-1/2, 2]$에서 이를 확대한 그림 (오른쪽). Papadimitriou의 보조정리로부터 이 다항식은 $\cot^2 \pi/9 \approx 7.54863$, $\cot^2 2\pi/9 \approx 1.42028$, $\cot^2 \pi/3 \approx 0.33333$, $\cot^2 4\pi/9 \approx 0.03109$를 영점으로 가진다.

Lemma 0.100 임의의 $k, n \in \mathbb{N}$에 대해 k가 충분히 크면 다음이 성립한다.

$$\sum_{i=1}^{k} \cot^{2n} \frac{i}{2k+1}\pi = (-1)^{n-1} \frac{2^{4n-1} B_{2n}}{(2n)!} k^{2n} + O(k^{2n-1})$$

여기서 big-O 표기법은 $k \to \infty$인 경우에 대한 근사로 n과는 무관하다.

Proof 수학적 귀납법을 사용하자. 먼저 $n = 1$인 경우에는 Papadimitriou의 보조정리로부터 $\cot^2 \pi/(2k+1), \cdots, \cot^2 k\pi/(2k+1)$이 방정식

$$\sum_{i=0}^{k} (-1)^i \binom{2k+1}{2i+1} x^{k-i} = 0$$

의 근이므로 근과 계수의 관계에서

$$\begin{aligned}
\sum_{i=1}^{k} \cot^2 \frac{i}{2k+1}\pi &= \binom{2k+1}{1} \Big/ \binom{2k+1}{3} \\
&= \frac{k(2k-1)}{3} \\
&= \frac{2}{3}k^2 + O(k)
\end{aligned}$$

이 되어 보조정리가 성립한다. 이제 보조정리가 $1, \cdots, n$에 대해 모두 성립한다고 가정하고 각 $i \in \mathbb{N}$에 대해 $S_i = \sum_{j=1}^{k} \cot^{2i} j\pi/(2k+1)$이라 하면 Newton의 power sum formula로부터 $k \in \mathbb{N}$가 충분히 클 때

$$\begin{aligned}
0 &= (-1)^{n+1}(n+1)\binom{2k+1}{2n+3} + \sum_{i=1}^{n+1} (-1)^{n+1-i} \binom{2k+1}{2n-2i+3} S_i \\
&= (2k+1)\left[\frac{(-1)^{n+1}(n+1)}{2k+1} \binom{2k+1}{2n+3} + \sum_{i=1}^{n} \frac{(-1)^{n+1-i}}{2k+1} \binom{2k+1}{2n-2i+3} S_i + S_{n+1} \right]
\end{aligned}$$

이 되어

$$S_{n+1} = \frac{(-1)^n(n+1)}{2k+1} \binom{2k+1}{2n+3} + \sum_{i=1}^{n} \frac{(-1)^{n-i}}{2k+1} \binom{2k+1}{2n-2i+3} S_i$$

인데, 여기서

$$\begin{aligned}
\frac{(-1)^n(n+1)}{2k+1} \binom{2k+1}{2n+3} &= (-1)^n \frac{(2k)^{2n+2}}{(2n+3)!}(n+1) \\
&= \frac{(-1)^n 2^{2n+2}(n+1)}{(2n+3)!} k^{2n+2} + O(k^{2n+1})
\end{aligned}$$

이고 가정으로부터 각 $i \leq n$에 대해

$$\begin{aligned}
&\frac{(-1)^{n-i}}{2k+1}\binom{2k+1}{2n-2i+3}S_i \\
&\qquad = \left[(-1)^{n-i}\frac{(2k)^{2n-2i+2}}{(2n-2i+3)!}\right]\left[(-1)^{i-1}\frac{2^{4i-1}B_{2i}}{(2i)!}k^{2i}+O(k^{2i-1})\right] \\
&\qquad = \frac{(-1)^{n-1}2^{2n+2i+1}B_{2i}}{(2i)!(2n-2i+3)!}k^{2n+2}+O(k^{2n+1})
\end{aligned}$$

이므로 이상을 종합하면 보조정리 0. 98으로부터 k가 충분히 클 때

$$\begin{aligned}
S_{n+1} &= \frac{(-1)^n 2^{2(n+1)}(n+1)}{(2n+3)!}k^{2n+2}+O(k^{2n+1}) \\
&\qquad\qquad + \sum_{i=1}^{n}\frac{(-1)^{n-1}2^{2n+2i+1}B_{2i}}{(2i)!(2n-2i+3)!}k^{2n+2}+O(k^{2n+1}) \\
&= (-1)^n 2^{2n+1}k^{2n+2}\left[\frac{1}{(2n+2)!}-\sum_{i=0}^{n}\frac{2^{2i}B_{2i}}{(2i)!(2n+3-2i)!}\right]+O(k^{2n+1}) \\
&= (-1)^n\frac{2^{4n+3}B_{2n+2}}{(2n+2)!}k^{2n+2}+O(k^{2n+1})
\end{aligned}$$

이 되어 증명이 끝난다. □

Theorem 0.101 (Basel problem) 임의의 $n \in \mathbb{N}$에 대해 다음이 성립한다.

$$\zeta(2n) = (-1)^{n-1}\frac{(2\pi)^{2n}B_{2n}}{2(2n)!}$$

Proof 임의의 $\theta \in (0, \pi/2)$에 대해 $\sin\theta < \theta < \tan\theta$이므로 $\cot^2\theta < \theta^{-2} < \operatorname{cosec}^2\theta = 1+\cot^2\theta$에서 $\cot^{2n}\theta < \theta^{-2n} < (1+\cot^2\theta)^n$이 임의의 $n \in \mathbb{N}$에 대해 성립한다. 따라서 이항정리와 위의 보조정리로부터 임의의 $k \in \mathbb{N}$에 대해 이가 충분히 크면

$$\begin{aligned}
(-1)^{n-1}\frac{2^{4n-1}B_{2n}}{(2n)!}k^{2n}+O(k^{2n-1}) &= \sum_{i=1}^{k}\cot^{2n}\frac{i}{2k+1}\pi \\
&< \frac{(2k+1)^{2n}}{\pi^{2n}}\sum_{i=1}^{k}\frac{1}{i^{2n}} \\
&< \sum_{i=1}^{k}\left(1+\cot^2\frac{i}{2k+1}\pi\right)^n \\
&= \sum_{i=1}^{k}\left[\cot^{2n}\frac{i}{2k+1}\pi+O\left(\cot^{2(n-1)}\frac{i}{2k+1}\pi\right)\right] \\
&= \sum_{i=1}^{k}\cot^{2n}\frac{i}{2k+1}\pi+O\left(\sum_{i=1}^{k}\cot^{2(n-1)}\frac{i}{2k+1}\pi\right)
\end{aligned}$$

$$= \sum_{i=1}^{k} \cot^{2n} \frac{i}{2k+1}\pi + O(k^{2n-2})$$
$$= (-1)^{n-1} \frac{2^{4n-1}B_{2n}}{(2n)!} k^{2n} + O(k^{2n-1})$$

이고, 곧 조임정리로부터

$$\begin{aligned}\zeta(2n) &= \lim_{k\to\infty} \sum_{i=1}^{k} \frac{1}{i^{2n}} \\ &= (-1)^{n-1} \frac{2^{4n-1}\pi^{2n}B_{2n}}{(2n)!} \lim_{k\to\infty} \frac{k^{2n}}{(2k+1)^{2n}} + \lim_{k\to\infty} \frac{O(k^{2n-1})}{(2k+1)^{2n}} \\ &= (-1)^{n-1} \frac{(2\pi)^{2n}B_{2n}}{2(2n)!}\end{aligned}$$

이 되어 증명이 끝난다. □

제타함수와 아무런 관련도 없어 보이던 Bernoulli 수가 $\zeta(2n)$의 값으로 등장한다는 것은 조금 놀라운 사실이다. 한편, 이와 한 끝 차이인 $\zeta(2n+1)$의 값을 구하는 문제는 아직도 미해결이며, 심지어 $\zeta(3) = \sum_{i=1}^{\infty} 1/i^3 \approx 1.2020569$의 값도 아직 닫힌 형태로 구해지지 않았다. 다만 Roger Apéry에 의해 $\zeta(3)$이 무리수라는 것이 밝혀져 **Apéry 상수 (− constant)** 라는 이름이 붙어있다. 시간이 남는다면 다들 한 번 도전해보자.

0.2.4 Special Functions Related to Cantor Set

이번에는 해석개론 연습문제에서 단골로 등장하는 Cantor 집합과 관련된 함수를 공부할 텐데, 이에 앞서 Cantor 집합에 대해서 조금 알아보도록 하자. Cantor 집합은 fractal의 일종으르 그 일부를 확대하면 이는 다시 Cantor 집합의 모습을 띄는 자기 유사성을 가진다. 또한, Cantor 집합은 길이가 1인 선분에서 거의 대부분을 제거하여 얻어지므로 이를테면 군데군데 흩어진 '점' 으로만 이루어진 집합이지만 실수와 같은 개수의 원소를 가지는 꽤나 반직관적인 면모도 가지고 있다. 이렇게 흥미로우면서도 우리의 직관에 반하는 듯한 묘한 특성 덕에 일반 대중들에게도 비교적 잘 알려져 있고, 해석학에서 각종 그럴싸한 명제의 반례를 만들어내는 데 유용하게 사용되고, Smith–Volterra–Cantor 집합[7]같은 아류도 조금 있다. 아무튼, 이 절은 Cantor 집합 자체가 주제는 아니니 각종 아류나 Hausdorff 차원[8]같은 fractal 기하와 관련된 결과는 잠시 치워두고 Cantor 집합과 관련된 특수함수를 논하는데 필요한 정도만 다루어보고자 한다. 다만, Cantor 집합이나 이와 관련된 함수의 성질에는 측도론적인 성질이 핵심적인 부분을 차지하므로 측도론이 아직 익숙하지 않다면 1장을 보고 다시 돌아오는 것을 추천하며, 이제부터는 기본적인 측도론의 지식을 전제한다.

보통 Cantor 집합은 $[0,1]$에서 중간 $1/3$을 빼고, 다시 남은 것들의 중간 $1/3$을 빼고 이런 식의 조작을 무한번 반복하면 얻을 수 있다는 식으로 정의된다. 이제 이걸 엄밀한 수학의 언어로 옮기는 작업이 관건인데, 가장 담백한 방법은 바로 집합열을 이용하는 방법이다.

Definition 0.102 집합열 $\{C_i\}$를

$$C_i = \begin{cases} C_{i-1}/3 \cup (C_{i-1}/3 + 2/3) & i > 0\text{인 경우} \\ [0,1] & \text{ow.} \end{cases}$$

로 두자. 이때, $\bigcap_{i=1}^{\infty} C_i$를 **Cantor 집합 (– set)**이라 하고 C로 쓴다.

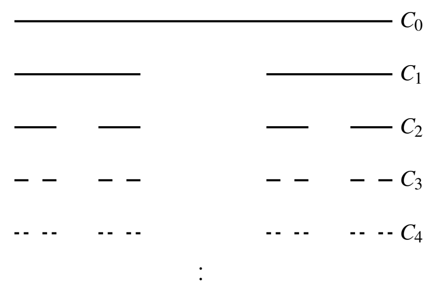


Figure 0.24 정의 0. 102에서의 집합열 $\{C_i\}$를 통해 Cantor 집합을 구성하는 과정. 흔히 아는 것과 같이 $[0,1]$에서 시작하여 이전 단계의 중간 $1/3$을 빼는 과정을 무한번 반복하여 Cantor 집합을 얻는다.

위 정의의 점화식을 조금만 주의깊게 살펴보면 이가 어떻게 '중간 $1/3$을 빼는' 과정을 담아내는 지 파악할 수 있다. 즉, 점화식에서 $C_{i-1}/3$에서는 C_{i-1}의 왼쪽 $1/3$을 택하고, $C_{i-1}/3 + 2/3$에서는 오른쪽 $1/3$을 택하여 이들을 합집합하면 결과적으로 C_i는 C_{i-1}의 중간 $1/3$을 뺀 집합이 된다. 그리고 이렇게 얻은 C_i를 모두 교집합하면 이것이 곧 $[0,1]$에서 중간 $1/3$을 빼는 조작을 무한번 반복하는 것이 되어 $\bigcap_{i=1}^{\infty} C_i$가 Cantor 집합이 된다. 한편, 여기서의 집합열 $\{C_i\}$는 비록 복잡하지만 그 일반항을 닫힌 형태로 구할 수 있다. 물론, Hermite 다항식의 닫힌 형태가 그닥 쓸모가 없었던 것처럼 이 또한 크게 쓸모는 없다.

Theorem 0.103 집합열 $\{C_i\}$를 정의 0. 102에서와 같이 두면 임의의 $i \in \mathbb{N}$에 대해

$$C_i = \bigcap_{j=1}^{i} \bigcup_{k=1}^{3^{j-1}} \left(\left[\frac{3k-3}{3^j}, \frac{3k-2}{3^j} \right] \cup \left[\frac{3k-1}{3^j}, \frac{3k}{3^j} \right] \right)$$

이다.

Proof 증명은 수학적 귀납법을 사용한다. 우선 $i = 1$인 경우에는 $C_1 = [0,1/3] \cup [2/3,1]$이 되어 정리가 자명하므로 귀납가정으로서 $i \in \mathbb{N}$에 대해 정리가 성립한다고 가정하고 $A = C_i/3 + (C_i/3 + 2/3)$, $B = \bigcap_{j=1}^{i+1} \bigcup_{k=1}^{3^{j-1}} \{[(3k-3)/3^j, (3k-2)/3^j] \cup [(3k-1)/3^j, 3k/3^j]\}$

라 하여 $A=B$임을 보이자. 먼저 임의의 $x\in A$를 택하면 $x\in C_i/3$이거나 $x\in C_i/3+2/3$이므로 경우를 나누어 생각하자. 전자의 경우 $3x\in C_i\subseteq[0,1]$에서 $x\in[0,1/3]$이고, 귀납가정으로부터 각 $j\le i$에 대해 $3x\in\bigcup_{k=1}^{3^{j-1}}\{[(3k-3)/3^j,(3k-2)/3^j]\cup[(3k-1)/3^j,3k/3^j]\}$이므로 $x\in\bigcup_{j=1}^{3^{j-1}}\{[(3k-3)/3^{j+1},(3k-2)/3^{j+1}]\cup[(3k-1)/3^{j+1},3k/3^{j+1}]\}$이 되어

$$\begin{aligned}
x&\in\left(\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]\right)\cap\bigcap_{j=1}^{i}\bigcup_{k=1}^{3^{j-1}}\left(\left[\frac{3k-3}{3^{j+1}},\frac{3k-2}{3^{j+1}}\right]\cup\left[\frac{3k-1}{3^{j+1}},\frac{3k}{3^{j+1}}\right]\right)\\
&=\left(\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]\right)\cap\bigcap_{j=2}^{i+1}\bigcup_{k=1}^{3^{j-2}}\left(\left[\frac{3k-3}{3^{j}},\frac{3k-2}{3^{j}}\right]\cup\left[\frac{3k-1}{3^{j}},\frac{3k}{3^{j}}\right]\right)\\
&\subseteq\left(\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]\right)\cap\bigcap_{j=2}^{i+1}\bigcup_{k=1}^{3^{j-1}}\left(\left[\frac{3k-3}{3^{j}},\frac{3k-2}{3^{j}}\right]\cup\left[\frac{3k-1}{3^{j}},\frac{3k}{3^{j}}\right]\right)\\
&=\bigcap_{j=1}^{i+1}\bigcup_{k=1}^{3^{j-1}}\left(\left[\frac{3k-3}{3^{j}},\frac{3k-2}{3^{j}}\right]\cup\left[\frac{3k-1}{3^{j}},\frac{3k}{3^{j}}\right]\right)\\
&=B
\end{aligned}$$

이다. 후자의 경우에도 비슷하게 $3x-2\in C_i\subseteq[0,1]$에서 $x\in[2/3,1]$이고, 귀납가정으로부터 각 $j\le i$에 대해 $3x-2\in\bigcup_{k=1}^{3^{j-1}}\{[(3k-3)/3^j,(3k-2)/3^j]\cup[(3k-1)/3^j,3k/3^j]\}$이므로 $x\in\bigcup_{k=1}^{3^{j-1}}([(3k-3)/3^{j+1}+2/3,(3k-2)/3^{j+1}+2/3]\cup[(3k-1)/3^{j+1}+2/3,3k/3^{j+1}+2/3])$이 되어

$$\begin{aligned}
x&\in\left(\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]\right)\cap\bigcap_{j=1}^{i}\bigcup_{k=1}^{3^{j-1}}\left(\left[\frac{3k-3}{3^{j+1}}+\frac{2}{3},\frac{3k-2}{3^{j+1}}+\frac{2}{3}\right]\cup\left[\frac{3k-1}{3^{j+1}}+\frac{2}{3},\frac{3k}{3^{j+1}}+\frac{2}{3}\right]\right)\\
&=\left(\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]\right)\\
&\quad\cap\bigcap_{j=1}^{i}\bigcup_{k=1}^{3^{j-1}}\left(\left[\frac{3(k+2\cdot3^{j-1})-3}{3^{j+1}},\frac{3(k+2\cdot3^{j-1})-2}{3^{j+1}}\right]\cup\left[\frac{3(k+2\cdot3^{j-1})-1}{3^{j+1}},\frac{3(k+2\cdot3^{j-1})}{3^{j+1}}\right]\right)\\
&=\left(\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]\right)\cap\bigcap_{j=1}^{i}\bigcup_{k=2\cdot3^{j-1}+1}^{3^{i}}\left(\left[\frac{3k-3}{3^{j+1}},\frac{3k-2}{3^{j+1}}\right]\cup\left[\frac{3k-1}{3^{j+1}},\frac{3k}{3^{j+1}}\right]\right)\\
&=\left(\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]\right)\cap\bigcap_{j=2}^{i+1}\bigcup_{k=2\cdot3^{j-2}+1}^{3^{j-1}}\left(\left[\frac{3k-3}{3^{j}},\frac{3k-2}{3^{j}}\right]\cup\left[\frac{3k-1}{3^{j}},\frac{3k}{3^{j}}\right]\right)\\
&\subseteq\left(\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]\right)\cap\bigcap_{j=2}^{i+1}\bigcup_{k=1}^{3^{j-1}}\left(\left[\frac{3k-3}{3^{j}},\frac{3k-2}{3^{j}}\right]\cup\left[\frac{3k-1}{3^{j}},\frac{3k}{3^{j}}\right]\right)
\end{aligned}$$

$$
\begin{aligned}
&= \bigcap_{j=1}^{i+1} \bigcup_{k=1}^{3^{j-1}} \left(\left[\frac{3k-3}{3^j}, \frac{3k-2}{3^j} \right] \cup \left[\frac{3k-1}{3^j}, \frac{3k}{3^j} \right] \right) \\
&= B
\end{aligned}
$$

에서 곧 어느 경우에나 $x \in B$ 이고, 이로부터 $A \subseteq B$ 임을 안다. 역으로 임의의 $x \in B$ 를 택하면 $x \in [0, 1/3] \cup [2/3, 1]$ 이므로 이번에도 경우를 나누어 생각해보자. 먼저 $x \in [0, 1/3]$ 인 경우에는 집합 B 의 정의로부터 각 $1 < j \le i+1$ 에 대해 $x \in \bigcup_{k=1}^{3^{j-2}} ([(3k-3)/3^j, (3k-2)/3^j] \cup [(3k-1)/3^j, 3k/3^j])$ 이므로 $3x \in \bigcup_{k=1}^{3^{j-2}} ([(3k-3)/3^{j-1}, (3k-2)/3^{j-1}] \cup [(3k-1)/3^{j-1}, 3k/3^{j-1}])$ 이 되어

$$
\begin{aligned}
3x &\in \bigcap_{j=2}^{i+1} \bigcup_{k=1}^{3^{j-2}} \left(\left[\frac{3k-3}{3^{j+1}}, \frac{3k-2}{3^{j+1}} \right] \cup \left[\frac{3k-1}{3^{j+1}}, \frac{3k}{3^{j+1}} \right] \right) \\
&= \bigcap_{j=1}^{i} \bigcup_{k=1}^{3^{j-1}} \left(\left[\frac{3k-3}{3^j}, \frac{3k-2}{3^j} \right] \cup \left[\frac{3k-1}{3^j}, \frac{3k}{3^j} \right] \right) \\
&= C_i
\end{aligned}
$$

에서 $x \in C_i/3 \subseteq A$ 이다. 한편, $x \in [2/3, 1]$ 인 경우에는 집합 B 의 정의로부터 각 $1 < j \le i+1$ 에 대해

$$
\begin{aligned}
x &\in \bigcup_{k=2\cdot 3^{j-1}+1}^{3^{j-1}} \left(\left[\frac{3k-3}{3^j}, \frac{3k-2}{3^j} \right] \cup \left[\frac{3k-1}{3^j}, \frac{3k}{3^j} \right] \right) \\
&= \bigcup_{k=1}^{3^{j-2}} \left(\left[\frac{3(k+2\cdot 3^{j-2})-3}{3^j}, \frac{3(k+2\cdot 3^{j-2})-2}{3^j} \right] \cup \left[\frac{3(k+2\cdot 3^{j-2})-1}{3^j}, \frac{3(k+2\cdot 3^{j-2})}{3^j} \right] \right) \\
&= \bigcup_{k=1}^{3^{j-2}} \left(\left[\frac{3k-3}{3^j} + \frac{2}{3}, \frac{3k-2}{3^j} + \frac{2}{3} \right] \cup \left[\frac{3k-1}{3^j} + \frac{2}{3}, \frac{3k}{3^j} + \frac{2}{3} \right] \right)
\end{aligned}
$$

이므로 $3x-2 \in \bigcup_{k=1}^{3^{j-2}} ([(3k-3)/3^{j-1}, (3k-2)/3^{j-1}] \cup [(3k-1)/3^{j-1}, 3k/3^{j-1}])$ 이 되어 방금과 비슷하게 $x \in C_i/3 + 2/3 \subseteq A$ 임을 보일 수 있다. 이로부터 곧 어느 경우에나 $x \in A$ 가 되어 $B \subseteq A$ 임을 알고, 증명은 이로써 충분하다. □

Corollary 0.104

$$
C = \bigcap_{i=1}^{\infty} \bigcup_{j=1}^{3^{i-1}} \left(\left[\frac{3j-3}{3^i}, \frac{3j-2}{3^i} \right] \cup \left[\frac{3j-1}{3^i}, \frac{3j}{3^i} \right] \right)
$$

PROOF 이는 Cantor 집합의 정의와 위의 정리로부터 자명하다. □

Proposition 0.105 집합열 $\{C_i\}$를 정의 0.102에서와 같이 두면 다음이 성립한다.

i. 집합열 $\{C_i\}$는 감소하는 compact한 집합열이다.
ii. 각 $i \in \mathbb{N}_0$에 대해 C_i는 2^i개의 서로소인 닫힌구간 $I_{i1}, \cdots, I_{i2^i}$로 이루어져 있으며 각 구간은 적당한 $n \in \mathbb{N}_0$에 대해 $[n/3^i, (n+1)/3^i]$로 쓸 수 있다.
iii. 집합 $[0,1] \setminus C$는 서로소인 열린구간의 열 $\{I_j\}$의 가산 합집합이며, 이때 각 I_j는 적당한 $n, i \in \mathbb{N}_0$에 대해 $(n/3^i, (n+1)/3^i)$로 쓸 수 있다.

Proof i. 이는 정리 0.103로부터 자명하다.

ii. 수학적 귀납법을 사용하자. 우선 $i = 0$인 경우에는 $C_0 = [0,1]$에서 자명하므로 귀납가정으로서 $i \in \mathbb{N}_0$에 대해 정리가 성립한다고 하면 $C_i = \bigsqcup_{j=1}^{2^i} I_{ij}$이고 이때 각 $j \le 2^i$에 대해 적당한 $n_j \in \mathbb{N}_0$가 존재하여 $I_{ij} = [n_j/3^i, (n_j+1)/3^i]$이다. 나아가 $C_{i+1} = C_i/3 \cup (C_i/3 + 2/3)$에서 $C_i \subseteq [0,1]$임이 분명한데, 이는 $C_i/3$와 $C_i/3 + 2/3$가 서로소임을 뜻하므로

$$\begin{aligned}
C_{i+1} &= \frac{C_i}{3} \sqcup \left(\frac{C_i}{3} + \frac{2}{3}\right) \\
&= \bigsqcup_{j=1}^{2^i} \frac{I_{ij}}{3} \sqcup \left(\bigsqcup_{j=1}^{2^i} \frac{I_{ij}}{3} + \frac{2}{3}\right) \\
&= \bigsqcup_{j=1}^{2^i} \left[\frac{n_j}{3^{i+1}}, \frac{n_j+1}{3^{i+1}}\right] \sqcup \bigsqcup_{j=1}^{2^i} \left[\frac{n_j + 2\cdot 3^i}{3^{i+1}}, \frac{n_j + 2\cdot 3^i + 1}{3^{i+1}}\right]
\end{aligned}$$

가 되어 $i+1$에 대해서도 정리가 성립함을 알고, 증명이 끝난다.

iii. 집합열 $\{D_i\}$를 $D_i := C_{i-1} \setminus C_i$로 정의하면 i에서 $\{C_i\}$는 감소하므로 $\bigcup_{i=1}^{\infty} D_i = [0,1] \setminus C$임이 자명하다. 이제 서로다른 임의의 $i, j \in \mathbb{N}$에 대해 WLOG, $i < j$라 하고 $x \in D_i \cap D_j$라 하면 $x \in D_i = C_{i-1} \setminus C_i$에서 $x \notin C_i$이지만 $x \in D_j \subseteq C_j \subseteq C_i$에서 $x \in C_i$의 모순이 발생하므로 D_i와 D_j는 서로소이다. 따라서 각 D_i가 서로소인 열린구간의 열 $\{I_{ij}\}$의 가산 합집합이며, 이때 각 I_{ij}는 적당한 $n, k \in \mathbb{N}_0$에 대해 $(n/3^k, (n+1)/3^k)$로 쓸 수 있다는 것만 보이면 증명은 끝난다.

이를 보이는 것은 어렵지 않은데, 임의의 $i \in \mathbb{N}$를 고정하고 C_i와 C_{i+1}의 형태에 대해 ii를 참조하면 C_i를 구성하는 닫힌구간의 끝점이 C_{i+1}에 다시 속하기만 하면 충분함을 알 수 있다. 이를 위해 $x \in [0,1]$를 C_i를 구성하는 닫힌구간의 한 끝점이라 하면 정리 0.103로부터 적당한 $k \le 3^{i-1}$가 존재하여 x는 $(3k-3)/3^i, (3k-2)/3^i, (3k-1)/3^i, 3k/3^i$ 중 하나로 쓸 수 있다. 이제 각 경우를 나누어 생각해보면 되는데, 간결한 논의를 위해 첫 번째 경우만 생각해보면 적당한 $k \le 3^{i-1}$에 대해 $x = (3k-3)/3^i = [3(3k-2)-3]/3^{i+1}$이므로 $3k-2 \le 3^i$에서 $x \in \bigcup_{k=1}^{3^i} \{[(3k-3)/3^{i+1}, (3k-2)/3^{i+1}] \cup [(3k-1)/3^{i+1}, 3k/3^{i+1}]\}$이 되어 $x \in C_{i+1}$이다. (다른 경우도 이와 비슷하게 하면 같은 결론을 얻는다.) 증명은 이로써 충분하다. □

그렇다면 과연 Cantor 집합에 속하는 원소는 어떤 것들일까? 이 질문에 답하기 위해서는 삼진법으로 생각하는 것이 편하다. 임의의 $x \in [0,1]$을 삼진법 소수로 썼을 때, 만약 소수점 아래 첫 번째 자리가 1이라면 이는 $[0,1]$에서 중간 $1/3$을 빼는 과정에서 제거되어 Cantor 집합에 속하지 못할 것이다. 만약 첫 번째 자리는 0이나 2였지만 두 번째 자리가 1이라면 이번에는 중간 $1/3$을 빼는 과정을 두 번 반복하면 제거되어 이 또한 Cantor 집합에 속하지 못할 것이다. 비슷하게 계속 생각해보면 x가 Cantor 집합에 속하기 위해서는 삼진법 소수로 썼을 때 절대로 1이 등장해서는 안된다는 결론에 이른다. 그리고 여기서 다음 정리가 나온다.

Theorem 0.106 Cantor 집합은 1을 사용하지 않고 삼진법 소수로 쓸 수 있는 $[0,1]$에 속하는 모든 실수의 집합이다.

Proof 표기의 편의를 위해 집합열 $\{C_i\}$를 정의 0. 102에서와 같이 두고 1을 사용하지 않고 삼진법 소수로 쓸 수 있는 $[0,1]$에 속하는 모든 실수의 집합을 A라 하여 $A \subseteq C$임을 먼저 보이자. 이를 위해서는 임의의 $i \in \mathbb{N}$에 대해 $A \subseteq C_i$임을 수학적 귀납법으로 보이면 충분하다. 우선 임의의 $x \in A$에 대해 $x_{(3)} = 0.x_1x_2\cdots$라 하면 x_1은 0이거나 1이므로 $x \le 0.0\overline{2}_{(3)} = 1/3$이거나 $x \ge 0.2_{(3)} = 2/3$에서 $x \in C_1$이 되어 $A \subseteq C_1$임을 안다. 이어서 귀납가정으로서 $i \in \mathbb{N}$에 대해 $A \subseteq C_i$가 성립한다고 가정하고 이번에도 임의의 $x \in A$에 대해 $x_{(3)} = 0.x_1x_2\cdots$라 하면 모든 x_j는 0이거나 1이다. 만약 $x_1 = 0$이면 $3x_{(3)} = 0.x_2x_3\cdots \in A \subseteq C_k$이고 $x_1 = 2$이면 $3x_{(3)} - 2 = 0.x_2x_3\cdots \in A \subseteq C_i$이므로 어느 경우에나 $x \in C_i/3 \cup (C_i/3 + 2/3) = C_{i+1}$이 되어 $A \subseteq C_{i+1}$이고, 곧 $A \subseteq C$이다.

다음으로, $C \subseteq A$임을 보이기 위해서는 명제 $P(i)$를 'C_i에 속하는 임의의 실수는 소수점 아래 i번째까지 1을 사용하지 않고 삼진법 소수로 쓸 수 있다' 로 두고 수학적 귀납법을 사용하면 된다. 우선 $C_1 = [0,1/3] \cup [2/3,1]$인데, 만약 $x < 1/3$이거나 $x > 2/3$이면 x의 삼진법 소수 표현에서 첫 번째 자리는 각각 $0, 2$이고 $x = 1/3 = 0.0\overline{2}_{(3)}, x = 2/3 = 0.2_{(3)}$으로 쓸 수 있으므로 $P(1)$은 자명하다. 이제 귀납가정으로서 $i \in \mathbb{N}$에 대해 $P(i)$가 성립한다고 가정하고 임의의 $x \in C_{i+1} \subseteq C_i$를 택하자. 그렇다면 x는 적어도 삼진법 소수 표현에서 소수점 아래 i번째까지는 1을 사용하지 않고 쓸 수 있으므로 0혹은 2인 $x_1, \cdots, x_i$에 대해 $x_{(3)} = 0.x_1\cdots x_i\cdots$로 쓸 수 있다. 한편, 정의로부터 $3x \in C_i$이거나 $3x - 2 \in C_i$인데, 만약 $3x \in C_i$이면 $x_1 = 0$이어야 하고, 곧 $3x = 0.x_2\cdots x_i\cdots \in C_i$에서 $3x$도 삼진법 소수 표현에서 소수점 아래 i번째 자리까지는 1을 사용하지 않고 0혹은 2인 $y_1, \cdots, y_i$에 대해 $3x = 0.y_1\cdots y_i\cdots$로 쓸 수 있다. 비슷하게, $3x - 2 \in C_i$인 경우에도 0혹은 2인 $y_1, \cdots, y_i$에 대해 $3x - 2 = 0.y_1\cdots y_i\cdots$로 쓸 수 있음을 쉽게 알 수 있다. 그렇다면 $x_{(3)} = 0.0y_1\cdots y_i\cdots$이거나 $x_{(3)} = 0.2y_1\cdots y_i\cdots$이므로 $P(i+1)$이 성립하여, 곧 $C \subseteq A$이다. 증명은 이로써 충분하다. □

위의 정리에서는 표현에 주의해야 한다. 삼진법 무한소수 표현이 유일한 것과 달리 일반적으로 삼진법 소수 표현 자체는 유일하지 않다. 예컨대 $1/3$의 경우 간단히 유한소수로

$0.1_{(3)}$ 으로 쓸 수도 있지만 이를 $0.0\overline{2}_{(3)}$ 과 같이 무한소수로 쓸 수도 있다. 그리고 위의 정리에 따르면 비록 $1/3 = 0.1_{(3)}$ 이지만 이는 $0.0\overline{2}_{(3)}$ 으로 1을 사용하지 않고 삼진법 소수로 쓸 수 있으므로 Cantor 집합에 속한다. 한편, 5/9의 경우 $0.12_{(3)}$ 나 $0.11\overline{2}_{(3)}$ 으로 쓸 수 있는데, 어느 경우에든 1이 등장하므로 Cantor 집합에 속하지 못한다. 한편, 위의 정리에서 '삼진법 무한소수 표현에서 1이 등장하지 않는' 과 같이 깔끔한 기술 대신 '1을 사용하지 않고 삼진법 소수로 쓸 수 있는' 으로 말장난을 칠 수 밖에 없는 것은 Cantor 집합을 구성하는 과정에서 중간 1/3을 제거할 때에 이를 닫힌구간이 아닌 열린구간으로 제거하는 데 크게 기인한다.

이제 Cantor 집합의 성질들을 간단히 살펴보자. 전술했다시피, 그 성질 중에는 당연하게 느껴지는 것들도 있고 전혀 당연하지 않게 생겨먹은 것들도 있다.

Theorem 0.107 다음이 성립한다.

i. Cantor 집합은 비어있지 않다.
ii. Cantor 집합은 compact하다.
iii. Cantor 집합은 totally disconnected이다. 즉, 두 개 이상의 원소를 가지는 연결 부분집합 $A \subseteq C$가 존재하지 않는다.
iv. Cantor 집합은 perfect하다. 즉, Cantor 집합에는 고립점이 존재하지 않는다.
v. Cantor 집합은 nowhere dense하다. 즉, Cantor 집합의 내부는 공집합이다.
vi. $|C| = \mathfrak{c}$.
vii. Cantor 집합은 Borel 영집합이다.

Proof i. 정리 0. 106로부터 $2/3 = 0.2_{(3)} \in C$이므로 이는 자명하다.

ii. 집합열 $\{C_i\}$를 정의 0. 102에서와 같이 두면 정리 0. 103로부터 이가 감소하는 compact한 집합열임이 분명하므로 C도 닫혀있고, C가 $[0,1]$에 의해 유계임은 자명하므로 곧 C는 compact하다.

iii. 모순을 유도하기 위해 두 개 이상의 원소를 가지는 연결 부분집합 $A \subseteq C$가 존재한다고 가정하고 서로다른 $x, y \in A$를 택하여 WLOG, $x < y$라 하면 $[x, y] \in A$이다. 한편, 정리 0. 106로부터 소수점 아래 1이 나타나지 않도록 x, y를 삼진법 소수로 $x_{(3)} = 0.x_1\cdots$, $y_{(3)} = 0.y_1\cdots$와 같이 표현할 수 있다. (유한소수인 경우 뒤에 0이 계속 붙은 것으로 생각한다.) 여기서 x, y가 소수점 아래 k번째까지 이러한 소수 표현이 일치한다고 하면 $x_{k+1} = 0, y_{k+2} = 2$이어야 하는데, $z = 0.x_1\cdots x_k 11_{(3)}$로 두면 $z \in [x, y] \in A \subseteq C$이지만 $z \notin C$의 모순이 발생한다. 따라서 Cantor 집합은 totally disconnected이다.

iv. 임의의 $x \in C$에 대해 정리 0. 106로부터 소수점 아래 1이 나타나지 않도록 x를 삼진법 소수로 $x_{(3)} = 0.x_1x_2\cdots$와 같이 표현할 수 있다. 이제 각 $i \in \mathbb{N}$에 대해 $x_k = 0.x_1\cdots x_{i(3)}$라 하여 수열 $\{x_i\}$를 구성하면 이는 명백히 C에 속하는 수열로 $x_i \to x$이다. 따라서 Cantor 집합은 perfect하다.

v. 만약 Cantor 집합의 내부가 비어있지 않다면 적당한 $x \in C$를 하나 택하고 C에 속하는 그 근방을 찾을 수 있다. 그러나 이는 Cantor 집합은 totally disconnected하다는 iii의 결과와 모순되므로 $C^\circ = \emptyset$이고, 곧 Cantor 집합은 nowhere dense하다.

vi. 우선 $C \subseteq [0,1]$에서 $|C| \le |[0,1]| = \mathfrak{c}$임은 자명하므로 $|C| \ge \mathfrak{c}$라는 사실만 보이면 충분하다. 이를 위해 임의의 $x \in [0,1]$의 이진 무한소수 표현을 $x_{(2)} = 0.x_1x_2\cdots$라 하고 함수 $f : [0,1] \to C$를 $f : x \mapsto \sum_{i=1}^{\infty} 2\mathbf{1}_{\{1\}}(x_i)/3^i$로 두자. 이진 무한소수 표현은 유일하므로 함수 f는 well-define된다. 생각해보면, 임의의 $x \in [0,1]$에 대해 f는 $0.x_1x_2\cdots$이었던 x의 이진 소수 표현에서 1만 2로 바꾼 뒤, 이를 삼진 소수 표현으로 재해석하여 C의 한 원소로 mapping한다. 이러한 f는 명백히 단사이므로 $\mathfrak{c} = |[0,1]| \le |C|$가 되어 증명이 끝난다.

vii. ii로부터 Cantor 집합이 compact하므로 Borel임은 자명하다. 같은 이유로, 집합열 $\{C_i\}$를 정의 0.102에서와 같이 두면 각 C_i도 Borel이다. 이제 모든 $i \in \mathbb{N}$에 대해 $\mu_1(C_i) = 1 - \sum_{j=1}^{i} 2^{j-1}/3^j$임을 수학적 귀납법으로 보이자. 우선 $i = 1$인 경우에는 $C_1 = [0,1/3] \cup [2/3,1]$이므로 $\mu_1(C_1) = 2/3$에서 자명하므로 귀납가정으로서 $i \in \mathbb{N}$에 대해 $\mu_1(C_i) = 1 - \sum_{j=1}^{i} 2^{j-1}/3^j$라 가정하자. 그렇다면 $\mu_1(C_{i+1}) = \mu_1(C_i/3 \cup (C_i/3 + 2/3))$인데, $C_i \in [0,1]$에서 $C_i/3$과 $C_i/3 + 2/3$은 서로소이므로 $\mu_1(C_{i+1}) = \mu_1(C_i/3) + \mu(C_i/3 + 2/3) = 2\mu_1(C_i)/3 = 2/3 - \sum_{j=1}^{i} 2^j/3^{j+1} = 1 - \sum_{j=1}^{i+1} 2^{j-1}/3^j$가 되어 곧 모든 $i \in \mathbb{N}$에 대해 $\mu_1(C_i) = 1 - \sum_{j=1}^{i} 2^{j-1}/3^j$임을 안다. 이제 $\{C_i\}$가 감소하는 집합열임이 정리 0.103로부터 자명하므로 $\mu_1(C_i) \downarrow \mu_1(C) = 0$이다. □

Cantor 집합과 관련하여 우리가 알아볼 특수함수는 바로 Cantor 함수이다.

Definition 0.108 닫힌구간 $[0,1]$에 속하는 실수 x를 삼진 무한소수로 $0.x_1x_2\cdots$와 같이 쓸 때, N_x를 소수점 아래에서 처음으로 1이 등장하는 위치라 하자. 만약 그 삼진 무한소수 표현에서 1이 등장하지 않으면 $N_x = \infty$라 한다. 이제 다음과 같이 정의되는 함수 $c : [0,1] \to \mathbb{R}$를 **Cantor 함수 (- function)**, **Lebesgue의 특이함수 (- singular function)** 혹은 **악마의 계단 (devil's staircase)**이라 한다.

$$c : x \mapsto \sum_{i=1}^{N_x - 1} \frac{x_i}{2^{i+1}} + \frac{1}{2^{N_x}}$$

Cantor 함수가 삼진 무한소수 표기로써 정의되는 점에서 Cantor 집합과 관련이 있으리라 추측 정도는 할 수 있지만 사실 정의만 보고 이가 Cantor 집합과 정확히 어떻게 관련된 것인지를 알아치리기는 쉽지 않다. Cantor 집합과 Cantor 함수의 정확한 관계를 파악하기 위해서는 Cantor 함수의 성질을 살펴보아야 한다. 한편, 위의 그래프에서 볼 수 있듯이 Cantor 함수도 Cantor 집합과 같이 fractal의 모습을 띄며, 그래프의 모양으로 미루어보건대 뭔가 범상치 않은 면모를 가지리라 짐작할 수 있다. 실제로 Cantor 함수는 Cantor 집합처럼 해석학에서 반례를 만들어내는 데 유용하게 쓰이며, 대표적으로 일반적인 조건에서의

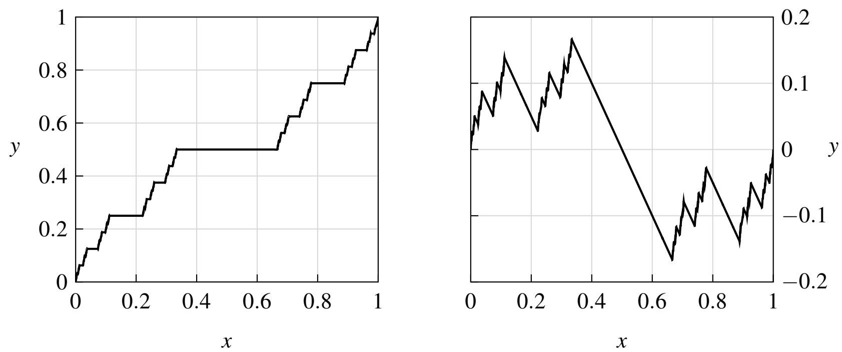


Figure 0.25 Cantor 함수 (왼쪽) 와 함수 $x \mapsto c(x) - x$ (오른쪽) 의 그래프. 단순 증가 성분을 뺀 오른쪽 그래프를 보면 Cantor 함수의 자기 유사성을 보다 뚜렷하게 느낄 수 있다. Computed by Wolfram MATHEMATICA.

FTC의 반례가 되어 이후 1장의 측도론에서 우리를 여러모로 골치아프게 할 것이다. 이에 대해서는 1장에서 하나의 절을 할애하여 소상히 논할 예정이니, 여기서는 잠시 넘어가도록 하고 우리는 Cantor 함수의 기본적인 성질들을 살펴보자.

Theorem 0.109 다음이 성립한다.

i. $0 \le c \le 1$.
ii. Cantor 함수는 증가한다.
iii. 집합 $[0,1] \setminus C$의 임의의 연결성분 I에 대해 I에서 Cantor 함수는 상수함수이다.
iv. $c(C) = [0,1]$.
v. Cantor 함수는 균등연속이지만 절대연속이 아니다. [9]
vi. Cantor 함수는 집합 $[0,1] \setminus C$에서 $\mathscr{C}^\infty$ 급이며 $n \in \mathbb{N}$에 대해 $(c|_{[0,1]\setminus C})^{(n)} = 0$이지만 임의의 $x \in C$에서는 미분불가능하다.

Proof 편의를 위해 $[0,1]$에 속하는 실수 x를 삼진 무한소수로 $0.x_1x_2\cdots$와 같이 쓰고, N_x를 소수점 아래에서 처음으로 1이 등장하는 위치라 하자.

i. 모순을 유도하기 위해 $c(x) > 1$인 $x \in [0,1]$가 존재한다고 하고 $N_x < \infty$인 경우를 먼저 생각하자. 그렇다면 $c(x) = \sum_{i=1}^{N_x-1} x_i/2^{i+1} + 1/2^{N_x} > 1$인데, 이는 $1 - 1/2^{N_x} < \sum_{i=1}^{N_x-1} x_i/2^{i+1} \le \sum_{i=1}^{N_x-1} 1/2^i = 1 - 1/2^{N_x}$의 모순을 함의하므로 이 경우는 불가능하다. 한편, $N_x = \infty$인 경우에도 비슷한 모순이 발생하므로 곧 $c \le 1$임을 안다. 이제 c가 음이 아님은 자명하므로 $0 \le c \le 1$이다.

ii. 임의의 $x, y \in [0,1]$를 택하여 $x < y$라 하자. 증명은 경우를 나누어 진행된다. 먼저 살펴볼 것은 $N_x = N_y =: N$인 경우이다. 이때, x, y의 삼진 무한소수 표현에서 처음으로 서로 다른 수가 나타나는 위치를 M이라 하면 $M = N$일 수는 없으므로 $M > N$이거나 $M < N$이다.

전자의 경우에는 정의로부터 $c(x) = c(y)$ 임이 분명하고, 후자의 경우에는 $x_M = 0, y_M = 2$ 이어야 하므로

$$\begin{aligned}
c(y) - c(x) &= \sum_{i=1}^{N-1} \frac{y_i - x_i}{2^{i+1}} \\
&= \sum_{i=M+1}^{N-1} \frac{y_i - x_i}{2^{i+1}} + \frac{1}{2^M} \\
&\geq \frac{1}{2^M} - \sum_{i=M+1}^{N-1} \frac{1}{2^i} \\
&\geq \frac{1}{2^M} - \sum_{i=M+1}^{\infty} \frac{1}{2^i} \\
&= 0
\end{aligned}$$

에서 $c(x) \leq c(y)$ 가 되어 어느 경우에나 $N_x = N_y$ 이면 $c(x) \leq c(y)$ 이다.

다음으로 살펴볼 것은 $N_x > N_y$ 인 경우인데, M 을 방금과 같이 두면 $M > N_y$ 일 수는 없으므로 $M < N_y$ 이거나 $M = N_y$ 이다. 전자의 경우에는 $x_M = 0, y_M = 2$ 이어야 하므로

$$\begin{aligned}
c(y) - c(x) &= \sum_{i=1}^{N_y-1} \frac{y_i}{2^{i+1}} - \sum_{i=1}^{N_x-1} \frac{x_i}{2^{i+1}} + \frac{1}{2^{N_y}} - \frac{1}{2^{N_x}} \\
&= \sum_{i=M+1}^{N_y-1} \frac{y_i}{2^{i+1}} - \sum_{i=M+1}^{N_x-1} \frac{x_i}{2^{i+1}} + \frac{1}{2^M} + \frac{1}{2^{N_y}} - \frac{1}{2^{N_x}} \\
&= \sum_{i=M+1}^{N_y-1} \frac{y_i - x_i}{2^{i+1}} - \sum_{i=N_y}^{N_x-1} \frac{x_i}{2^{i+1}} + \frac{1}{2^M} + \frac{1}{2^{N_y}} - \frac{1}{2^{N_x}} \\
&\geq \frac{1}{2^M} - \sum_{i=M+1}^{N_y-1} \frac{1}{2^i} - \sum_{i=N_y}^{N_x-1} \frac{1}{2^i} \\
&= \frac{1}{2^M} - \sum_{i=M+1}^{N_x-1} \frac{1}{2^i} + \frac{1}{2^M} \\
&\geq \frac{1}{2^M} - \sum_{i=M+1}^{\infty} \frac{1}{2^i} \\
&= 0
\end{aligned}$$

에서 $c(x) \leq c(y)$ 이고 후자의 경우에는 $x_M = 0, y_M = 1$ 이어야 하므로

$$c(y) - c(x) = \sum_{i=1}^{N_y-1} \frac{y_i}{2^{i+1}} - \sum_{i=1}^{N_x-1} \frac{x_i}{2^{i+1}} + \frac{1}{2^{N_y}} - \frac{1}{2^{N_x}}$$

$$
\begin{aligned}
&= \frac{1}{2^{N_y}} - \frac{1}{2^{N_x}} - \sum_{i=N_y+1}^{N_x-1} \frac{x_i}{2^{i+1}} \\
&\geq \frac{1}{2^{N_y}} - \frac{1}{2^{N_x}} - \sum_{i=N_y+1}^{N_x-1} \frac{1}{2^i} \\
&= \frac{1}{2^{N_y}} - \frac{1}{2^{N_x}} - \frac{1}{2^{N_y}} \left(1 - \frac{1}{2^{N_x-N_y-1}}\right) \\
&= \frac{1}{2^{N_x-1}} - \frac{1}{2^{N_x}} \\
&\geq 0
\end{aligned}
$$

에서 $c(x) \leq c(y)$가 되어 어느 경우에나 $N_x > N_y$이면 $c(x) \leq c(y)$이다. 이와 비슷하게 $N_x < N_y$인 경우에도 $c(x) \leq c(y)$임을 쉽게 보일 수 있으므로 곧 어느 경우에나 $c(x) \leq c(y)$가 되어 c가 증가함을 안다.

iii. 명제 0.105의 iii으로부터 $[0,1] \setminus C$의 연결성분은 열린구간 $I \subseteq [0,1]$이며 적당한 $n, i \in \mathbb{N}$에 대해 $I = (n/3^i, (n+1)/3^i)$로 쓸 수 있음을 알 수 있다. 이제 서로다른 임의의 $x, y \in I$에 대해 WLOG, $x < y$라 하고 $c(x) = c(y)$임을 보이기만 하면 된다. 나아가, M을 ii의 증명에서와 같이 둔다면 이는 $N_x = N_y \leq M$임을 보이는 것으로 충분하다. 이를 위해 먼저 $N_x < N_y$라 가정하고 $z = 0.x_1 \cdots x_{N_x-1}2_{(3)}$를 생각하면 정리 0.106으로부터 $z \in [x, y] \subseteq I \subseteq [0,1] \setminus C$이지만[10] $z \in C$인 모순이 발생한다. 비슷한 방법으로 $N_x > N_y$의 경우도 불가능함을 보일 수 있으므로 $N_x = N_y =: N$이다. 증명을 끝내기 위해 $N > M$이라 가정하고 $z = 0.x_1 \cdots x_{M-1}0\overline{2}_{(3)} = 0.x_1 \cdots x_{M-1}1_{(3)}$를 생각하면 이번에도 방금과 같은 모순이 발생하므로 $N \leq M$이 되어 $c(x) = c(y)$임을 안다.

iv. 임의의 $x \in [0,1]$를 택하고 이의 이진 무한소수 전개를 $0.x_1x_2\cdots$라 하여 $y = \sum_{i=1}^{\infty} 2\mathbf{1}_{\{1\}}(x_i)/3^i$를 생각하면 이는 정리 0.106으로부터 명백히 C에 속하고 $c(y) = x$이므로 $c(C) = [0,1]$이다.

v. iii으로부터 $[0,1] \setminus C$에서 c가 연속임은 자명하므로 임의의 $x \in C$를 고정하고 x에서 c가 연속임을 보이면 c가 연속임을 보이기에 충분하다. 이를 위해 임의의 $\varepsilon > 0$을 택하면 $2^{-k+2} < \varepsilon$인 충분히 큰 $k \in \mathbb{N}$가 존재하므로 $|x - y| < 3^{-k}$인 $y \in [0,1]$를 택할 수 있다. 간결한 논의를 위해 $x \leq y$인 경우만 생각하기로 하여 M을 ii의 증명에서와 같이 두고 경우를 나누어 생각한다. ($x > y$인 경우에도 비슷하게 같은 결론을 얻을 수 있다.) 먼저 $N_x = N_y =: N$인 경우에는 $M > N$이라면 $c(x) = c(y)$가 되어 $|c(x) - c(y)| < \varepsilon$임이 자명하다. 한편, $M < N$인 경우에는 $x_M = 0, y_M = 2$인데, 만약 $M < k$라면 $y - x \geq 2/3^M + \sum_{i=M+1}^{\infty}(y_i - x_i)/3^i \geq 2/3^M - \sum_{i=M+1}^{\infty} 2/3^i = 1/3^M > 1/3^k$의 모순이 발생하므로 $M \geq k$이다. 이로부터

$$
c(y) - c(x) = \sum_{i=1}^{N-1} \frac{y_i - x_i}{2^{i+1}}
$$

$$
\begin{aligned}
&= \sum_{i=M+1}^{N-1} \frac{y_i - x_i}{2^{i+1}} + \frac{1}{2^M} \\
&\le \sum_{i=M+1}^{N-1} \frac{1}{2^i} + \frac{1}{2^M} \\
&\le \sum_{i=M+1}^{\infty} \frac{1}{2^i} + \frac{1}{2^M} \\
&= \frac{1}{2^{M-1}} \\
&\le \frac{1}{2^{k-1}} \\
&< \varepsilon
\end{aligned}
$$

이 되어 ii로부터 어느 경우에나 $N_x = N_y$ 이면 $|c(x) - c(y)| < \varepsilon$ 이다.

다음으로 $N_x > N_y$ 인 경우를 보면 $M < N_y$ 라면 $x_M = 0, y_M = 2$ 이고 방금과 비슷한 이유로 $M \ge k$ 이므로

$$
\begin{aligned}
c(y) - c(x) &= \sum_{i=1}^{N_y-1} \frac{y_i}{2^{i+1}} - \sum_{i=1}^{N_x-1} \frac{x_i}{2^{i+1}} + \frac{1}{2^{N_y}} - \frac{1}{2^{N_x}} \\
&= \sum_{i=M+1}^{N_y-1} \frac{y_i}{2^{i+1}} - \sum_{i=M+1}^{N_x-1} \frac{x_i}{2^{i+1}} + \frac{1}{2^M} + \frac{1}{2^{N_y}} - \frac{1}{2^{N_x}} \\
&= \sum_{i=M+1}^{N_y-1} \frac{y_i - x_i}{2^{i+1}} - \sum_{i=N_y}^{N_x-1} \frac{x_i}{2^{i+1}} + \frac{1}{2^M} + \frac{1}{2^{N_y}} - \frac{1}{2^{N_x}} \\
&\le \sum_{i=M+1}^{N_y-1} \frac{1}{2^i} + \frac{1}{2^M} + \frac{1}{2^{N_y}} \\
&\le \sum_{i=M+1}^{\infty} \frac{1}{2^i} + \frac{1}{2^M} + \frac{1}{2^{N_y}} \\
&= \frac{1}{2^{M-1}} + \frac{1}{2^{N_y}} \\
&< \frac{1}{2^{M-2}} \\
&\le \frac{1}{2^{k-2}} \\
&< \varepsilon
\end{aligned}
$$

에서 ii로부터 $|c(x) - c(y)| < \varepsilon$ 이다. 한편, $M = N_y$ 라면 $x_M = 0, y_M = 1$ 인데, 이 경우는 조금 복잡하다. 우선 소수점 아래 M 번째 자리에서 처음으로 다른 수가 나타난 이후 x 의 무한소수 전개에서 처음으로 2가 아닌 수가 나타나는 위치를 N'_x 라 하고 $N'_x < k$ 라 하면

$y-x=1/3^M+\sum_{i=M+1}^{\infty}(y_i-x_i)/3^i \geq 1/3^M-\sum_{i=M+1}^{N_x'-1}2/3^i-\sum_{i=N_x'+1}^{\infty}2/3^i \geq 1/3^{N_x'} > 1/3^k$의 모순이 발생하므로 $N_x' \geq k$이다. 나아가, $N_x' \leq N_x$이고 $x_{M+1}=\cdots x_{N_x'-1}=2$임은 분명하므로 이상으로부터

$$\begin{aligned}
c(y)-c(x) &= \sum_{i=1}^{N_y-1}\frac{y_i}{2^{i+1}}-\sum_{i=1}^{N_x-1}\frac{x_i}{2^{i+1}}+\frac{1}{2^{N_y}}-\frac{1}{2^{N_x}} \\
&= \frac{1}{2^{N_y}}-\frac{1}{2^{N_x}}-\sum_{i=N_y+1}^{N_x'-1}\frac{1}{2^i}-\sum_{i=N_x'}^{N_x-1}\frac{x_i}{2^{i+1}} \\
&\leq \frac{1}{2^{N_y}}-\sum_{i=N_y+1}^{N_x'-1}\frac{1}{2^i} \\
&= \frac{1}{2^{N_x'-1}} \\
&\leq \frac{1}{2^{k-1}} \\
&< \varepsilon
\end{aligned}$$

에서 ii로부터 $|c(x)-c(y)|<\varepsilon$이 되어 곧 어느 경우에나 $N_x > N_y$이면 $|c(x)-c(y)|<\varepsilon$이다. 이제 이와 비슷하게 $N_x < N_y$인 경우에도 $|c(x)-c(y)|<\varepsilon$임을 보일 수 있으므로 c가 $x \in C$에서도 연속임을 안다.

이제 남은 부분은 어렵지 않다. 우선 연속인 c가 정의된 $[0,1]$이 compact하므로 균등연속임이 분명하다. 마지막으로 모순을 유도하기 위해 c가 절대연속이라고 가정하고 집합열 $\{C_i\}$를 정의 0.102에서와 같이 두면 적당한 $\delta>0$가 존재하여 임의의 서로소인 반열린구간 $I_1=[a_1,b_1),\cdots,I_k=[a_k,b_k)\subseteq\mathbb{R}$에 대해 $\sum_{i=1}^{k}(b_i-a_i)<\delta$이면 $\sum_{i=1}^{k}[c(b_i)-c(a_i)]<1$이다. 이에 대한 반례를 찾기 위해 $(2/3)^k<\delta$를 만족하는 충분히 큰 $k\in\mathbb{N}$를 택하여 집합 C_k를 생각하면 명제 0.105의 ii로부터 이는 2^k개의 서로소인 닫힌구간 $J_1,\cdots,J_{2^k}$의 합집합이며, 적당한 $n_1<\cdots<n_{2^k}\in\mathbb{N}_0$에 대해 각 J_i는 $[n_i/3^k,(n_i+1)/3^k]$로 쓸 수 있다. 그리고 정리 0.103로부터 $n_1=0, n_{2^k}=3^k-1$임이 분명하다. 그렇다면 자연스럽게 각 $i<2^k$에 대해 열린구간 $((n_i+1)/3^k, n_{i+1}/3^k)$은 서로소이고 이들의 합집합이 $[0,1]\setminus C_k$가 되어 iii과 방금 보인 c의 연속성으로부터 $c((n_i+1)/3^k)=c(n_{i+1}/3^k)$이다. 이제 각 $i\leq 2^k$에 대해 $I_i=[n_i/3^k,(n_i+1)/3^k)$로 두면 $\sum_{i=1}^{2^k}1/3^k=(2/3)^k<\delta$이지만 $\sum_{i=1}^{2^k}[c((n_i+1)/3^k)-c(n_i/3^k)]=c(1)-c(0)=1$이 되어 모순이 발생한다. 따라서 c는 절대연속이 아니다.

vi. Cantor 함수가 $[0,1]\setminus C$에서 $\mathscr{C}^\infty$이고 $(c|_{[0,1]\setminus C})^{(n)}=0$임은 ii로부터 자명하므로 임의의 $x\in C$를 고정하고 x에서 c가 미분가능하지 않다는 것만 보이면 된다. 이번에도 모순을 유도하기 위해 x에서 c가 미분가능하다고 하면 적당한 $\delta>0$가 존재하여 $|x-y|<\delta$이지만 $y\neq x$인 $y\in[0,1]$에 대해 $|c(x)-c(y)|/|x-y|<1$이다. 이제 경우를 나누어 생각한다. 먼저 x의 삼진 무한소수 표현에서 1이 나타나지 않아 $N_x=\infty$인 경우를 생각해보면, 그 소수

표현에는 2가 무한히 많이 나타나는데, $1/3^k < \delta$를 만족하는 충분히 큰 $k \in \mathbb{N}$를 택하여 x의 소수 표현에서 k번째로 나타나는 2를 0으로 바꾸어 얻는 실수를 $y \in [0,1]$라 하여 M을 ii의 증명에서와 같이 두면 $M \geq k$에서 $|x-y| = 2/3^M \leq 2/3^k < \delta$이고 $x \neq y$이지만 $|c(x)-c(y)|/|x-y| = (1/2^{M+1})/(2/3^M) = (3/2)^M \geq (3/2)^k > 1$의 모순이 발생한다.

다음으로 x의 삼진 무한소수 표현에서 1이 나타나는 경우를 생각해보면 정리 0.106의 결과를 고려했을 때 $x_{(3)} = 0.x_1 \cdots x_{N_x}\overline{2}$의 형태가 되어야 함을 알 수 있다. 이번에도 $1/3^k < \delta$를 만족하는 충분히 큰 $k \in \mathbb{N}$를 택하여 x의 무한소수 표현에서 1인 N_k번째 수를 2로 바꾸고 2인 $N_x+1, \cdots, N_x+k$번째 수를 모두 0으로 바꾸어 얻는 실수를 $y \in [0,1]$라 하면 $|x-y| = |\sum_{i=1}^{k} 2/3^{N_x+k} - 1/3^{N_x}| = 1/3^{N_x+k} \leq 1/3^k < \delta$이지만

$$\begin{aligned}\left|\frac{c(x)-c(y)}{x-y}\right| &= 3^{N_x+k}\left|\sum_{i=1}^{N_x-1}\frac{x_i}{2^{i+1}} + \frac{1}{2^{N_x}} - \sum_{i=1}^{\infty}\frac{y_i}{2^{i+1}}\right| \\ &= 3^{N_x+k}\sum_{i=N_x+k+1}^{\infty}\frac{1}{2^i} \\ &= \frac{3^{N_x+k}}{2^{N_x+k}} \\ &> 1\end{aligned}$$

의 모순이 발생한다. 곧 어느 경우에나 모순이 발생하므로 임의의 $x \in C$에서는 c가 미분불가능함을 안다. □

0.2.5 Special Functions Related to Sine

이어서 우리에게 친숙한 사인 함수와 관련된 특수함수를 조금 살펴보자. 물론, 사인 함수를 다루는 것이 어렵지는 않지만 그렇다고 해서 이와 관련된 특수함수가 모두 다루기 쉬운 것은 아니니 조금은 긴장할 필요가 있다. 처음으로 살펴볼 특수함수는 고등학교 문제집에서 한두 번쯤 본 적이 있을 수도 있다.

Definition 0.110 다음과 같이 정의되는 함수 $\mathrm{sinc} : \mathbb{R} \to \mathbb{R}$를 **sinc 함수 (− function)**라 한다.

$$\mathrm{sinc} : x \mapsto \begin{cases} \sin x/x & x \neq 0 \text{인 경우} \\ 1 & \text{ow.} \end{cases}$$

Theorem 0.111 다음이 성립한다.

i. sinc 함수는 우함수이다.

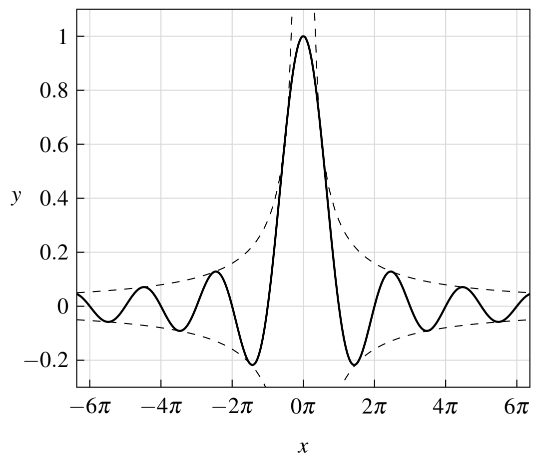


Figure 0.26 sinc 함수(실선)와 함수 $x \mapsto \pm 1/x$(점선)의 그래프. Computed by Wolfram MATHEMATICA.

ii. sinc 함수는 각 $k \in \mathbb{N}$에 대해 $\pm k\pi$에서만 영점을 가진다.
iii. sinc 함수는 0에서 최댓값 1을 가지고, 적당한 $x_* \in (\pi, 3\pi/2)$가 존재하여 $\pm x_*$에서 최솟값을 가진다.
iv. $\lim_{x\to-\infty} \operatorname{sinc}(x) = \lim_{x\to\infty} \operatorname{sinc}(x) = 0$.
v. sinc 함수는 $\mathscr{C}^\infty$급이다.

PROOF i, ii, iv. 이는 sinc 함수의 정의로부터 자명하다.

v. 이는 $x \in \mathbb{R}$에 대해 $\operatorname{sinc}(x) = \sum_{i=0}^{\infty} (-1)^i x^{2i}/(2i+1)!$이라는 점에서 자명하다.

iii. 우선 임의의 $x \in \mathbb{R}$에 대해 $x \neq 0$이면 $\sin x < |x|$이므로 $\sin x$가 0에서 최댓값으로 1을 가지는 것은 자명하다. 한편,

$$\operatorname{sinc}' : x \mapsto \begin{cases} \dfrac{x\cos x - \sin x}{x^2} & x \neq 0\text{인 경우} \\ 0 & \text{ow.} \end{cases}$$

이므로 $\mathbb{R}^+$에서 방정식 $x = \tan x$의 근을 $x_1 < x_2 < \cdots$라 하면 $\{x_{2i-1}\}$는 $\mathbb{R}^+$에서 sinc 함수가 극소가 되는 모든 점들로 구성된 수열이다. 여기서 각 $i \in \mathbb{N}$에 대해 탄젠트의 주기성으로부터 $(2i-1)\pi < x_i < (4i-1)\pi/2$이고 $\tan x_i = x_i < x_{i+1} < \tan x_{i+1}$에서 $\pi < x_i - 2i\pi < x_{i+1} - 2(i+1)\pi < 3\pi/2$이므로 $\operatorname{sinc}(x_i) = \sin x_i/x_i = \cos x_i < \cos x_{i+1} = \sin x_{i+1}/x_{i+1} = \operatorname{sinc}(x_{i+1})$이다. 이로부터 sinc 함수가 우함수라는 점과 0에서 최댓값을 가진다는 점을 생각하면 이가 $x_1 \in (\pi, 3\pi/2)$에서 최솟값을 가짐을 안다. □

Euler는 위의 정리의 ii에 착안하여 $\operatorname{sinc}(x) = \prod_{i=1}^{\infty}(1 - x^2/\pi^2 i^2)$로 sinc 함수를 ‘인수분해’ 함으로써 앞서 살펴보았던 Basel 문제를 해결하였다. 물론, 이는 오늘날의 기준에서

보면 조금 억지이다. sinc 함수가 다항식이면 모르겠지만, 다항식도 아닌 함수가 단순히 x_0를 영점으로 가진다고 해서 $x-x_0$를 인자로 가진다고 할 수는 없기 때문이다. 그럼에도 불구하고, 위의 식이 성립하기는 한다.

Lemma 0.112 임의의 $x_1, \cdots, x_k \in [0,1]$에 대해 $\prod_{i=1}^{k}(1-x_i) \geq 1-\sum_{i=1}^{k} x_i$이다.

Proof 수학적 귀납법을 사용하자. 우선 $k=1$인 경우에는 보조정리가 자명하므로 귀납가정으로서 $k \in \mathbb{N}$에 대해 보조정리가 성립한다고 가정하자. 그렇다면 임의의 $x_1, \cdots, x_{k+1} \in [0,1]$에 대해 $\prod_{i=1}^{k+1}(1-x_i) \geq (1-x_{k+1})(1-\sum_{i=1}^{k} x_i) = 1-\sum_{i=1}^{k+1} x_i + x_{k+1}\sum_{i=1}^{k} x_i \geq 1-\sum_{i=1}^{k+1} x_i$가 되어 증명이 끝난다. □

Theorem 0.113 임의의 $x \in \mathbb{R}$에 대해 $\operatorname{sinc}(x) = \prod_{i=1}^{\infty}(1-x^2/\pi^2 i^2)$이다. [11]

Proof 먼저 임의의 $k \in \mathbb{N}$와 임의의 $x \in \mathbb{R}$를 택하면 임의의 $\theta, \varphi \in \mathbb{R}$에 대해 $\sin\theta = 2\sin(\theta/2)\cos(\theta/2) = 2\sin(\theta/2)\sin((\theta+\pi)/2)$이고 $\sin(\theta+\varphi)\sin(\theta-\varphi) = (\cos 2\varphi - \cos 2\theta)/2 = \sin^2\theta - \sin^2\varphi$이므로

$$\begin{aligned}
\sin x &= 2\sin\frac{x}{2}\sin\left(\frac{x+\pi}{2}\right) \\
&= 8\sin\frac{x}{4}\sin\left(\frac{x+\pi}{4}\right)\sin\left(\frac{x+2\pi}{4}\right)\sin\left(\frac{x+3\pi}{4}\right) \\
&= \cdots \\
&= 2^{2^k-1}\prod_{i=0}^{2^k-1}\sin\frac{x+i\pi}{2^k} \\
&= 2^{2^k-1}\sin\frac{x}{2^k}\sin\frac{x+2^{k-1}\pi}{2^k}\left(\prod_{i=1}^{2^{k-1}-1}\sin\frac{x+i\pi}{2^k}\right)\left[\prod_{i=1}^{2^{k-1}-1}\sin\frac{x+(2^k-i)\pi}{2^k}\right] \\
&= -2^{2^k-1}\sin\frac{x}{2^k}\cos\frac{x}{2^k}\prod_{i=1}^{2^{k-1}-1}\sin\left(\frac{x+i\pi}{2^k}\right)\sin\left(\frac{x-i\pi}{2^k}\right) \\
&= 2^{2^k-1}\sin\frac{x}{2^k}\cos\frac{x}{2^k}\prod_{i=1}^{2^{k-1}-1}\left(\sin^2\frac{i\pi}{2^k}-\sin^2\frac{x}{2^k}\right)
\end{aligned}$$

가 성립한다. 이로부터 0이 아닌 충분히 작은 x에 대해

$$\frac{\sin x}{\sin x/2^k} = 2^{2^k-1}\cos\frac{x}{2^k}\prod_{i=1}^{2^{k-1}-1}\left(\sin^2\frac{i\pi}{2^k}-\sin^2\frac{x}{2^k}\right)$$

가 되어 $x \to 0$의 극한을 취하면 $2^{2^k-1} = 2^k/\prod_{i=1}^{2^{k-1}-1}\sin^2 i\pi/2^k$임을 안다. 여기서 다시 0이 아닌 임의의 $x \in \mathbb{R}$를 고정하고 이상을 종합하면

$$\begin{aligned}\sin x &= \frac{2^k}{\prod_{i=1}^{2^{k-1}-1} \sin^2 i\pi/2^k} \sin\frac{x}{2^k}\cos\frac{x}{2^k}\prod_{i=1}^{2^{k-1}-1}\left(\sin^2\frac{i\pi}{2^k} - \sin^2\frac{x}{2^k}\right)\\ &= 2^k\sin\frac{x}{2^k}\cos\frac{x}{2^k}\prod_{i=1}^{2^{k-1}-1}\left(1-\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\right)\end{aligned}$$

를 얻는다. 이제 충분히 큰 $k, l \in \mathbb{N}$을 택하여 $l \le 2^{k-1}-1$이고 $(x/2l)^2 < 1$이라 하면 임의의 $\theta \in (0, \pi/2)$에 대해 $2\theta/\pi < \sin\theta < \theta$이므로 각 $l \le i \le 2^{k-1}-1$에 대해

$$\begin{aligned}1 &> 1-\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\\ &> 1-\left(\frac{x/2^k}{i/2^{k-1}}\right)^2\\ &= 1-\left(\frac{x}{2i}\right)^2\\ &\ge 1-\left(\frac{x}{2l}\right)^2\\ &> 0\end{aligned}$$

이 되어

$$\prod_{i=1}^{2^{k-1}-1}\left(1-\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\right) \le \prod_{i=1}^{l}\left(1-\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\right)$$

이다. 한편, 같은 이유에서 위의 보조정리로부터 k와 l이 충분히 크다면

$$\begin{aligned}\prod_{i=l+1}^{2^{k-1}-1}\left(1-\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\right) &\ge 1-\sum_{i=l+1}^{2^{k-1}-1}\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\\ &\ge 1-\sum_{i=l+1}^{2^{k-1}-1}\left(\frac{x/2^k}{i/2^{k-1}}\right)^2\\ &= 1-\frac{x^2}{4}\sum_{i=l+1}^{2^{k-1}-1}\frac{1}{i^2}\\ &\ge 1-\frac{x^2}{4}\sum_{i=l+1}^{\infty}\frac{1}{i^2}\\ &> 0\end{aligned}$$

이므로

$$\prod_{i=1}^{2^{k-1}-1}\left(1-\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\right) \geq \left(1-\frac{x^2}{4}\sum_{i=l+1}^{\infty}\frac{1}{i^2}\right)\prod_{i=1}^{l}\left(1-\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\right)$$

이다. 이상으로부터 충분히 큰 k, l에 대해 $l \leq 2^{k-1}-1$이면

$$\begin{aligned}\left(1-\frac{x^2}{4}\sum_{i=l+1}^{\infty}\frac{1}{i^2}\right)\prod_{i=1}^{l}&\left(1-\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\right)\\ &\leq \prod_{i=1}^{2^{k-1}-1}\left(1-\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\right) \leq \prod_{i=1}^{l}\left(1-\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\right)\end{aligned}$$

이고,

$$\begin{aligned}\lim_{k\to\infty}\prod_{i=1}^{2^{k-1}-1}\left(1-\frac{\sin^2 x/2^k}{\sin^2 i\pi/2^k}\right) &= \lim_{k\to\infty}\frac{\sin x}{2^k\sin(x/2^k)\cos(x/2^k)}\\ &= \frac{\sin x}{x}\end{aligned}$$

이므로 위의 부등식에서 각 변에 $k\to\infty$의 극한을 취하면

$$\left(1-\frac{x^2}{4}\sum_{i=l+1}^{\infty}\frac{1}{i^2}\right)\prod_{i=1}^{l}\left(1-\frac{x^2}{i^2\pi^2}\right) \leq \frac{\sin x}{x} \leq \prod_{i=1}^{l}\left(1-\frac{x^2}{i^2\pi^2}\right)$$

이다. 이는 곧

$$\left|\frac{\sin x}{x}-\prod_{i=1}^{l}\left(1-\frac{x^2}{i^2\pi^2}\right)\right| \leq \frac{x^2}{4}\left(\sum_{i=l+1}^{\infty}\frac{1}{i^2}\right)\prod_{i=1}^{l}\left|1-\frac{x^2}{i^2\pi^2}\right|$$

임을 의미하는데,

$$\begin{aligned}\prod_{i=1}^{l}\left|1-\frac{x^2}{i^2\pi^2}\right| &\leq \prod_{i=1}^{l}\left(1+\frac{x^2}{i^2\pi^2}\right)\\ &\leq \prod_{i=1}^{l}\exp\left(\frac{x^2}{i^2\pi^2}\right)\\ &= \exp\left(\frac{x^2}{\pi^2}\sum_{i=1}^{l}\frac{1}{i^2}\right)\\ &\leq \exp\left(\frac{x^2}{\pi^2}\sum_{i=1}^{\infty}\frac{1}{i^2}\right)\\ &= e^{x^2/6}\end{aligned}$$

이므로

$$\left|\frac{\sin x}{x} - \prod_{i=1}^{l}\left(1 - \frac{x^2}{i^2\pi^2}\right)\right| \le \frac{x^2}{4}\left(\sum_{i=l+1}^{\infty}\frac{1}{i^2}\right)e^{x^2/6}$$

이고, 양변에 $l \to \infty$의 극한을 취하면 $\sum_{i=l+1}^{\infty} 1/i^2 \to 0$에서 $\sin(x)/x = \prod_{i=1}^{\infty}(1 - x^2/\pi^2 i^2)$임을 안다. 한편, $x = 0$인 경우에는 정리가 자명하므로 증명은 이로써 충분하다. □

다음으로 살펴볼 함수는 아마 난생 처음 보는 함수일 것이다.

Definition 0.114 자연수 $k \in \mathbb{N}$에 대해 다음과 같이 정의된 함수 $D_k : \mathbb{R} \to \mathbb{R}$를 **Dirichlet kernel**이라 한다.

$$D_k : x \mapsto \begin{cases} \dfrac{\sin(2k+1)\pi x}{\sin \pi x} & x \notin \mathbb{Z}\text{인 경우} \\ 2k+1 & \text{ow.} \end{cases}$$

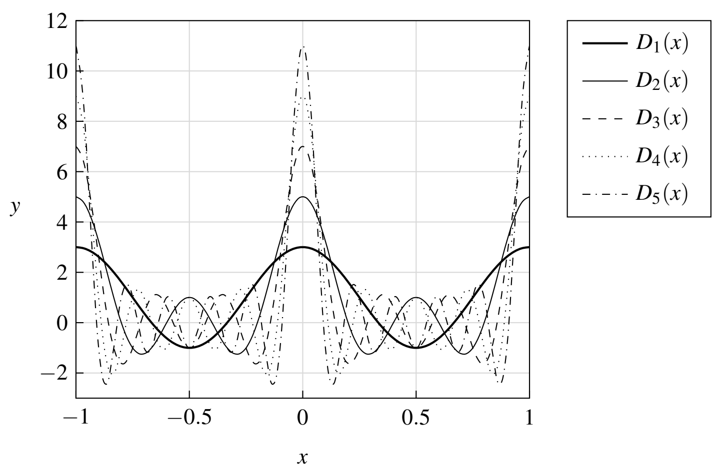


Figure 0.27 Dirichlet kernel의 그래프. Computed by Wolfram MATHEMATICA.

함수의 형태는 둘째치고, 함수가 아닌 kernel이라 명명된 것이 일단 가장 눈에 띌 것이다. 이는 이 함수의 주된 용도와 깊은 관련이 있다. 보통 대부분의 특수함수는 그 자체로 나름의 용도가 있거나 특별한 의미를 가져 독립적으로 사용된다. 그러나 어떤 특수함수는 독립적으로 사용되기보단 다른 함수와 함께 연산을 취하는 용도로 사용되는 경우가 있는데, 특히 다른 함수와 곱하여져 적분하는 용도로 널리 사용되는 함수를 kernel이라 부른다. 그렇다면 왜 Dirichlet kernel을 다른 함수와 곱하여 적분하는 용도로 사용하느냐? 이 질문에 대해 지금 깔끔하게 답하기는 어렵지만, 적당히 둘러대자면 Dirichlet kernel과 어떤 함수 f를 곱하여 적분하면 f의 '특성치'가 추출되어 나오기 때문이다. 이에 대해 보다 자세히 알기

위해서는 Fourier 해석이 필요하므로 일단은 이 정도의 답으로 만족하기로 하고, 우리는 Dirichlet kernel의 성질을 살펴보도록 하자.

Theorem 0.115 임의의 $k \in \mathbb{N}$와 임의의 $x \in \mathbb{R}$에 대해 $D_k(x) = \sum_{i=-k}^{k} e^{2\pi \mathbf{i} ix} = \sum_{i=-k}^{k} \cos 2\pi ix$이다.

Proof 만약 $x \in \mathbb{Z}$이면 정리가 자명하므로 $x \notin \mathbb{Z}$라 가정하자. 그렇다면

$$\begin{aligned}\sum_{i=-k}^{k} e^{2\pi \mathbf{i} ix} &= \frac{e^{-2\pi \mathbf{i} kx}[1 - e^{2\pi \mathbf{i}(2k+1)x}]}{1 - e^{2\pi \mathbf{i} x}} \\ &= \frac{e^{\pi \mathbf{i}(2k+1)x} - e^{-\pi \mathbf{i}(2k+1)x}}{e^{\pi \mathbf{i} x} - e^{-\pi \mathbf{i} x}} \\ &= \frac{\sin(2k+1)\pi x}{\sin \pi x} \\ &= D_k(x)\end{aligned}$$

에서 어느 경우에나 $D_k(x) = \sum_{i=-k}^{k} e^{2\pi \mathbf{i} ix}$이다. 이제 Dirichlet kernel이 실함수라는 점을 생각하면 증명이 끝난다. □

Theorem 0.116 임의의 $k \in \mathbb{N}$에 대해 다음이 성립한다.

i. Dirichlet kernel은 우함수이다.
ii. Dirichlet kernel은 주기가 1인 주기함수이다.
iii. 함수 D_k는 $[0, 1]$에서 $1/(2k+1), \cdots, 2k/(2k+1)$만을 영점으로 가진다.
iv. 함수 D_k는 각 $n \in \mathbb{Z}$에 대해 n에서 최댓값 $2k+1$을 가진다.
v. 임의의 $x \in \mathbb{R}$에 대해 $x \in \mathbb{Z}$이면 $\lim_{k\to\infty} D_k(x) = \infty$이고 $x \notin \mathbb{Z}$이면 $\lim_{k\to\infty} D_k(x) = 0$이다.
vi. Dirichlet kernel은 $\mathscr{C}^\infty$급이다.

Proof v를 제외한 나머지는 Dirichlet kernel의 정의와 정리 0.115로부터 자명하다. v의 경우 $x \in \mathbb{Z}$인 경우에는 자명하고, $x \notin \mathbb{Z}$인 경우에는 다시 정리 0.115로부터

$$\begin{aligned}\lim_{k\to\infty} D_k(x) &= \sum_{i=-\infty}^{\infty} e^{2\pi \mathbf{i} ix} \\ &= 1 + \sum_{i=1}^{\infty} e^{2\pi \mathbf{i} ix} + \sum_{i=1}^{\infty} e^{-2\pi \mathbf{i} ix} \\ &= 1 + e^{2\pi \mathbf{i} x}/(1 - e^{2\pi \mathbf{i} x}) + e^{-2\pi \mathbf{i} x}/(1 - e^{-2\pi \mathbf{i} x}) \\ &= 0\end{aligned}$$

이다. □

다음 특수함수로 넘어가기 전에, Dirichlet kernel의 쓰임새를 하나만 더 소개하도록 하겠다. 물리학에서 애용하는 '함수' 중에 **Dirac 델타함수 (− delta function)**라는, 흔히 $\delta(x)$로 쓰는 '함수' 가 있다. 이는 참 골때리는 성질을 가지고 있는데, 0이 아닌 곳에서는 함숫값이 0이고 0에서는 ∞의 함숫값을 가지지만 $\mathbb{R}$에서의 적분은 1이 된다. 당연히, 수학적으로 이는 절대로 존재할 수 없는 함수이다. 함숫값으로 ∞를 가지는 것이야 공역으로 $\overline{\mathbb{R}}$을 주면 되니 상관이 없지만 이 '함수' 는 거의 대부분의 곳에서 0이므로 $\mathbb{R}$에서의 적분값은 절대 1이 될 수 없다. 이런 상황에서 충분히 큰 $k \in \mathbb{N}$에 대해 D_k를 적당히 scaling하여 δ의 근사식으로 사용될 수 있다.[12]

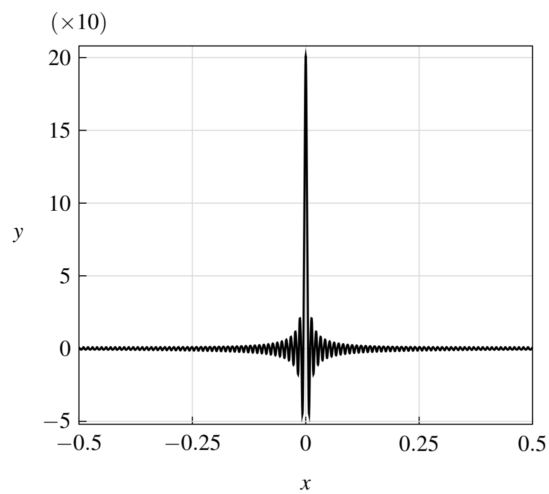


Figure 0.28 Dirichlet kernel $D_{100}(x)$의 그래프. 닫힌구간 $[-1/2, 1/2]$에서 Dirac 델타함수와 꽤 유사하다. Computed by Wolfram MATHEMATICA.

마지막으로 알아볼 특수함수는 sinc 함수의 적분으로 정의된다.

Definition 0.117 다음과 같이 정의된 함수 $\mathrm{Si} : \mathbb{R} \to \mathbb{R}$를 **사인적분함수 (sine integral function)**이라 한다.

$$\mathrm{Si} : x \mapsto \int_0^x \mathrm{sinc}(t)\, dt$$

해석개론 시간에 조건수렴에 그치는 이상적분의 예시로 $\int_0^\infty \sin x/x\, dx$를 본 기억이 새록새록 날 것이다. 이 적분은 Gauss 적분과 더불어 유명한 적분 중 하나로 Dirichlet 적분이라는 이름이 붙어 있으며 $\pi/2$로 조건수렴한다는 사실이 잘 알려져 있다. 물론, 조건수렴에 그치기 때문에 Gauss 적분보다는 다루기 조금 어렵다. 이론적으로 엄밀한 증명을 위해서는 Riemann-Lebesgue의 정리가 필요한데, 이의 증명에서 적분가능한 함수의 근사가 사용되

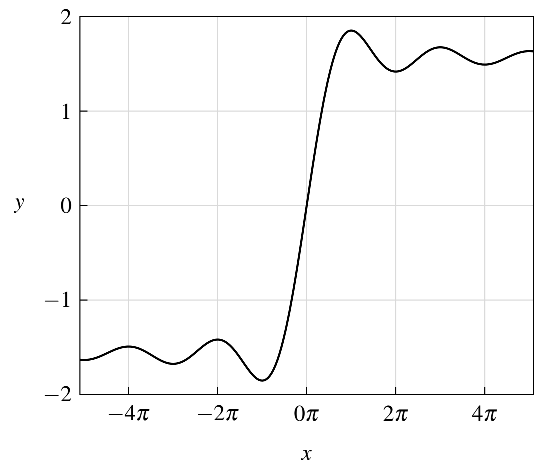


Figure 0.29 사인적분함수의 그래프. Computed by Wolfram MATHEMATICA.

므로 잠시 1장으로 건너가 정리 ? ? 의 결과를 참조하는 것을 추천한다. 이 정리는 이후 2 장의 확률론에서 조금 더 일반적인 형태로 다시 증명될 것이다.

Theorem 0.118 (Riemann-Lebesgue) 적분가능한 함수 $f : \mathbb{R} \to \overline{\mathbb{R}}$에 대해 $t \to \infty$이면

$$\int_{\mathbb{R}} f(x) e^{\mathbf{i}tx}\, d\lambda_1(x) \to 0$$

이다.

Proof 함수 f가 적분가능하므로 임의의 $\varepsilon > 0$을 택하면 정리 ? ? 의 i로부터 단순함수 $g : \mathbb{R} \to \mathbb{R}$가 존재하여 이는 적당한 $a_1, \cdots, a_k \in \mathbb{R}$와 유계인 $B_1, \cdots, B_k \in \mathscr{S}_1$에 대해 $g = \sum_{i=1}^k a_i \mathbf{1}_{B_i}$로 쓸 수 있으며 $\int_{\mathbb{R}} |f - g|\, d\lambda_1 < \varepsilon/2$이다. 여기서 각 $i \le k$에 대해 $B_i = (b_i, c_i]$라 하면

$$\begin{aligned}
\lim_{t\to\infty} \int_{\mathbb{R}} g(x) e^{\mathbf{i}tx}\, d\lambda_1(x) &= \sum_{i=1}^k a_i \lim_{t\to\infty} \int_{b_i}^{c_i} e^{\mathbf{i}tx}\, d\lambda_1(x) \\
&= \sum_{i=1}^k a_i \lim_{t\to\infty} \frac{e^{\mathbf{i}tc_i} - e^{\mathbf{i}tb_i}}{\mathbf{i}t} \\
&= 0
\end{aligned}$$

이므로 적당한 $M > 0$이 존재하여 $t \ge M$이면 $|\int_{\mathbb{R}} g(x) e^{\mathbf{i}tx}\, d\lambda_1(x)| < \varepsilon/2$이다. 이로부터 $t \ge M$이면

$$\left| \int_{\mathbb{R}} f(x) e^{\mathbf{i}tx}\, d\lambda_1(x) \right| \le \left| \int_{\mathbb{R}} [f(x) - g(x)] e^{\mathbf{i}tx}\, d\lambda_1(x) \right| + \left| \int_{\mathbb{R}} g(x) e^{\mathbf{i}tx}\, d\lambda_1(x) \right|$$

$$
\begin{aligned}
&< \int_{\mathbb{R}} |f(x) - g(x)|\, d\lambda_1(x) + \frac{\varepsilon}{2} \\
&< \varepsilon
\end{aligned}
$$

이 되어 정리가 성립한다. □

Corollary 0.119 적분가능한 $f : \mathbb{R} \to \overline{\mathbb{R}}$에 대해 $t \to \infty$이면

$$\int_{\mathbb{R}} f(x) \sin(tx)\, d\lambda_1(x),\ \int_{\mathbb{R}} f(x) \cos(tx)\, d\lambda_1(x) \to 0$$

이다.

Proof 이는 Riemann-Lebesgue의 정리로부터 자명하다. □

Lemma 0.120 임의의 $k \in \mathbb{N}$에 대해 $\int_0^{1/2} D_k(x)\, dx = 1/2$이다.

Proof 정리 0. 115로부터

$$
\begin{aligned}
\int_{-1/2}^{1/2} D_k(x)\, dx &= \sum_{i=-k}^{k} \int_{-1/2}^{1/2} e^{2\pi \mathbf{i} i x}\, dx \\
&= \sum_{i=1}^{k} \int_{-1/2}^{1/2} e^{2\pi \mathbf{i} i x} + e^{-2\pi \mathbf{i} i x}\, dx + 1 \\
&= \sum_{i=1}^{k} \left[\frac{e^{2\pi \mathbf{i} i x} - e^{-2\pi \mathbf{i} i x}}{2\pi \mathbf{i} i} \right]_{-1/2}^{1/2} + 1 \\
&= \sum_{i=1}^{k} \frac{e^{\pi \mathbf{i} i} - e^{-\pi \mathbf{i} i}}{\pi \mathbf{i} i} + 1 \\
&= 2 \sum_{i=1}^{k} \frac{\sin \pi i}{\pi i} + 1 \\
&= 1
\end{aligned}
$$

인데, Dirichlet kernel이 우함수이므로 보조정리가 성립한다. □

Theorem 0.121 (Dirichlet integral)

$$\int_0^{\infty} \operatorname{sinc}(x)\, dx = \frac{\pi}{2} \qquad \text{(conditionally)}$$

Proof 충분히 큰 $M \in \mathbb{R}$에 대해 $\int_1^M \sin x / x\, dx = [-\cos x / x]_1^M - \int_1^M \cos x / x^2\, dx = \cos M / M - \cos 1 - \int_1^M \cos x / x^2\, dx$인데, $\int_1^{\infty} |\cos x| / x^2\, dx \le \int_1^{\infty} 1/x^2\, dx = 1 < \infty$이므로 $\int_0^{\infty} \operatorname{sinc}(x)\, dx$가 조건수렴함은 분명하다. 이제 함수 $f : (0, 1/2] \to \mathbb{R}$를 $f : x \mapsto 1/\pi x - 1/\sin \pi x$로 두면 L'Hospital의 법칙으로부터

$$\begin{aligned}
\lim_{x\downarrow 0} f(x) &= \lim_{x\downarrow 0} \frac{\sin \pi x - \pi x}{\pi x \sin \pi x} \\
&= \lim_{x\downarrow 0} \frac{\cos \pi x - 1}{\sin \pi x + \pi x \cos \pi x} \\
&= \lim_{x\downarrow 0} \frac{-\sin \pi x}{2\cos \pi x - x \sin \pi x} \\
&= 0
\end{aligned}$$

이므로 f는 유계이고, 곧 적분가능하다. 그렇다면 위의 보조정리와 Riemann-Lebesgue의 정리로부터

$$\begin{aligned}
\int_0^\infty \operatorname{sinc}(x)\,dx &= \lim_{k\to\infty} \int_0^{(2k+1)\pi/2} \frac{\sin x}{x}\,dx \\
&= \lim_{k\to\infty} \int_0^{1/2} \frac{\sin(2k+1)\pi x}{x}\,dx \\
&= \pi \lim_{k\to\infty} \int_0^{1/2} \left[f(x) + \frac{1}{\sin \pi x} \right] \sin(2k+1)\pi x\,dx \\
&= \pi \lim_{k\to\infty} \left[\int_0^{1/2} f(x) \sin(2k+1)\pi x\,dx + \int_0^{1/2} D_k(x)\,dx \right] \\
&= \frac{\pi}{2}
\end{aligned}$$

를 얻는다. □

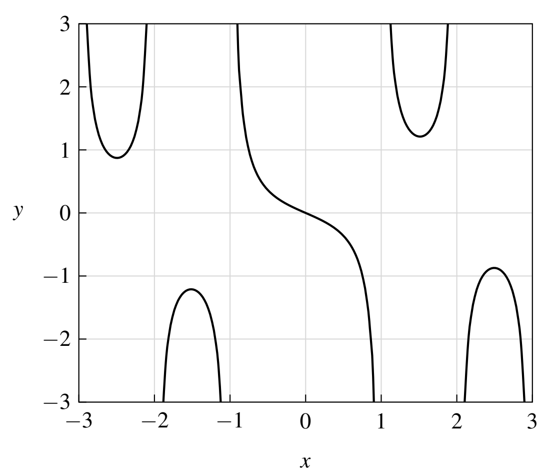


Figure 0.30 Dirichlet 적분의 증명에서의 함수 $f(x)$의 그래프. Computed by Wolfram MATHEMATICA.

이제 사인적분함수의 성질을 간단히 알아보는 것으로 사인 함수와 관련된 특수함수에 대한 논의를 마치도록 하겠다.

Theorem 0.122 다음이 성립한다.

i. 사인적분함수는 기함수이다.
ii. 사인적분함수는 0에서 유일한 영점을 가진다.
iii. 사인적분함수는 각 $k \in \mathbb{N}$에 대해 $(2k\pi, (2k+1)\pi), (-(2k+1)\pi, 2k\pi), (-\pi, \pi)$에서 순증가하고 $((2k-1)\pi, 2k\pi), (-2k\pi, -(2k-1)\pi)$에서 순감소하며 $\pi, -\pi$에서 각각 최댓값과 최솟값을 가진다.
iv. $\lim_{x\to-\infty} \mathrm{Si}(x) = -\pi/2$.
v. $\lim_{x\to\infty} \mathrm{Si}(x) = \pi/2$.
vi. 사인적분함수는 $\mathscr{C}^\infty$급이다.

Proof i, iv, v, vi. 이는 사인적분함수의 정의와 sinc 함수의 성질 그리고 Dirichlet 적분으로부터 자명하다.

ii, iii. 우선 $\mathrm{Si}' = \mathrm{sinc}$라는 점에서 사인적분함수의 증감에 대한 부분은 자명하다. 한편, 남은 부분을 보이기 위해 각 $i \in \mathbb{N}_0$에 대해 $x_i = (2i+1)\pi$라 하면 $\{x_i\}$는 Si가 $\mathbb{R}_0^+$에서 극대가 되는 모든 점들로 구성된 수열이다. 여기서 각 $i \in \mathbb{N}$에 대해

$$
\begin{aligned}
\mathrm{Si}(x_{i+1}) - \mathrm{Si}(x_i) &= \int_{(2i+1)\pi}^{(2i+3)\pi} \frac{\sin t}{t}\,dt \\
&= \left[-\frac{\cos t}{t}\right]_{(2i+1)\pi}^{(2i+3)\pi} - \int_{(2i+1)\pi}^{(2i+3)\pi} \frac{\cos t}{t^2}\,dt \\
&= \frac{1}{(2i+3)\pi} - \frac{1}{(2i+1)\pi} - \int_{(2i+1)\pi}^{(2i+3)\pi} \frac{\cos t}{t^2}\,dt \\
&< \frac{1}{(2i+3)\pi} - \frac{1}{(2i+1)\pi} + \int_{(2i+1)\pi}^{(2i+3)\pi} \frac{1}{t^2}\,dt \\
&= 0
\end{aligned}
$$

이므로 $\mathrm{Si}(x_{i+1}) < \mathrm{Si}(x_i)$가 되어 $\mathbb{R}_0^+$에서의 Si의 최댓값은 $\mathrm{Si}(\pi)$이다. 이제 임의의 $x \geq 0$에 대해 $\mathrm{Si}(x) \geq 0$임을 보이면 i로부터 증명이 끝난다. 이를 위해 각 $i \in \mathbb{N}$에 대해 $y_i = 2i\pi$라 하면 $\{y_i\}$는 Si가 $\mathbb{R}_0^+$에서 극소가 되는 모든 점들로 구성된 수열이고, 방금과 같이 하면 $\mathrm{Si}(y_i) < \mathrm{Si}(y_{i+1})$임을 쉽게 보일 수 있다. 그런데 충분히 작은 임의의 $\varepsilon > 0$에 대해

$$
\begin{aligned}
\mathrm{Si}(2\pi) - \mathrm{Si}(\varepsilon) &= \int_\varepsilon^{2\pi} \frac{\sin t}{t}\,dt \\
&= \left[-\frac{\cos t}{t}\right]_\varepsilon^{2\pi} - \int_\varepsilon^{2\pi} \frac{\cos t}{t^2}\,dt
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\cos\varepsilon}{\varepsilon} - \frac{1}{2\pi} - \int_{\varepsilon}^{2\pi} \frac{\cos t}{t^2}\,dt \\
&> \frac{\cos\varepsilon}{\varepsilon} - \frac{1}{2\pi} - \int_{\varepsilon}^{\pi/2} \frac{1}{t^2}\,dt - \int_{3\pi/2}^{2\pi} \frac{1}{t^2}\,dt \\
&= \frac{\cos\varepsilon - 1}{\varepsilon} + \frac{4}{3\pi}
\end{aligned}
$$

이고, $\varepsilon \downarrow 0$이면 $(\cos\varepsilon - 1)/\varepsilon \to 0$이므로 충분히 작은 $\varepsilon > 0$에 대해 $\mathrm{Si}(2\pi) > \mathrm{Si}(\varepsilon)$이어서 vi으로부터 $\mathrm{Si}(2\pi) \geq \lim_{\varepsilon\downarrow 0} \mathrm{Si}(\varepsilon) = 0$이다. iii의 증명은 이로써 충분하고 ii 또한 증명 과정으로부터 자명하다. □

0.2.6 More Special Functions

여기서는 이후의 논의에서 사용되지만 다른 특수함수와 특별한 접점이 없는 함수들을 모아 다루어보고자 한다. 먼저 살펴볼 특수함수는 급수로 정의되는 함수로 다른 특수함수를 포함한 여러 복잡한 함수를 표현하는 용도로 널리 쓰이는 함수이다.

Definition 0.123 벡터 $a \in \mathbb{R}^m, b \in \mathbb{R}^n$에 대해 well-define되는 멱급수 $\sum_{i=0}^{\infty} a^{\overline{i\mathbf{1}}} z^i / b^{\overline{i\mathbf{1}}} i!$의 수렴반경을 $R \geq 0$이라 하자. 이때, 다음과 같이 정의되는 함수를 **(일반화된) 초기하함수 ((generalized) hypergeometric function)**라 하고 ${}_mF_n(a;b;\cdot) : B(R) \to \mathbb{C}$로 쓴다.

$$
{}_mF_n(a;b;\cdot) : z \mapsto \sum_{i=0}^{\infty} \frac{a^{\overline{i\mathbf{1}}}}{b^{\overline{i\mathbf{1}}} i!} z^i
$$

특별히, ${}_1F_1(a;b;\cdot)$을 **제1종 합류 초기하함수 (confluent hypergeometric function of the first kind)**, ${}_2F_1(a;b;\cdot)$을 **(Gaussian) 초기하함수 (− hypergeometric function)**라 하며, ${}_1F_1(a;b;\cdot)$은 $M(a;b;\cdot)$으로 쓰기도 한다.

보통 m, n이 0이 되는 경우도 허용하여 각각 $a^{\bar{i}}$와 $b^{\bar{i}}$ 자리에 1이 대신 들어가는 것으로 정의하지만, 이 책에서는 $m, n \in \mathbb{N}$인 경우만 생각해도 충분하다. 한편, 위의 정의에는 급수 $\sum_{i=0}^{\infty} a^{\overline{i\mathbf{1}}} z^i / b^{\overline{i\mathbf{1}}} i!$이 well-define되어야 한다는 조건이 붙어 있는데, 이는 바꾸어 말하면 어떤 $a \in \mathbb{R}^m, b \in \mathbb{R}^n$에 대해서는 방금의 급수가 잘 정의되지 않을 수도 있다는 뜻으로, 이를 판단하는 데에는 다음 정리가 유용하다.

Theorem 0.124 임의의 $a \in \mathbb{R}^m, b \in \mathbb{R}^n$에 대해 다음이 성립한다.

i. 만약 b가 양이 아닌 정수를 그 성분으로 가지면 ${}_mF_n(a;b;\cdot)$는 정의되지 않는다.
ii. 만약 b가 양이 아닌 정수를 그 성분으로 가지지 않고 a가 양이 아닌 정수를 그 성분으로 가지면 ${}_mF_n(a;b;\cdot)$는 유한차수 다항식이다.

iii. 만약 a, b가 양이 아닌 정수를 그 성분으로 가지지 않으면 다음이 성립한다.

a. 만약 $m < n+1$이면 ${}_mF_n(a;b;\cdot)$의 수렴반경은 ∞이다.
b. 만약 $m = n+1$이면 ${}_mF_n(a;b;\cdot)$의 수렴반경은 1이다.
c. 만약 $m > n+1$이면 ${}_mF_n(a;b;\cdot)$의 수렴반경은 0이다.

PROOF i. 만약 b가 그 성분으로 양이 아닌 정수를 가지면 충분히 큰 $i \in \mathbb{N}_0$에 대해 $b^{\overline{i\mathbf{1}}} = 0$이 되어 ${}_mF_n(a;b;\cdot)$가 정의되지 않는다.

ii. 우선 b가 그 성분으로 양이 아닌 정수를 가지지 않으므로 ${}_mF_n(a;b;\cdot)$은 well-define 되는데, a가 그 성분으로 양이 아닌 정수를 가지면 적당한 $i_0 \in \mathbb{N}_0$에 대해 $i \geq i_0$이면 $a^{\overline{i_0\mathbf{1}}} = 0$이 되어 ${}_mF_n(a;b;\cdot)$이 유한차수 다항식이 된다.

iii. 우선 a, b가 그 성분으로 양이 아닌 정수를 가지지 않으므로 ${}_mF_n(a;b;\cdot)$은 well-define 되며 유한차수 다항식이 아니다. 이제 만약 $m < n+1$이면

$$\begin{aligned}\lim_{i\to\infty}\left|\frac{a^{\overline{i\mathbf{1}}}/b^{\overline{i\mathbf{1}}}i!}{a^{\overline{(i+1)\mathbf{1}}}/b^{\overline{(i+1)\mathbf{1}}}(i+1)!}\right| &= \lim_{i\to\infty}\left|\frac{\prod_{j=1}^{n}(b_j+i)}{\prod_{k=1}^{m}(a_k+i)}\right|(i+1)\\ &= \infty\end{aligned}$$

에서 ${}_mF_n(a;b;\cdot)$의 수렴반경이 ∞임을 알고, 비슷하게 하면 $m = n+1$인 경우와 $m > n+1$인 경우에 ${}_mF_n(a;b;\cdot)$의 수렴반경이 각각 $1, 0$임을 안다. □

앞서 살펴봤던 다른 특수함수들과는 달리, 초기하함수는 다양한 형태를 가질 수 있어서 특기할만한 공통적인 성질이 많지 않다. 미분 공식 정도만 간단히 살펴보고 앞서 배운 특수함수들이 초기하함수로써 어떻게 표현되는지를 보는 것으로 마무리하자.

Theorem 0.125 양이 아닌 정수를 그 성분으로 가지지 않는 $a \in \mathbb{R}^l, b \in \mathbb{R}^m$에 대해 급수 $\sum_{i=0}^{\infty} a^{\overline{i\mathbf{1}}}x^i/b^{\overline{i\mathbf{1}}}i!$의 수렴반경을 $R \geq 0$이라 하면 함수 ${}_lF_m(a;b;\cdot) : (-R,R) \to \mathbb{R}$는 $\mathscr{C}^\infty$급이며 $n \in \mathbb{N}_0$에 대해

$$\frac{d^n}{dx^n}{}_lF_m(a;b;x) = \frac{a^{\overline{n\mathbf{1}}}}{b^{\overline{n\mathbf{1}}}}{}_lF_m(a+n\mathbf{1};b+n\mathbf{1};x)$$

이다. 만약 a가 양이 아닌 정수를 그 성분으로 가져 ${}_lF_m(a;b;\cdot) : \mathbb{R} \to \mathbb{R}$가 k차 다항식이 되는 경우에도 동일한 결과가 성립한다.

PROOF 우선 ${}_lF_m(a;b;\cdot)$가 $\mathscr{C}^\infty$급임은 자명하고, 수렴반경 안에서 멱급수로 정의된 함수의 미분은 각 항의 미분으로 주어지므로

$$\frac{d^n}{dx^n}{}_lF_m(a;b;x) = \frac{d^n}{dx^n}\sum_{i=0}^{\infty}\frac{a^{\overline{i\mathbf{1}}}}{b^{\overline{i\mathbf{1}}}i!}x^i$$

$$
\begin{aligned}
&= \sum_{i=n}^{\infty} \frac{a^{\overline{i\mathbf{1}}}}{b^{\overline{i\mathbf{1}}}(i-n)!} x^{i-n} \\
&= \sum_{i=0}^{\infty} \frac{a^{\overline{(i+n)\mathbf{1}}}}{b^{\overline{(i+n)\mathbf{1}}} i!} x^{i} \\
&= \frac{a^{\overline{n\mathbf{1}}}}{b^{\overline{n\mathbf{1}}}} \sum_{i=0}^{\infty} \frac{(a+n\mathbf{1})^{\overline{i\mathbf{1}}}}{(b+n\mathbf{1})^{\overline{i\mathbf{1}}} i!} x^{i} \\
&= \frac{a^{\overline{n\mathbf{1}}}}{b^{\overline{n\mathbf{1}}}} {}_lF_m(a+n\mathbf{1}; b+n\mathbf{1}; x)
\end{aligned}
$$

가 성립한다. 한편, a가 양이 아닌 정수를 그 성분으로 가지는 경우에는 이가 k차 다항식이므로 동일한 결과를 얻는다. □

Theorem 0.126 임의의 $k \in \mathbb{N}_0$, 임의의 $\alpha, \beta > 0$, 임의의 $y \geq 0$에 대해 다음이 성립한다.

i. $He_{2k}(x) = (-1)^k[(2k)!/k!]{}_1F_1(-k; 1/2; x^2)$.
ii. $He_{2k+1}(x) = (-1)^k[(2k+1)!/k!]2x{}_1F_1(-k; 3/2; x^2)$.
iii. $\mathrm{erf}(x) = (2x/\sqrt{\pi}){}_1F_1(1/2; 3/2; -x^2)$.
iv. $\gamma(x, y) = (y^x/x){}_1F_1(x; x+1; -y)$.
v. $\mathrm{B}(x; \alpha, \beta) = (x^\alpha/\alpha){}_2F_1(\alpha, 1-\beta; \alpha; x)$.
vi. $\mathrm{Si}(x) = x{}_1F_2(1/2; 3/2, 3/2; -x^2/4)$.

여기서 $x \in \mathbb{R}$는 임의의 실수이지만, iv와 v에서는 각각 양수, 1보다 작은 음이 아닌 실수로 생각한다.

Proof i, ii. 정리 0. 52로부터

$$
\begin{aligned}
He_{2k}(x) &= (2k)! \sum_{i=0}^{k} \frac{(-1)^i}{i!(2k-2i)!} (2x)^{2k-2i} \\
&= (2k)! \sum_{i=0}^{k} \frac{(-1)^{k-i}}{(k-i)!(2i)!} (2x)^{2i} \\
&= (-1)^k \frac{(2k)!}{k!} \sum_{i=0}^{k} \frac{i!2^i}{(2i)!/2^i} \cdot \frac{(-1)^i k!}{(k-i)!i!} x^{2i} \\
&= (-1)^k \frac{(2k)!}{k!} \sum_{i=0}^{k} \frac{(-k)^{\overline{i}}}{(1/2)^{\overline{i}} i!} x^{2i} \\
&= (-1)^k \frac{(2k)!}{k!} \sum_{i=0}^{\infty} \frac{(-k)^{\overline{i}}}{(1/2)^{\overline{i}} i!} x^{2i} \\
&= (-1)^k \frac{(2k)!}{k!} {}_1F_1\left(-k; \frac{1}{2}; x^2\right)
\end{aligned}
$$

이 되어 i이 성립하고, ii도 이와 비슷하게 보일 수 있다.

iii. 닫힌구간 $[0,x]$에서 $e^{-t^2} = \sum_{i=0}^{\infty}(-t^2)^i/i!$이 균등수렴하므로

$$\begin{aligned}
\operatorname{erf}(x) &= \frac{2}{\sqrt{\pi}}\int_0^x \sum_{i=0}^{\infty}\frac{(-1)^i}{i!}t^{2i}\,dt \\
&= \frac{2}{\sqrt{\pi}}\sum_{i=0}^{\infty}\frac{(-1)^i}{i!}\int_0^x t^{2i}\,dt \\
&= \frac{2x}{\sqrt{\pi}}\sum_{i=0}^{\infty}\frac{(-x^2)^i}{(2i+1)i!} \\
&= \frac{2x}{\sqrt{\pi}}\sum_{i=0}^{\infty}\frac{(2i-1)!!/2^i}{(2i+1)!!/2^i}\cdot\frac{(-x^2)^i}{i!} \\
&= \frac{2x}{\sqrt{\pi}}\sum_{i=0}^{\infty}\frac{(1/2)^{\bar{i}}}{(3/2)^{\bar{i}}i!}(-x^2)^i \\
&= \frac{2x}{\sqrt{\pi}}\,{}_1F_1\left(\frac{1}{2};\frac{3}{2};-x^2\right)
\end{aligned}$$

이다.

iv. 닫힌구간 $[0,y]$에서 충분히 큰 $i_0\in\mathbb{N}$에 대해 $t^{x-1}e^{-t} = \sum_{i=i_0}^{\infty}(-1)^i t^{x+i-1}/i!$이 균등수렴하므로

$$\begin{aligned}
\gamma(x,y) &= \int_0^y \sum_{i=0}^{\infty}\frac{(-1)^i}{i!}t^{x+i-1}\,dt \\
&= \sum_{i=0}^{\infty}\frac{(-1)^i}{i!}\int_0^y t^{x+i-1}\,dt \\
&= y^x\sum_{i=0}^{\infty}\frac{(-y)^i}{(x+i)i!} \\
&= \frac{y^x}{x}\sum_{i=0}^{\infty}\frac{x^{\bar{i}}}{(x+1)^{\bar{i}}i!}(-y)^i \\
&= \frac{y^x}{x}\,{}_1F_1(x;x+1;-y)
\end{aligned}$$

이다.

v. 닫힌구간 $[0,x]$에서 충분히 큰 $i_0\in\mathbb{N}$에 대해 $t^{\alpha-1}(1-t)^{\beta-1} = \sum_{i=i_0}^{\infty}(-1)^i(\beta-1)^{\underline{i}}t^{\alpha+i-1}/i!$가 균등수렴하므로

$$\begin{aligned}
\mathrm{B}(x;\alpha,\beta) &= \int_0^x \sum_{i=0}^{\infty}(-1)^i\frac{(\beta-1)^{\underline{i}}}{i!}t^{\alpha+i-1}\,dt \\
&= \sum_{i=0}^{\infty}(-1)^i\frac{(\beta-1)^{\underline{i}}}{i!}\int_0^y t^{\alpha+i-1}\,dt
\end{aligned}$$

$$
\begin{aligned}
&= x^\alpha \sum_{i=0}^{\infty} \frac{(1-\beta)^{\bar{i}}}{(\alpha+i)i!} x^i \\
&= \frac{x^\alpha}{\alpha} \sum_{i=0}^{\infty} \frac{\alpha^{\bar{i}}(1-\beta)^{\bar{i}}}{(\alpha+1)^{\bar{i}} i!} x^i \\
&= \frac{x^\alpha}{\alpha} {}_2F_1(\alpha, 1-\beta; \alpha; x)
\end{aligned}
$$

이다.

vi. 닫힌구간 $[0, x]$에서 $\operatorname{sinc}(t) = \sum_{i=0}^{\infty} (-1)^i t^{2i}/(2i+1)!$이 균등수렴하므로

$$
\begin{aligned}
\operatorname{Si}(x) &= \int_0^x \operatorname{sinc}(t)\, dt \\
&= \int_0^x \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i+1)!} t^{2i}\, dt \\
&= \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i+1)!} \int_0^x t^{2i}\, dt \\
&= \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i+1)(2i+1)!} x^{2i+1}\, dt \\
&= x \sum_{i=0}^{\infty} \frac{(2i-1)!!}{(2i)!![(2i+1)!!]^2/4^i} \left(-\frac{x^2}{4}\right)^i dt \\
&= x \sum_{i=0}^{\infty} \frac{(1/2)^{\bar{i}}}{[(3/2)^{\bar{i}}]^2} \left(-\frac{x^2}{4}\right)^i dt \\
&= x\, {}_1F_2\left(\frac{1}{2}; \frac{3}{2}, \frac{3}{2}; -\frac{x^2}{4}\right)
\end{aligned}
$$

이다. □

다음으로 살펴볼 것은 정확히 말하면 특수함수는 아니고 특별한 수 정도 된다.

Definition 0.127 자연수 $m, n \in \mathbb{N}_0$에 대해 m개의 서로다른 원소로 만들 수 있는 n개의 서로소인 cycle로 구성된 순열의 수를 **(부호 없는) 제1종 Stirling 수 ((unsigned) - number of the first kind)**라 하며 $c(m, n)$ 혹은 다음과 같이 쓴다.

$$
\begin{bmatrix} m \\ n \end{bmatrix}
$$

한편, m개의 서로다른 원소를 n개의 비어있지 않은 부분집합으로 분할하는 경우의 수를 **(부호 없는) 제2종 Stirling 수 ((unsigned) - number of the second kind)**라 하며 $s(m, n)$ 혹은 다음과 같이 쓴다.

$$\begin{Bmatrix} m \\ n \end{Bmatrix}$$

조합론적으로 정의된 Stirling 수가 수리통계학에 필요한 이유는 이가 상향계승 혹은 하향계승과 계승을 이어주는 역할을 하기 때문이다. 그리고 이런 역할을 생각해보면

Theorem 0.128 임의의 $m, n \in \mathbb{N}_0$에 대해 다음이 성립한다.

i. $c(m, n) = \begin{cases} 0 & m = 0 \text{ 혹은 } n = 0 \text{ 혹은 } m < n \text{인 경우} \\ mc(m-1, n) + c(m-1, n-1) & \text{ow.} \end{cases}$

ii. $s(m, n) = \begin{cases} 0 & m = 0 \text{ 혹은 } n = 0 \text{ 혹은 } m < n \text{인 경우} \\ ns(m-1, n) + s(m-1, n-1) & \text{ow.} \end{cases}$

Proof i. 우선 $m = 0$ 혹은 $n = 0$ 혹은 $m < n$인 경우에 $c(m, n) = 0$임은 자명하므로 $m, n \geq 1$이라 하자. 그렇다면 m개의 서로다른 원소로 n개의 서로소인 cycle로 구성된 순열을 만드는 방법은 두 가지가 있다. 첫째로 마지막 원소를 제외한 $m-1$개의 원소로 $n-1$개의 서로소인 cycle로 구성된 순열을 만든 후에 남은 하나의 원소를 고정점으로 하는 방법이 있다. 두번째로 마지막 원소를 제외한 $m-1$개의 원소로 n개의 서로소인 cycle로 구성된 순열을 만든 후에 남은 하나의 원소를 n개의 cycle 중 하나에 적당히 끼워넣는 방법이 있다. 전자의 방법으로 만들어지는 순열의 수가 $c(m-1, n-1)$이고, 후자의 방법으로 만들어지는 순열의 수가 $mc(m-1, n)$이므로 $c(m, n) = mc(m-1, n) + c(m-1, n-1)$이다.

ii. 이번에도 $m = 0$ 혹은 $n = 0$ 혹은 $m < n$인 경우에 $c(m, n) = 0$임은 자명하므로 $m, n \geq 1$이라 하자. 그렇다면 m개의 서로다른 원소를 n개의 비어있지 않은 부분집합으로 분할하는 방법은 두 가지가 있다. 첫째로 마지막 원소를 제외한 $m-1$개의 원소를 $n-1$개의 부분집합으로 분할한 후에 남은 하나의 원소로 한원소 집합을 구성하는 방법이 있다. 두번째로 마지막 원소를 제외한 $m-1$개의 원소를 n개의 부분집합으로 분할한 후에 남은 하나의 원소를 n개의 부분집합 중 하나에 적당히 끼워넣는 방법이 있다. 전자의 방법으로 만들 수 있는 분할의 수가 $s(m-1, n-1)$이고, 후자의 방법으로 만들 수 있는 분할의 수가 $ns(m-1, n)$이므로 $s(m, n) = ns(m-1, n) + s(m-1, n-1)$이다. □

Theorem 0.129 임의의 $n \in \mathbb{N}_0$에 대해 다음이 성립한다.

i. $x^n = \sum_{k=0}^{n} (-1)^{n-k} s(n, k) x^{\overline{k}} = \sum_{k=0}^{n} s(n, k) x^{\underline{k}}$.
ii. $x^{\underline{n}} = \sum_{k=0}^{n} c(n, k) x^k$.
iii. $x^{\overline{n}} = \sum_{k=0}^{n} (-1)^{n-k} c(n, k) x^k$.

Proof i. 수학적 귀납법을 사용하자. 우선 $n = 0$인 경우에는 정리가 자명하므로 귀납가정으로서 $n \in \mathbb{N}_0$에 대해 정리가 성립한다고 가정하면 정리 0.128로부터

$$
\begin{aligned}
x^{n+1} &= x\sum_{k=0}^{n}(-1)^{n-k}s(n,k)x^{\overline{k}} \\
&= \sum_{k=0}^{n}(-1)^{n-k}s(n,k)(x^{\overline{k+1}}-kx^{\overline{k}}) \\
&= \sum_{k=1}^{n+1}(-1)^{n-k+1}s(n,k-1)x^{\overline{k}}+\sum_{k=1}^{n}(-1)^{n-k+1}ks(n,k)x^{\overline{k}} \\
&= \sum_{k=1}^{n+1}(-1)^{n-k+1}[s(n,k-1)+ks(n,k)]x^{\overline{k}} \\
&= \sum_{k=1}^{n+1}(-1)^{n-k+1}s(n+1,k)x^{\overline{k}} \\
&= \sum_{k=0}^{n+1}(-1)^{n-k+1}s(n+1,k)x^{\overline{k}}
\end{aligned}
$$

이므로 $n+1$에 대해서도 정리가 성립하여 곧 증명이 끝난다. 이제 남은 부분도 이와 비슷하게 하면 된다.

ii, iii. 이번에도 수학적 귀납법을 사용한다. 우선 $n=0$인 경우에는 정리가 자명하므로 귀납가정으로서 $n\in\mathbb{N}_0$에 대해 정리가 성립한다고 가정하면 정리 0. 128로부터

$$
\begin{aligned}
x^{\underline{n+1}} &= (x+n)\sum_{k=0}^{n}c(n,k)x^{k} \\
&= \sum_{k=0}^{n}c(n,k)x^{k+1}+\sum_{k=0}^{n}nc(n,k)x^{k} \\
&= \sum_{k=1}^{n+1}c(n,k-1)x^{k}+\sum_{k=1}^{n}nc(n,k)x^{k} \\
&= x^{n+1}+\sum_{k=1}^{n}[c(n,k-1)+nc(n,k)]x^{k} \\
&= x^{n+1}+\sum_{k=1}^{n}c(n+1,k)x^{k} \\
&= \sum_{k=0}^{n+1}c(n+1,k)x^{k}
\end{aligned}
$$

이므로 $n+1$에 대해서도 정리가 성립하여 곧 증명이 끝난다. 이제 iii도 이와 비슷하게 하면 된다. □

Notes

1 일반적으로 열린집합 $U \subseteq \mathbb{R}$에서 정의된 미분가능한 함수 $f: U \to \mathbb{R}$에 대해 $x_0 \in \mathbb{R}$에서 $f(x_0) = 0$이고 $x \to \infty$일 때 $f(x) \to 0$이면 f'은 (x_0, ∞)에서 영점을 가지며, 이는 Rolle의 정리의 확장으로 생각할 수 있다. 이를 보이기 위해 먼저 f가 (x_0, ∞)에서 영점을 가지는 경우를 생각하여 그 영점을 $x_1 > x_0$이라 하면 f에 대한 (x_0, x_1)에서의 Rolle의 정리로부터 f'이 $(x_0, x_1) \subseteq (x_0, \infty)$에서 영점을 가짐을 안다. 반대로 f가 (x_0, ∞)에서 영점을 가지지 않는다면 이는 곧 f가 (x_0, ∞)에서 부호가 바뀌지 않음을 의미하고, 임의의 $y > x_0$를 택하면 가정으로부터 충분히 큰 $x_1' > y$에 대해 $|f(x_1')| < |f(y)|$이다. 또한, IVT로부터 적당한 $x_0' \in (x_0, y)$이 존재하여 $f(x_0') = f(x_1')$이므로 이번에도 f에 대한 (x_0', x_1')에서의 Rolle의 정리로부터 f'이 $(x_0', x_1') \subseteq (x_0, \infty)$에서 영점을 가짐을 안다.

2 증명에는 Galois 이론을 비롯한 대수학의 지식이 필요하므로 여기에서는 참고할 만한 문헌을 풍부하게 제시하는 것으로 증명을 갈음한다.

- A. D. Fitt and G. T. Q. Hoare, "The Closed-form Integration of Arbitrary Functions", *The Mathematical Gazette*, vol. 77, no. 479, 1993.
- I. Kaplansky, *An Introduction to Differential Algebra*, Hermann, 1957.
- E. R. Kolchin, *Differential Algebra and Algebraic Groups*, Academic Press, 1973.
- A. R. Magid, *Lectures on Differential Galois Theory*, AMS, 1994.
- E. Marchisotto and G. Zakeri, "An Invitation to Integration in Finite Terms", *The College Mathematics Journal*, vol. 25, no. 4, 1994.
- J. F. Ritt, *Integration in Finite Terms*, Columbia, 1948.
- J. F. Ritt, *Differential Algebra*, AMS, 1950.
- M. Rosenlicht, "Liouville's Theorem on Functions With Elementary Integrals", *Pacific Journal of Mathematics*, vol. 24, no. 1, 1968.
- M. Rosenlicht, "Integration in Finite Terms", *The American Mathematics Monthly*, vol. 79, no. 9, 1972.
- G. N. Watson, *A Treatise on The Theory of Bessel Functions*, Cambridge, 1962.

3 점 x_0 근방의 $x > 0$에 대해 적당한 $M > 0$이 존재하여 $|x| \leq M$이므로 각 $i \in \mathbb{N}$에 대해 $x/i(x+i) \leq M/i^2$이고, $\sum_{i=1}^{\infty} M/i^2 = M\pi^2/6 < \infty$이므로 Weierstrass의 $M-$판정법으로부터 $\sum_{i=1}^{\infty} x/i(x+i)$는 x_0의 근방에서 균등수렴한다.

4 이 급수는 사실 $\mathbb{R}^+$ 전체에서 균등수렴한다. 임의의 $x > 0$와 각 $i \in \mathbb{N}$에 대해 $1/(x+i)^2 \leq 1/i^2$이고, $\sum_{i=1}^{\infty} 1/i^2 = \pi^2/6 < \infty$이므로 Weierstrass의 $M-$판정법으로부터 $\sum_{i=1}^{\infty} 1/(x+i)^2$은 $\mathbb{R}^+$에서 균등수렴한다. 한편, 임의의 $k \in \mathbb{N}$에 대해 $\sum_{i=0}^{\infty} (-1)^{k+2}(k+1)!/(x+i)^{k+2}$가 x_0에서 균등수렴함도 비슷하게 보일 수 있다.

5 점 x_0 근방의 $x > 1$에 대해 적당한 $m > 1$이 존재하여 $x \geq m$이므로 곧

$$\begin{aligned}\sum_{i=1}^{\infty} \frac{(\log i)^{n+1}}{i^x} &\leq \sum_{i=1}^{\infty} \frac{(\log i)^{n+1}}{i^m} \\ &< \infty\end{aligned}$$

가 되어 Weierstrass의 $M-$판정법으로부터 $\sum_{i=1}^{\infty} (-\log i)^{n+1}/i^x$는 x_0의 근방에서 균등수렴한다.

6 이는 거의 자명하다. 임의의 $x \geq 2$에 대해 $\sum_{i=1}^{\infty} 1/i^x \leq \sum_{i=1}^{\infty} 1/i^2 = \pi^2/6 < \infty$이므로 Weierstrass의 $M-$판정법으로부터 $\sum_{i=1}^{\infty} 1/i^x$는 $[2, \infty)$에서 균등수렴한다.

7 Smith–Volterra–Cantor 집합은 흔히 SVC 집합이라 약칭되며 Cantor 집합과 같이 $[0,1]$에서 시작하지만 매번 중간 $1/3$을 빼는 과정을 반복하는 대신 처음에는 중간에서 $1/4$ 만큼을 빼고, 그 다음에는 중간에서 $1/16$ 만큼씩을 빼고, 그 다음에는 중간에서 $1/64$ 만큼씩을 빼는 식으로 i번째 단계에서 중간에서 $1/4^i$ 만큼씩을 제거하는 시행을 반복하여 얻어진다. 즉, Cantor 집합의 구성에서는 매번 이전 단계의 $1/3$이라는 고정된 '비율' 을 중간에서 제거하였다면, SVC 집합의 구성에서는 매번 지수적으로 감소하는 '길이' 를 중간에서 제거한다. 이러한 점이 SVC 집합과 Cantor 집합 사이에 미묘한 차이를 만든다.

우선, Cantor 집합과 SVC 집합은 위상적인 성질들을 공유한다. 즉, 둘 모두 NSP로부터 비어있지 않은 compact한 집합이고, 그 구성으로부터 totally disconnected이고 nowhere dense하며 perfect한 집합임을 쉽게 알 수 있다. (사실 Cantor 집합과 SVC 집합은 위상동형이다.) 그러나 둘의 측도론적인 성질은 서로 다른데, Cantor 집합이 영집합이었던 것과 달리 SVC 집합의 구성에서는 i번째 단계에서 제거되는 부분의 길이가 $2^{i-1}/4^i = 1/2^{i+1}$이므로 SVC 집합은 $1-\sum_{i=1}^{\infty} 1/2^{i+1} = 1/2$의 측도를 가진다.

이런 위상적인 성질과 측도론적인 성질의 묘한 조합으로 인하여 재미있는 결과를 몇 개 얻을 수 있다. 먼저, SVC 집합은 그 경계가 양의 측도를 가지는 nowhere dense한 집합이다. 또한, SVC 집합을 S라 하면 Riemann-Lebesgue 정리로부터 함수 $\mathbf{1}_S$는 $[0,1]$에서 Riemann 적분불가능이다. 나아가, SVC 집합을 이용하여 $(0,1)$에서 미분가능하고 그 도함수가 유계임에도 Riemann 적분불가능한 함수를 구성할 수 있다. 이 함수는 **Volterra 함수 (− function)**라 불리는데, SVC 집합에서 함수 $x \mapsto x^2 \sin(1/x)$를 적당히 복사하고 대칭하여 이어 붙이는 방식으로 구성할 수 있다. 구체적인 구성 및 관련 성질의 증명은 [1]의 연습문제나 [12]를 참조하기 바란다.

8 Fractal 기하에 있어 Hausdorff 차원은 핵심적인 개념으로 자리잡고 있다. '차원' 이라는 이름에서 알 수 있듯이 Hausdorff 차원은 fractal이 어느 정도의 공간을 차지하는지를 정량적으로 측정하는 역할을 하며, 흔히 fractal 차원이라 알려진 것의 엄밀한 수학적 형식화로 볼 수 있다. 일단 준비가 조금 필요하다.

거리공간 (M,d)에 대해 $\mathscr{P}(M)$ 위의 외측도 μ^*를 생각하자. 만약 임의의 $A,B \subseteq M$에 대해 $d(A,B) := \inf_{x\in A, y\in B} d(x,y) > 0$일 때 $\mu^*(A\cup B) = \mu^*(A)+\mu^*(B)$가 성립하면 이때의 외측도 μ^*를 **metric outer measure**라 한다. 즉, 쉽게 생각하면 metric outer measure는 서로 떨어진 집합에 대해 가법성이 성립하는 외측도이다. 이제 μ^*가 metric outer measure라 하면 이를 Carathéodory의 확장정리로써 확장하여 측도공간 $(M,\mathscr{M},\mu)$를 얻을 수 있는데, 이에 대해 (M,d)의 Borel σ−대수를 $\mathscr{B}_M$이라 하면 $\mathscr{B}_M \subseteq \mathscr{M}$이 성립한다. 이를 보이는 것은 크게 어렵지 않다. 거리공간 (M,d)의 모든 열린집합의 모임 $\mathscr{U}$가 $\mathscr{B}_M$의 생성자이고 σ−대수의 정의를 생각해보면 (M,d)의 모든 닫힌집합의 모임 $\mathscr{F}$도 $\mathscr{B}_M$의 생성자임을 쉽게 보일 수 있다. (Hint: 이는 $\mathscr{F} \subseteq \sigma(\mathscr{U})$와 $\mathscr{U} \subseteq \sigma(\mathscr{F})$를 보이는 것으로 충분하다.) 따라서 임의의 닫힌집합 $F \subseteq M$와 임의의 $S \subseteq M$를 고정하고 $\mu^*(S) \geq \mu^*(S\cap F)+\mu^*(S\cap F^c)$임을 보이면 증명은 끝난다. 만약 $\mu^*(S)=\infty$이면 이는 자명하게 성립하므로 $\mu^*(S)<\infty$라 가정하고 증가하는 집합열 $\{S_i\}$를 $S_i := \{x \in S\setminus F : \rho(x,F) := \inf_{y\in F} d(x,y) \geq 1/i\}$로 정의하면 각 $i \in \mathbb{N}$에 대해 $d(S_i, S\cap F) \geq d(S_i,F) \geq 1/i$이다. 그렇다면 μ^*가 metric outer measure이므로 각 $i\in\mathbb{N}$에 대해 $\mu^*(S_i \cup (S\cap F)) = \mu^*(S_i)+\mu^*(S\cap F)$가 성립한다. 나아가 F가 닫혀있으므로 임의의 $x \in S\setminus F$에 대해 $d(x,F)>0$이어서 $S_i \uparrow S\setminus F$이고, 곧 $S=(S\cap F)\cup(S\cap F^c) = (S\cap F)\cup\bigcup_{i=1}^{\infty} S_i = \bigcup_{i=1}^{\infty}[(S\cap F)\cup S_i]$가 되어 이상의 결과를 종합하면 각 $i\in\mathbb{N}$에 대해 $\mu^*(S) \geq \mu^*((S\cap F)\cup S_i) = \mu^*(S\cap F)+\mu^*(S_i)$이다. 이로부터 우리는 $\mu^*(S_i) \to \mu(S\cap F^c)$만 보이면 된다. 집합열 $\{T_i\}$를 $T_i := S_{i+1}\setminus S_i$로 두자. 만약 각 $i\in\mathbb{N}$와 임의의 $x\in T_{i+1}$와 임의의 $y\in M$에 대해 $d(x,y)<1/i(i+1)$라면 $d(y,F) \leq d(x,y)+d(x,F) < 1/i(i+1)+1/(i+1) = 1/i$에서 $y \notin S_i$이므로 $d(T_{i+1},S_i) \geq 1/i(i+1)$이다. 이로부터 μ^*가 metric outer measure라는 사실을 상기한다면 각 $i\in\mathbb{N}$에 대해

$$\begin{aligned}\mu^*(S_{2i+1}) &= \mu^*(T_{2i}\cup S_{2i})\\ &\geq \mu^*(T_{2i}\cup S_{2i-1})\\ &= \mu^*(T_{2i})+\mu^*(S_{2i-1})\\ &\geq \cdots\\ &= \sum_{j=1}^{i}\mu^*(T_{2j})+\mu^*(S_1)\\ &\geq \sum_{j=1}^{i}\mu^*(T_{2j})\end{aligned}$$

이고, 비슷하게 $\mu^*(S_{2i})\geq\sum_{j=1}^{i}\mu^*(T_{2j-1})$임을 알 수 있다. 여기서 두 급수 $\sum_{j=1}^{i}\mu^*(T_{2j}),\sum_{j=1}^{i}\mu^*(T_{2j-1})$는 모두 $\mu^*(S)<\infty$에 의해 위로 유계이므로 MSP로부터 수렴하고, 곧 $\sum_{i=1}^{\infty}\mu^*(T_i)$는 적당한 $L\leq 2\mu^*(S)$로 수렴한다. 그렇다면 각 $i\in\mathbb{N}$에 대해 $S_i\subseteq S\setminus F$이고 $\mu^*(S\setminus F)=\mu^*(\bigcup_{j=1}^{\infty}S_j)=\mu^*(S_i\cup\bigcup_{j=i}^{\infty}T_j)\leq\mu^*(S_i)+\sum_{j=i}^{\infty}\mu^*(T_j)$에서 우변의 합이 $i\to\infty$이면 0으로 수렴하므로 $\mu^*(S\setminus F)\leq\liminf_{i\to\infty}\mu^*(S_i)\leq\limsup_{i\to\infty}\mu^*(S_i)\leq\mu^*(S\setminus F)$이다. 이는 곧 $\mu^*(S_i)\to\mu^*(S\cap F^c)$임을 뜻하므로 증명이 끝난다.
다음으로, 거리공간 (M,d)의 부분집합 $A\subseteq M$와 $p\geq 0,\delta>0$에 대해 함수 $\mathscr{H}_\delta^p:\mathscr{P}(M)\to\overline{R}_0^+$를

$$\mathscr{H}_\delta^p:A\mapsto\inf\left\{\sum_{i=1}^{\infty}(\operatorname{diam}A_i)^p\in\overline{\mathbb{R}}_0^+:\{A_i\}\text{는 }A\text{의 가산 덮개이고 각 }A_i\text{에 대해 }\operatorname{diam}A_i<\delta\text{이다.}\right\}$$

로 정의하면 이는 외측도임을 쉽게 보일 수 있다. (Hint: 명제 ? ? 의 iii의 증명과 똑같이 하면 된다.) 이제 함수 $\mathscr{H}^p:\mathscr{P}(M)\to\overline{R}_0^+$를 $\mathscr{H}^p:A\mapsto\lim_{\delta\downarrow 0}\mathscr{H}_\delta^p(A)$로 정의하고 이를 p**차원 Hausdorff 외측도 (p dimensional − outer measure)**라 한다. 여기서 임의의 $A\subseteq M$에 대해 $\mathscr{H}_\delta^p(A)$가 δ에 대해 감소함이 자명하므로 극한 $\lim_{\delta\downarrow 0}\mathscr{H}_\delta^p(A)$이 존재하여 $\mathscr{H}^p$는 well-define된다. 한편, 이렇게 정의한 $\mathscr{H}^p$는 metric outer measure가 됨을 보일 수 있다. 먼저 $\mathscr{H}^p(\emptyset)=0$임은 자명하고 임의의 $A,B\subseteq M$에 대해 $A\subseteq B$라면 $\mathscr{H}^p(A)=\lim_{\delta\downarrow 0}\mathscr{H}_\delta^p(A)\leq\lim_{\delta\downarrow 0}\mathscr{H}_\delta^p(B)=\mathscr{H}^p(B)$에서 단조성도 분명하다. 나아가, 임의의 M에 속하는 집합열 $\{A_i\}$에 대해서 MCT로부터 $\mathscr{H}^p(\bigcup_{i=1}^{\infty}A_i)=\lim_{\delta\downarrow 0}\mathscr{H}_\delta^p(\bigcup_{i=1}^{\infty}A_i)\leq\lim_{\delta\downarrow 0}\sum_{i=1}^{\infty}\mathscr{H}_\delta^p(A_i)=\sum_{i=1}^{\infty}\lim_{\delta\downarrow 0}\mathscr{H}_\delta^p(A_i)=\sum_{i=1}^{\infty}\mathscr{H}^p(A_i)$가 되어 $\sigma-$반가법성도 가지므로 $\mathscr{H}^p$는 외측도이다. 이제 임의의 $A,B\subseteq M$에 대해 $d(A,B)>0$라 하고 $\mu^*(A\cup B)\geq\mu^*(A)+\mu^*(B)$임만 보이면 된다. 이를 위해 $\delta<d(A,B)$인 임의의 $\delta>0$를 택하자. 만약 $A\cup B$의 가산 덮개이면서 덮개의 각 원소의 지름이 δ보다 작은 덮개가 존재하지 않는다면 $\mathscr{H}_\delta^p(A\cup B)=\infty$가 되어 자명하므로 이러한 조건을 만족하는 덮개가 존재한다고 가정하고, 이를 $\{C_i\}$로 쓰자. 그렇다면 각 $i\in\mathbb{N}$에 대해 $\operatorname{diam}C_i<\delta<d(A,B)$이므로 C_i는 A,B 모두와 비어있지 않은 교집합을 가지지 못하고 WLOG, $\{C_i\}$의 원소 중 A와 비어있지 않은 교집합을 가지는 것들의 모임 $\{C_{ai}\}$, B와 비어있지 않은 교집합을 가지는 것들의 모임 $\{C_{bi}\}$에 대해 $\{C_i\}=\{C_{ai}\}\sqcup\{C_{bi}\}$로 쓸 수 있다. 이로부터 $\{C_{ai}\}$와 $\{C_{bi}\}$는 각각 A,B의 가산 덮개이므로 $\mathscr{H}_\delta^p(A\cup B)\geq\sum_i(\operatorname{diam}C_i)^p=\sum_i(\operatorname{diam}C_{ai})^p+\sum_i(\operatorname{diam}C_{bi})^p\geq\mathscr{H}_\delta^p(A)+\mathscr{H}_\delta^p(B)$에서 $\mathscr{H}^p(A\cup B)\geq\mathscr{H}^p(A)+\mathscr{H}^p(B)$가 되어 증명이 끝난다.
이제 거리공간 (M,d)에서 정의된 p차원 Hausdorff 외측도 $\mathscr{H}^p$에 대해 이의 Carathéodory 확장을 표기를 남용하여 $\mathscr{H}^p$로 쓰면 앞선 결과로부터 $(M,\mathscr{B}_M,\mathscr{H}^p)$가 측도공간이 된다. 우리는 Borel 집합 $A\subseteq M$에 대해 $\inf\{p\in\mathbb{R}_0^+:\mathscr{H}^p(A)=0\}$를 A의 **Hausdorff 차원 (− dimension)**이라 하고 $\dim_H A$로 쓴다. 그 복잡한 정의에서 알 수 있듯이 일반적으로 Borel 집합의 Hausdorff 차원을 계산하는 것은 쉽지 않다. 이에 $p\in\mathbb{R}_0^+$이면 $0<\mathscr{H}^p(A)<\infty$이면 $\dim_H A=p$라는 결과가 도움이 된다. 이 또한 보이는 것은 크게 어렵지 않다. 먼저 임의의 $q>p$에 대해 $\mathscr{H}^q(A)=0$임을 보이자. 임의의 $\delta>0$를 택하면

$\mathscr{H}^p_\delta(A) \le \mathscr{H}^p(A) < \infty$이고 적당한 A의 가산 덮개 $\{A_i\}$가 존재하여 각 A_i에 대해 $\mathrm{diam}A_i < \delta$이고 $\sum_i(\mathrm{diam}A_i)^p \le \mathscr{H}^p_\delta(A)+1 \le \mathscr{H}^p(A)+1$이다. 이로부터

$$\begin{aligned}\mathscr{H}^q_\delta(A) &\le \sum_i(\mathrm{diam}A_i)^q \\ &= \sum_i(\mathrm{diam}A_i)^{q-p}(\mathrm{diam}A_i)^p \\ &\le \delta^{q-p}\sum_i(\mathrm{diam}A_i)^p \\ &\le \delta^{q-p}[\mathscr{H}^p(A)+1]\end{aligned}$$

이므로 $\mathscr{H}^q(A) = \lim_{\delta\downarrow 0}\mathscr{H}^q_\delta(A) = 0$이다. 이로부터 만약 $\dim_H A < p$이면 적당한 $q < p$가 존재하여 $\mathscr{H}^q(A) = 0 < \infty$에서 $\mathscr{H}^p(A) = 0$의 모순이 발생하므로 $\dim_H A \ge p$이다. 한편, 만약 $\dim_H A > p$이면 임의의 $q > p$에 대해 $\mathscr{H}^q(A) = 0$이 되어 이번에도 모순이 발생하므로 $\dim_H A = p$이어야 한다.

이상의 내용을 바탕으로 Cantor 집합의 Hausdorff 차원이 $p := \log 2/\log 3 \approx 0.63093$임을, 나아가 $\mathscr{H}^p(C) = 1$임을 보일 수 있다. 이를 위해 집합열 $\{C_i\}$를 정의 0. 102에서와 같이 두고 임의의 $\delta > 0$와 $1/3^k < \delta$인 충분히 큰 $k \in \mathbb{N}$를 택하자. 그렇다면 C_k는 길이가 $1/3^k$인 2^k개의 닫힌구간으로 이루어져 있는데, 이 닫힌구간들이 C의 가산 덮개임은 자명하므로 $\mathscr{H}^p_\delta(C) \le 2^k(1/3^k)^p = 1$에서 $\mathscr{H}^p(C) \le 1$이 성립한다. 이제 $\mathscr{H}^p(C) \ge 1$이라는 것만 보이면 증명이 끝나는데, 이게 조금 어렵다. 이를 위해 C의 임의의 가산 덮개를 $\{A_i\}$라 하고 $\sum_i(\mathrm{diam}A_i)^p \ge 1$임을 보이자. 사전 준비가 살짝 필요한데, 먼저 모든 $\mathbb{R}$의 부분집합은 가산개의 구간의 합집합으로 쓸 수 있으므로 (Hint: 임의의 $A \subseteq \mathbb{R}$은 그 연결성분의 서로소 합집합으로 쓸 수 있고, $\mathbb{R}$에서의 연결집합은 모두 구간이다.) WLOG, 각 A_i를 구간이라 해도 된다. 이어서, 임의의 $\varepsilon > 0$을 택하여 각 A_i를 양 옆으로 $\varepsilon/2\cdot 3^i$씩 늘려 만든 닫힌구간을 I_i라 하면 $\{I_i\}$는 여전히 C의 가산 덮개이고 $\sum_i(\mathrm{diam}I_i)^p = \sum_i(\mathrm{diam}A_i)^p + \varepsilon$이다. 또한, $\{I_i^\circ\}$도 명백히 C의 덮개이므로 C가 compact하다는 점으로부터 적당한 부분덮개를 가지는데, WLOG, 그 부분덮개를 $\{I_1^\circ, \cdots, I_k^\circ\}$라 할 수 있다. 마지막으로, 필요하다면 $I_1, \cdots, I_k$에서 필요없는 양 끝 부분을 제거하여 WLOG, 각 $i \le k$에 대해 적당한 $j_i, j_i' \in \mathbb{N}$가 존재하여 각각 $C_{j_i}, C_{j_i'}$의 연결성분인 닫힌구간 J_i, J_i'과 적당한 $[0,1]\setminus C$의 연결성분인 열린구간 K_i로써 $I_i = J_i \sqcup K_i \sqcup J_i'$과 같이 쓸 수 있다. 이로써 준비는 끝났다. 이상을 종합하면 각 $i \le k$에 대해 $\mathrm{diam}J_i, \mathrm{diam}J_i' \le \mathrm{diam}K_i$이므로 (Hint: 그림을 그려보자.)

$$\begin{aligned}(\mathrm{diam}I_i)^p &= (\mathrm{diam}J_i + \mathrm{diam}K_i + \mathrm{diam}J_i')^p \\ &\ge \left[\frac{3(\mathrm{diam}J_i + \mathrm{diam}J_i')}{2}\right]^p \\ &= 2\left(\frac{\mathrm{diam}J_i + \mathrm{diam}J_i'}{2}\right)^p \\ &\ge (\mathrm{diam}J_i)^p + (\mathrm{diam}J_i')^p\end{aligned}$$

이고, 이로부터 $\sum_{i=1}^k[(\mathrm{diam}J_i)^p + (\mathrm{diam}J_i')^p] \le \sum_{i=1}^k(\mathrm{diam}I_i)^p \le \sum_i(\mathrm{diam}A_i)^p + \varepsilon$이다. 이제 이러한 과정을 유한번 반복하면 적당한 $l \in \mathbb{N}$이 존재하여 C_l의 연결성분인 닫힌구간 $L_1, \cdots, L_{k'}$에 대해 이들이 C의 덮개이면서 $\sum_{i=1}^{k'}(\mathrm{diam}L_i)^p \le \sum_i(\mathrm{diam}A_i)^p + \varepsilon$이도록 할 수 있다. 그런데 $L_1, \cdots, L_{k'}$이 C의 덮개이기 위해서는 C_l의 모든 연결성분이 적어도 한 번은 등장해야 하므로 $k' \ge 2^l$이 되어 $\sum_i(\mathrm{diam}A_i)^p + \varepsilon \ge \sum_{i=1}^{2^l}(\mathrm{diam}L_i)^p = 1$이고, 여기서 ε이 임의의 양수였다는 사실을 떠올리면 증명이 끝난다.

9 1장에서 함수의 절대연속은 $\mathbb{R}^n$ 에서 정의된 함수에 대해서만 정의되었는데, 이 정의를 부분집합 $A \subseteq \mathbb{R}^n$ 에서 정의된 함수 $f : A \to \mathbb{R}$ 로 어떻게 확장할 것인지에 대해서 한 번 쯤은 생각해 볼 만하다. 여러 방법이 있겠지만, 가장 쉽고 자연스러운 방법은 $\mathbb{R}^n \setminus A$ 에서의 함숫값을 0으로 하여 f 를 $\mathbb{R}^n$ 으로 확장하는 방법이다. Cantor 함수에 대해서도 이를 적용하여 임의의 $x \in \mathbb{R} \setminus [0,1]$ 에 대해서는 $c(x) = 0$ 으로 생각한다.

10 가정으로부터 $M \le N_x$ 인데, 먼저 $M < N_x$ 인 경우를 생각하면 $x_M = 0, y_M = 2$ 에서 $z \le y$ 임이 분명하고, $x_{N_x} = 1$ 에서 $x \le z$ 임이 분명하므로 곧 $z \in [x, y]$ 이다. 한편, $M = N_x$ 인 경우에도 $x_M = 1, y_M = 2$ 에서 $z \in [x, y]$ 임이 분명하다.

11 Euler가 이를 이용하여 어떻게 $\zeta(2)$ 의 값을 구하였는지 살펴보자. 우선 사인 함수의 Taylor 전개로부터 임의의 $x \in \mathbb{R}$ 에 대해 $\sin \pi x = \sum_{i=1}^{\infty} (-1)^i (\pi x)^{2i-1}/(2i-1)!$ 인데, 우리가 증명한 바에 따르면 $\sin \pi x = \pi x \prod_{i=1}^{\infty} (1 - x^2/i^2)$ 이다. 이제 각 표현에서 x^3 의 계수를 비교하면 $-\pi^3/6 = -\pi \sum_{i=1}^{\infty} 1/i^2$ 이 되어 $\zeta(2) = \pi^2/6$ 임을 바로 알 수 있다.

12 현대 수학에서는 이런 '함수' 를 **분포 (distribution)** 라 하여 (당연히, 이는 확률론에서의 분포와는 다르다.) 어떤 함수를 실수로 대응시키는 일종의 연산자로 생각하여 다룬다. 예컨대 $\int_{\mathbb{R}} \delta = 1$ 을 정말 δ 라는 함수를 적분하여 1을 얻는다고 생각하는 것이 아니라 δ 가 1이라는 상수함수를 1로 mapping 한다고 보는 것이다. 그리고 이런 아이디어를 확장하여 분포의 미분과 각종 연산을 정의할 수 있다. 물론, 분포에 대한 이론을 엄밀히 다루기 위해서는 꽤나 고급 수학이 필요하다. 분포가 아무런 함수나 실수로 대응시킬 수 있으면 좋겠지만, 분포가 실수로 대응시킬 수 있는 함수, 즉 분포와 곱해서 '적분' 할 수 있는 함수가 제한되어 있으므로 이런 '적분' 이 가능한 함수들의 공간을 먼저 구성하고 (이를 시험함수공간이라 한다.) 이에 적당한 위상을 주는 등의 각종 준비가 필요하다. 흥미가 생긴다면 [12] 를 참고하기 바란다.

References

[1] Jerrold E. Marsden and Michael J. Hoffman, *Elementary Classical Analysis 2nd Edition*, W. H. Freeman, 1993.

[2] Patrick Billingsley, *Probability and Measure 3rd Edition*, Wiley, 1995.

[3] 이인석,『선형대수와 군』개정판, 서울대학교출판문화원, 2015.

[4] Kenneth Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, Wiley, 2014.

[5] Kenneth Falconer, *The Geometry of Fractal Sets*, Cambridge University Press, 1995.

[6] Steven G. Johnson, *Notes on The Equivalence of Norms*, MIT, 2012.

[7] Mohammed Dahleh, Munther A. Dahleh, and George Verghese, *Lectures on Dynamic Systems and Control*, MIT, 2011.

[8] N. H. Bingham, *Analytic Number Theory*, Imperial College London, 2014.

[9] John K. Hunter, *Introduction to Analysis*, University of California Davis, 2015.

[10] Harold P. Boas, *Lecture Notes on Several Complex Variables*, Texas A&M University, 2013.

[11] Jordan Bell, *Hausdorff measure*, University of Toronto, 2014.

[12] José Luis Rodrigo, *Advanced Real Analysis*, University of Warwick, 2010

[13] Mathematics Stack Exchange, Available: https://math.stackexchange.com.

[14] Wolfram MathWorld, Available: http://mathworld.wolfram.com.

[15] Wikipedia, Available: https://en.wikipedia.org.

[16] 나무위키, Available: https://namu.wiki.

Part I

Probability and Measure

Chapter 1

Measure Theory

Abstract 측도란, $\mathbb{R}^2$에서의 넓이나 $\mathbb{R}^3$에서의 부피를 임의의 공간으로 일반화한 개념으로, 그 용어가 뜻하는 사전적인 의미와 같이 어떤 집합을 재는(measure) 함수이다. 언듯 보기에 이러한 추상적인 개념으로 수리통계학을 시작하는 것이 부자연스러워 보일 수 있다. 그럼에도 불구하고, 우리가 이 책을 측도론으로 시작하는 것에는 크게 두 가지 이유가 있다. 우선 첫째로 확률이 그 자체로 측도이기 때문이다. 러시아의 대수학자 Kolmogorov가 정립한 현대의 확률론에서는 확률을 '사건'이라는 집합에 0과 1사이의 수를 부여하는 측도로 파악한다. 따라서 측도는 확률의 본질이 되어 이를 알지 아니하고서는 확률에 대한 엄밀한 이론 전개가 어렵다. 둘째 이유는 보다 더 일반적인 것으로, 우리가 기존에 사용하던 적분론을 조금 손볼 필요가 있기 때문이다. 우리가 여태까지 사용해오던 적분은 Riemann 적분론인데, Riemann 적분론에서는 적분과 극한의 순서를 바꾸기 위해 균등수렴과 같은 상당히 강한 조건들이 필요하다. 이렇게 적분과 극한이 서로 매끄럽게 상호작용하지 못한다는 점은 Riemann 적분론의 단점으로, 이를 보완해 줄 수 있는 새로운 적분론이 필요한데, 그것이 바로 Lebesgue 적분론이다. 그리고 Lebesgue 적분론이 측도론을 기초로 하기에 우리는 측도론을 배워야 한다.

1.1 Measurable Spaces

본 장 전반에 걸쳐, 우리는 Lebesgue 측도를 구성할 것이다. Lebesgue 측도는 우리가 직관적으로 파악하는 넓이나 부피에 대응하는 $\mathbb{R}^n$ 위의 측도로 $n = 3$인 경우에 점이나 선에는 0의 측도를, unit cube에는 1의 측도를, 유계가 아닌 cube에는 무한대의 측도를 부여하는 이상적인 측도이다. 그러나 겉보기와는 달리, Lebesgue 측도를 구성하는 일은 생각보다 만만치 않다. 이는 놀랍게도 모든 집합에 이러한 측도를 부여하는 것이 일반적으로는 불가능하기 때문이다. 즉, $\mathbb{R}^n$의 부분집합이라고 해서 모두 잴 수 있는 것이 아니다. 이런 '잴 수 없

는' 집합은 분명히 존재하며, 여기서는 다루지 않겠지만 이러한 집합을 직접 구성할 수도 있다.[1]따라서 우리는 Lebesgue 측도를 구성하기에 앞서 우선 측도를 부여할 $\mathbb{R}^n$의 부분집합들을 먼저 선정해야 한다. 그렇다면 과연 측도를 부여할 부분집합들을 어떻게 골라야 할까? 너무 많이 고르면 우리가 원하는 Lebesgue 측도의 이상적인 성질을 유지하기 힘들어지고, 그렇다고 너무 적게 고르면 실용성이 떨어지는 측도가 되고 만다. Lebesgue 측도를 구성한다고 함은, 위의 질문에 대한 아름다운 답을 찾아나가는 과정이라 볼 수 있다. 그 첫 번째 단계로 이번 절에서는 집합들의 모임인 집합족을 다루는 법을 배운다.

Definition 1.1 공집합이 아닌 집합 X에 대해 $\mathscr{A}$를 X의 부분집합의 모임이라 하고 비어있지 않다고 하자. 만약 $\mathscr{A}$가

i. 임의의 $A, B \in \mathscr{A}$에 대해 $A \cup B \in \mathscr{A}$이다.
ii. 임의의 $A \in \mathscr{A}$에 대해 $A^c \in \mathscr{A}$이다.

를 만족하면 이때의 $\mathscr{A}$를 X 위의 **대수(algebra)** 혹은 **체(field)**라 한다. 나아가, 만약 $\mathscr{A}$에 속하는 임의의 수열 $\{A_i\}$에 대해 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$이면 이때의 $\mathscr{A}$를 X 위의 **σ−대수(σ−algebra)** 혹은 **σ−체(σ−field)**라 한다.

Proposition 1.2 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 다음이 성립한다.

i. $\emptyset, X \in \mathscr{A}$.
ii. 임의의 $A, B \in \mathscr{A}$에 대해 $A \cap B, A \setminus B \in \mathscr{A}$이다.

나아가, 만약 $\mathscr{A}$가 σ−대수이면 $\mathscr{A}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $\bigcap_{i=1}^{\infty} A_i \in \mathscr{A}$이다.

Proof ii. 이는 $A \cap B = (A^c \cup B^c)^c$와 $A \setminus B = A \cap B^c = (A^c \cup B)^c$에서 자명하다.

i. 집합족 $\mathscr{A}$가 비어있지 않으므로 적어도 하나의 원소 $A \in \mathscr{A}$를 택할 수 있다. 그렇다면 ii로부터 $\emptyset = A \setminus A \in \mathscr{A}$이고 $X = A \cup A^c \in \mathscr{A}$이다.

한편, $\mathscr{A}$가 σ−대수인 경우를 생각한다면, $\mathscr{A}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $\{A_i^c\}$도 $\mathscr{A}$에 속하므로 $\bigcap_{i=1}^{\infty} A_i = (\bigcup_{i=1}^{\infty} A_i^c)^c \in \mathscr{A}$이다. □

정의로부터, 대수는 교집합, 합집합, 차집합, 여집합에 대해 모두 닫혀있으며, 수학적 귀납법을 사용하면 이러한 연산을 유한번 적용하는 경우에도 여전히 닫혀있음을 쉽게 알 수 있다. 이에 더하여, σ−대수의 경우 이러한 연산을 가산번 반복 적용하는 경우에도 닫혀있다. 우리가 아는 집합의 연산이 대부분이 위의 네 가지 연산의 조합으로 이루어진다는 사실을 생각해보면, 대수는 굉장히 큰 모임이라고 생각할 수 있다. 실제로 대수에 속한 집합들을 연산하여 얻는 어지간한 결과들은 모두 그 대수 안에 다시 속한다.

Definition 1.3 공집합이 아닌 집합 X 위의 σ−대수 $\mathscr{A}$에 대해 tuple $(X, \mathscr{A})$를 **가측공간(measurable space)**이라 한다. 또한, $\mathscr{A}$에 속하는 임의의 집합을 **($\mathscr{A}$−)가측집합(($\mathscr{A}$−)measurable set)**이라 한다.

갑자기 가측공간의 정의가 덜컥 주어지니 당황스럽기 그지없을 것이다. 아직 우리는 측도가 무엇인지도 확실히 모르므로 이는 당연한 반응이다. 지금으로서는, 그냥 위와 같은 정의를 만족하는 tuple을 가측공간이라 부른다는 사실만 받아들이고, 일단 용어에 익숙해지도록 하자. 이후에 Lebesgue 측도를 본격적으로 구성하게 되면 가측공간은 자연스럽게 주인공으로 떠오르게 된다. 한편, 가측공간을 정의함에 있어서 측도라는 개념이 직접적으로 필요하지 않다는 점도 기억해둘 필요가 있다. (이후 가측공간에 측도를 더해 만들어지는 측도공간과 가측공간을 혼동하면 안된다.) 가측공간의 정의에서 유일하게 필요한 개념은 $\sigma-$대수 뿐이다. 이는 $\sigma-$대수가 측도론에서 중요한 역할을 한다는 점을 시사하기도 한다.

Definition 1.4 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 $\mathscr{C}$를 포함하는 최소의 $\sigma-$대수를 **$\mathscr{C}$가 생성하는 $\sigma-$대수($\sigma-$algebra generated by $\mathscr{C}$)**라 하고 $\sigma(\mathscr{C})$로 쓴다. 여기서 $\mathscr{C}$는 $\sigma(\mathscr{C})$의 **생성자(generator)**라 한다.

Proposition 1.5 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 이가 생성하는 $\sigma-$대수 $\sigma(\mathscr{C})$는 유일하게 존재한다. 따라서, $\sigma(\mathscr{C})$는 well-define된다.

Proof $\sigma(\mathscr{C})$가 존재하기만 한다면, 이의 유일성은 그 정의로부터 자명하다. 존재성을 보이기 위해 Σ를 $\mathscr{C}$를 포함하는 모든 $\sigma-$대수의 모임이라 하면 $\mathscr{P}(X)$가 $\mathscr{C}$를 포함하는 $\sigma-$대수임이 분명하므로 Σ는 비어있지 않다. 이제 이의 교집합인 $\bigcap\Sigma$가 $\sigma-$대수라 주장한다. 이를 보이기는 어렵지 않다. 정의 ? ? 의 조건을 하나하나 따져보면 되는데, 연습삼아 두 번째 조건을 보자. 임의의 $A \in \bigcap\Sigma$에 대해 A는 Σ에 속하는 모든 $\sigma-$대수에 속하므로 곧 A^c도 Σ에 속하는 모든 $\sigma-$대수에 속하고, 이는 곧 $A^c \in \bigcap\Sigma$임을 뜻한다. 다른 조건들도 이와 같이 쉽게 따져볼 수 있다. 그렇다면 $\bigcap\Sigma$이 바로 $\sigma(\mathscr{C})$이다. □

당연히, 어떤 $\sigma-$대수를 생성하는 생성자는 유일하지 않다. 즉, 같은 $\sigma-$대수를 생성하는 서로다른 다양한 집합족이 있을 수 있다.

Definition 1.6 공집합이 아닌 집합 X에 대해 $\mathscr{P}$를 X의 부분집합의 모임이라 하고 비어있지 않다고 하자. 만약 임의의 $A, B \in \mathscr{P}$에 대해 $A \cap B \in \mathscr{P}$가 성립하면 이때의 $\mathscr{P}$를 X 위의 **$\pi-$system**이라 한다.

Definition 1.7 공집합이 아닌 집합 X에 대해 $\mathscr{L}$을 X의 부분집합의 모임이라 하자. 만약 $\mathscr{L}$이

i. $X \in \mathscr{L}$.
ii. 임의의 $A \in \mathscr{L}$에 대해 $A^c \in \mathscr{L}$이다.
iii. 집합족 $\mathscr{L}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 $\bigsqcup_{i=1}^{\infty} A_i \in \mathscr{L}$이다.

를 만족하면 이때의 $\mathscr{L}$을 X 위의 **$\lambda-$system**, **$d-$system** 혹은 **Dynkin class**라 한다.[2]

정의로부터, π−system과 λ−system은 σ−대수보다 더 단순한 구조로 이루어져 있다. π−system은 교집합에 대해서만 닫혀 있고, λ−system은 여집합과 가산번의 서로소 합집합에 대해서만 닫혀 있다. 물론 이전과 같이 수학적 귀납법을 사용하면 각각 유한번의 교집합과 여집합에 대해서도 닫혀 있음을 쉽게 알 수 있지만 그래도 σ−대수의 구조에 비하면 π−system과 λ−system의 구조는 한참 단순하다. 그러나 이 둘과 σ−대수는 Dynkin의 π−λ 정리에 의해 아름답게 연결된다.

Definition 1.8 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 $\mathscr{C}$를 포함하는 최소의 λ−system을 **$\mathscr{C}$가 생성하는 λ−system(− generated by $\mathscr{C}$)**라 하고 $\lambda(\mathscr{C})$로 쓴다. 여기서 $\mathscr{C}$는 $\lambda(\mathscr{C})$의 **생성자(generator)**라 한다.

Proposition 1.9 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 이가 생성하는 λ−system $\lambda(\mathscr{C})$는 유일하게 존재한다. 따라서, $\lambda(\mathscr{C})$는 well-define된다.

Proof $\lambda(\mathscr{C})$가 존재하기만 한다면, 이의 유일성은 그 정의로부터 자명하다. 존재성을 보이기 위해 Λ를 $\mathscr{C}$를 포함하는 모든 λ−system의 모임이라 하면 $\mathscr{P}(X)$가 $\mathscr{C}$를 포함하는 λ−system임이 분명하므로 Λ는 비어있지 않다. 이제 이의 교집합인 $\bigcap\Lambda$가 λ−system이라 주장한다. 이를 보이기는 어렵지 않다. 정의 ? ? 의 조건을 하나하나 따져보면 되는데, 연습삼아 첫 번째 조건을 보자. 임의의 $A \in \bigcap\Lambda$에 대해 A는 Λ에 속하는 모든 λ−system에 속하므로 곧 A^c도 Λ에 속하는 모든 λ−system에 속하고, 이는 곧 $A^c \in \bigcap\Lambda$임을 뜻한다. 다른 조건들도 이와 같이 쉽게 따져볼 수 있다. 그렇다면 $\bigcap\Lambda$이 바로 $\lambda(\mathscr{C})$이다. □

Lemma 1.10 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{A}$가 π−system인 동시에 λ−system이라면 이는 σ−대수이다.

Proof 집합족 $\mathscr{A}$가 가산 합집합에 대해 닫혀있음을 보이면 충분하다. 즉, $\mathscr{A}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$임을 보이면 된다. 이를 위해 집합열 $\{B_i\}$를 $B_i := A_i \setminus \bigcup_{j=1}^{i-1} A_j = A_i \cap \bigcap_{j=1}^{i-1} A_j^c$로 정의하면 $\mathscr{A}$가 유한번의 교집합과 여집합에 대해 닫혀있으므로 이는 명백히 $\mathscr{A}$에 속하는 서로소인 집합열이다. 또한, $\bigcup_{i=1}^{\infty} A_i = \bigsqcup_{i=1}^{\infty} B_i$이므로 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$가 되어 원하는 결론을 얻는다. □

Theorem 1.11 (Dynkin's π-λ theorem) 공집합이 아닌 집합 X 위의 π−system $\mathscr{P}$에 대해 $\sigma(\mathscr{P}) = \lambda(\mathscr{P})$이다.

Proof $\lambda(\mathscr{P}) \subseteq \sigma(\mathscr{P})$임이 분명하므로 역의 포함관계만 보이면 되고, 다시 위의 보조정리로부터 $\lambda(\mathscr{P})$가 π−system임을 보이는 것으로 충분하다. 이를 위해 임의의 $A \in \lambda(\mathscr{P})$를 택하여 집합 $\lambda_A = \{B \subseteq X : A \cap B \in \lambda(\mathscr{P})\}$를 생각하고 $\lambda(\mathscr{P}) \subseteq \lambda_A$임을 보이자. 우선 $A \in \lambda_A$이므로 $X \in \lambda_A$임은 분명하고, 임의의 $B \in \lambda_A$에 대해 $A \cap B^c = A \setminus (A \cap B)$이므로 $B^c \in \lambda_A$이

다. 마지막으로 λ_A에 속하는 임의의 서로소인 집합열 $\{B_i\}$에 대해 $\{A \cap B_i\}$가 $\lambda(\mathscr{P})$에 속하는 서로소인 집합열이므로 $\bigsqcup_{i=1}^{\infty} B_i \in \lambda_A$가 성립하여 λ_A가 $\lambda-$system임을 안다. 이상의 결론이 임의의 $A \in \lambda(\mathscr{P})$에 대해 성립함을 상기한다면, 임의의 $B \in \mathscr{P}$에 대해 λ_B도 $\lambda-$system인데, λ_B의 정의에서 $\mathscr{P} \subseteq \lambda_B$임이 분명하므로 곧 $\lambda(\mathscr{P}) \subseteq \lambda_B$이고, 따라서 $A \in \lambda_B$이다. 그렇다면 다시 λ_B의 정의에서 $B \in \lambda_A$이며, 곧 $\mathscr{P} \subseteq \lambda_A$가 되고, 증명은 이로써 충분하다. □

앞서 대수가 꽤나 큰 집합족이라는 사실을 설명한 바 있는데, 이는 마치 양날의 검과 같다. 우선 그 크기가 충분히 커서 우리가 필요로 하는 대부분의 집합들을 포함시킬 수 있으며, 어지간한 연산에 대해 닫혀있다는 점은 분명 대수의 장점이다. 그러나 대수의 이런 거대한 크기 때문에 때로는 이를 직접 다루기 힘든 경우도 있다. 이러한 어려움을 극복하기 위해 중간다리의 역할을 하는 구조가 필요한데, 바로 semi-algebra이다.

Definition 1.12 공집합이 아닌 집합 X에 대해 $\mathscr{A}$를 X의 부분집합의 모임이라 하자. 만약 $\mathscr{A}$가

i. $X \in \mathscr{A}$.
ii. 임의의 $A, B \in \mathscr{A}$에 대해 $A \cap B \in \mathscr{A}$이다.
iii. 임의의 $A \in \mathscr{A}$에 대해 A^c를 $\mathscr{A}$에 속하는 서로소인 집합들의 유한 합집합으로 쓸 수 있다.

를 만족하면 이때의 $\mathscr{A}$를 X 위의 **semi-algebra**라 한다.

Proposition 1.13 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$에 대해 다음이 성립한다.

i. $\emptyset \in \mathscr{A}$.
ii. 임의의 $A, B \in \mathscr{A}$에 대해 $A \setminus B$를 $\mathscr{A}$에 속하는 서로소인 집합들의 유한 합집합으로 쓸 수 있다.

Proof i. $X \in \mathscr{A}$이므로 $\emptyset = X^c$는 적당한 서로소인 집합들 $B_1, \cdots, B_k \in \mathscr{A}$에 대해 이들의 서로소 합집합으로 쓸 수 있는데, 이는 곧 $B_1 = \cdots = B_k = \emptyset$임을 뜻하고, 따라서 $\emptyset \in \mathscr{A}$이다.

ii. 적당한 서로소인 집합들 $C_1, \cdots, C_k \in \mathscr{A}$에 대해 $B^c = \bigsqcup_{i=1}^{k} C_i$이고 여기서 각 $i \le k$에 대해 $A \cap C_i \in \mathscr{A}$이며 이들이 서로소인데, $A \setminus B = A \cap B^c = A \cap \bigsqcup_{i=1}^{k} C_i = \bigsqcup_{i=1}^{k} (A \cap C_i)$가 되어 증명이 끝난다. □

대수와 semi-algebra의 가장 큰 차이점은 semi-algebra의 경우 더 이상 여집합에 대해 닫혀있지 않다는 점이다. 위의 정의에서 볼 수 있듯이 여집합에 대해 닫혀있어야 한다는 대수의 조건이 조건 iii과 같이 약화되었다. 그러나 이번에도, semi-algebra와 대수를 연결해주는 정리가 존재한다.

Definition 1.14 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 $\mathscr{C}$를 포함하는 최소의 대수를 **$\mathscr{C}$가 생성하는 대수(algebra generated by $\mathscr{C}$)**라 하고 $A(\mathscr{C})$로 쓴다. 여기서 $\mathscr{C}$는 $A(\mathscr{C})$의 **생성자(generator)**라 한다.

Proposition 1.15 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 이가 생성하는 대수 $A(\mathscr{C})$는 유일하게 존재한다. 따라서, $A(\mathscr{C})$는 well-define된다.

PROOF $A(\mathscr{C})$가 존재하기만 한다면, 이의 유일성은 그 정의로부터 자명하다. 존재성을 보이기 위해 Γ를 $\mathscr{C}$를 포함하는 모든 대수의 모임이라 하면 $\mathscr{P}(X)$가 $\mathscr{C}$를 포함하는 대수임이 분명하므로 Γ는 비어있지 않다. 이제 이의 교집합인 $\bigcap\Gamma$가 대수라 주장한다. 이를 보이기는 어렵지 않다. 정의 **? ?** 의 조건을 하나하나 따져보면 되는데, 연습삼아 두 번째 조건을 보자. 임의의 $A \in \bigcap\Gamma$에 대해 A는 Γ에 속하는 모든 대수에 속하므로 곧 A^c도 Γ에 속하는 모든 대수에 속하고, 이는 곧 $A^c \in \bigcap\Gamma$임을 뜻한다. 다른 조건들도 이와 같이 쉽게 따져볼 수 있다. 그렇다면 $\bigcap\Gamma$가 바로 $A(\mathscr{C})$이다. □

Theorem 1.16 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$에 대해 $A(\mathscr{A})$는 $\mathscr{A}$에 속하는 서로소인 집합들의 유한 합집합으로 만들 수 있는 모든 집합들의 집합이다.

PROOF 집합 $\mathscr{C} = \{\bigsqcup_{i=1}^k A_i \subseteq X : A_1, \cdots, A_k \in \mathscr{A}\}$를 생각하면 $\mathscr{C} \subseteq A(\mathscr{A})$임은 분명하므로 역의 포함관계를 보이는 것으로 충분하다. 이를 위해, $\mathscr{C}$가 대수임을 보이자. 만약 이를 보일 수 있으면 $\mathscr{A} \subseteq \mathscr{C}$임이 분명하므로 증명이 끝난다. 이하의 증명에서 A, B는 $\mathscr{C}$의 임의의 두 원소로 적당한 서로소인 집합 $A_1, \cdots, A_k \in \mathscr{A}$와 $B_1, \cdots, B_l \in \mathscr{A}$이 존재하여 $A = \bigsqcup_{i=1}^k A_i, B = \bigsqcup_{j=1}^l B_j$로 쓸 수 있다고 하자. 먼저 $A_1 \cap B_1, \cdots, A_k \cap B_l \in \mathscr{A}$이 모두 서로소이고 $A \cap B = \bigsqcup_{i=1}^k \bigsqcup_{j=1}^l (A_i \cap B_j)$라는 점에서 $\mathscr{C}$가 교집합에 대해 닫혀있음을 안다. 다음으로, $A^c = (\bigsqcup_{i=1}^k A_i)^c = \bigcap_{i=1}^k A_i^c$이고 각 $i \le k$에 대해 A_i^c는 $\mathscr{A}$에 속하는 서로소인 집합들의 유한 합집합으로 쓸 수 있으므로 명백히 $\mathscr{C}$에 속하여 곧 $A^c \in \mathscr{C}$이고 $\mathscr{C}$가 여집합에 대해 닫혀있음도 안다. 또한, 만약 A, B가 서로소라면 $A_1 \cdots, A_k, B_1, \cdots, B_l$이 모두 서로소가 되어 $A \sqcup B = \bigsqcup_{i=1}^k A_i \sqcup \bigsqcup_{j=1}^l B_j$에서 $\mathscr{C}$는 서로소 합집합에 대해서도 닫혀있다. 이제 마지막으로 A, B가 서로소가 아니라 하더라도 $A \cup B = (A \setminus B) \sqcup (B \setminus A) \sqcup (A \cap B)$에서 $A \setminus B = A \cap B^c, B \setminus A = B^c \cap A, A \cap B$가 모두 $\mathscr{C}$에 속하므로 $A \cup B \in \mathscr{C}$이고 곧 $\mathscr{C}$가 합집합에 대해 닫혀있음을 안다. 이로써 $\mathscr{C}$가 대수의 조건을 모두 만족함을 안다. □

1.2 Measures and Premeasures

지금까지 우리는 다양한 구조의 집합족을 다루는 방법들과 서로다른 구조를 연결하는 몇몇 유용한 정리들을 익혔다. 이번 절에서는 측도가 무엇인지 알아보고 이전 절에서 덩그러니

던져진 가측공간에 측도를 부여하는 과정을 알아보도록 한다. 그리고 이로써, 우리의 첫 번째 Lebesgue 측도의 구성이 본격적으로 시작된다.

Definition 1.17 가측공간 $(X, \mathscr{A})$에 대해 함수 $\mu : \mathscr{A} \to \overline{\mathbb{R}}_0^+$가

i. $\mu(\emptyset) = 0$.
ii. ($\sigma-$가법성) $\sigma-$대수 $\mathscr{A}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 $\mu(\bigsqcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$이다.

를 만족하면 이때의 μ를 $\mathscr{A}$ 위의 **측도(measure)**라 하고, triple $(X, \mathscr{A}, \mu)$를 **측도공간(measure space)**이라 한다.

보통 측도는 그리스 문자 $\mu, \nu, \cdots$로 표기한다. 한편, 위의 정의를 유심히 보면 '측정'의 본질을 측도라는 개념에 어떻게 수학적으로 담아내었는지 살펴볼 수 있다. 우선, 측정의 결과는 반드시 0 이상의 실수이거나 무한대이다. 둘째로, 비어있는 집합을 측정한 결과는 반드시 0이다. 마지막으로 서로 겹치지 않는 가산개의 집합을 측정한 결과는 각각의 집합을 측정한 후에 이를 모두 더한 것과 일치한다. (이때, 더하는 항이 모두 양수이므로 덧셈의 순서는 문제가 되지 않는다.) 특히 마지막 성질을 흔히 **$\sigma-$가법성($\sigma-$additivity)** 혹은 **가산가법성(countable additivity)**이라 부르며 어떤 함수가 측도인지를 결정하는 데 있어 가장 강한 조건으로 작용한다. 한편, 서로소인 유한개의 집합들에 공집합을 dummy term으로 추가하여 이를 서로소인 집합열로 만들고 이에 위의 정의의 두 조건을 적용하면 유한개의 서로소인 집합들에 대해서도 $\sigma-$가법성과 비슷한 성질이 성립함을 알 수 있는데, 이를 **(유한)가법성((finite) additivity)**이라 한다. 방금 설명한 것과 같이 측도는 $\sigma-$가법성과 가법성을 모두 가진다.

이제 측도의 정의를 알았으니 본격적으로 Lebesgue 측도의 구성을 시작해보자. Lebesgue 측도의 첫 번째 구성은 이른바 '확장정리' 들을 이용한다. 이전 절에서 언급한 바와 같이 $\sigma-$대수가 너무 거대한 까닭에 그 위에 곧바로 측도를 구성하기는 어려우니, 일단 그 구조가 간단한 $\mathbb{R}^n$ 위의 semi-algebra 하나를 적당히 택하여 이 위에 궁극적으로 측도가 될 함수를 정의하고, 이를 몇 단계를 거쳐 확장함으로써 최종적으로 Lebesgue 측도를 구성해내는 것이다. 여기서 처음에 semi-algebra 위에 정의되는 함수는 측도가 아님에 유의하자. (측도가 되려면 정의역이 $\sigma-$대수이어야 한다.) 이렇게 확장을 거치기 전의 '궁극적으로 측도가 될' 함수를 premeasure라 한다.

Definition 1.18 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$에 대해 함수 $\rho : \mathscr{A} \to \overline{\mathbb{R}}_0^+$가

i. $\rho(\emptyset) = 0$.
ii. ($\sigma-$가법성) Semi-algebra $\mathscr{A}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 $\bigsqcup_{i=1}^{\infty} A_i \in \mathscr{A}$이면 $\mu(\bigsqcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$이다.

를 만족하면 이때의 ρ를 $\mathscr{A}$ 위의 **premeasure**라 한다.

보통 premeasure는 그리스 문자 $\rho, \cdots$로 표기한다. 위의 정의에서 알 수 있듯이 premeasure는 $\sigma-$대수가 아닌 semi-algebra 위에서 보다 일반적으로 정의된다는 점을 제외하면 측도와 거의 같은 개념이며 당연히 측도는 premeasure이다. 다만, 더 이상 정의역이 $\sigma-$대수가 아니므로 $\sigma-$가법성의 경우 서로소 합집합의 결과가 정의역에 다시 포함된다는 추가적인 조건이 붙어야 한다. (일반적으로 정의역이 $\sigma-$대수가 아니어서 이러한 조건이 필수적인 경우, $\sigma-$가법성은 이러한 조건을 암묵적으로 포함하는 것으로 본다. 이는 가법성의 경우에도 마찬가지이다.) 이번 절의 남은 부분은 모두 premeasure와 측도의 성질에 관한 내용들로, 측도는 premeasure이므로 premeasure의 성질은 모두 측도에 대해서도 동일하게 성립한다는 점에 유의하기 바란다.

Lemma 1.19 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$에 대해 함수 $f:\mathscr{A}\to\overline{\mathbb{R}}_0^+$가 가법성을 갖는다고 하면 임의의 집합 $A, B_1, \cdots, B_k \in \mathscr{A}$에 대해 다음이 성립한다.

i. 만약 $B_1, \cdots, B_k$가 서로소이고 $\bigsqcup_{i=1}^k B_i \subseteq A$이면 $\sum_{i=1}^k f(B_i) \le f(A)$이다.
ii. 만약 $A \subseteq \bigcup_{i=1}^k B_i$이면 $f(A) \le \sum_{i=1}^k f(B_i)$이다.

나아가, 만약 f가 $\sigma-$가법성을 가지면 ii보다 일반적인 명제가 성립한다. 즉, 임의의 $A \in \mathscr{A}$와 $\mathscr{A}$에 속하는 임의의 집합열 $\{B_i\}$에 대해 다음이 성립한다.

ii° 만약 $A \subseteq \bigcup_{i=1}^\infty B_i$이면 $f(A) \le \sum_{i=1}^\infty f(B_i)$이다.

Proof i. 명제 ? ? 의 ii로부터 적당한 서로소인 $C_1, \cdots, C_l \in \mathscr{A}$이 존재하여 $A \setminus \bigsqcup_{i=1}^k B_i = A \setminus B_1 \setminus \cdots \setminus B_k = \bigsqcup_{j=1}^l C_j$이고, 따라서

$$\begin{aligned} f(A) &= f\left(\bigsqcup_{i=1}^k B_i \sqcup \left(A \setminus \bigsqcup_{i=1}^k B_i\right)\right) \\ &= f\left(\bigsqcup_{i=1}^k B_i \sqcup \bigsqcup_{j=1}^l C_j\right) \\ &= \sum_{i=1}^k f(B_i) + \sum_{j=1}^l f(C_j) \\ &\ge \sum_{i=1}^k f(B_i) \end{aligned}$$

이다.

ii. 각 $i \le k$에 대해 집합 A_i를 $A_i := A \cap (B_i \setminus \bigcup_{j=1}^{i-1} B_j) = (A \cap B_i) \setminus B_1 \setminus \cdots \setminus B_{i-1}$로 두면 $A_1, \cdots, A_k$는 서로소이다. 또한 각 $i \le k$에 대해 $A_i \subseteq B_i$이며 명제 ? ? 의 ii로부터 적당한 서로소인 $C_{i1}, \cdots, C_{il_i} \in \mathscr{A}$가 존재하여 $A_i = \bigsqcup_{j=1}^{l_i} C_{ij}$이므로 방금 보인 i로부터

$$\begin{aligned} f(A) &= f\left(\bigsqcup_{i=1}^{k}\bigsqcup_{j=1}^{l_i} C_{ij}\right) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{l_i} f(C_{ij}) \\ &\le \sum_{i=1}^{k} f(B_i) \end{aligned}$$

이다.

한편, f가 $\sigma-$가법성을 가지는 경우에 ii°에 대해서는 집합열 $\{A_i\}$를 $A_i := A\cap(B_i \setminus \bigcup_{j=1}^{i-1} B_j) = (A\cap B_i)\setminus B_1 \setminus \cdots \setminus B_{i-1}$로 두고 방금 보인 ii의 증명과 비슷하게 하면 된다. □

Proposition 1.20 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하면 다음이 성립한다.

i. (단조성) 임의의 $A, B \in \mathscr{A}$에 대해 $A \subseteq B$이면 $\rho(A) \le \rho(B)$이다. 나아가 $B\setminus A \in \mathscr{A}$이면 $\rho(B) = \rho(A) + \rho(B\setminus A)$이다.
ii. ($\sigma-$반가법성) Semi-algebra $\mathscr{A}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$이면 $\rho(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \rho(A_i)$이다.

Proof i. 이는 위의 보조정리의 i로부터 자명하다. 한편, 만약 $B\setminus A \in \mathscr{A}$이면 당연히 $\rho(B) = \rho(A) + \rho(B\setminus A)$이다.

ii. 이는 위의 보조정리의 ii°로부터 자명하다. □

따라서 premeasure는 $\sigma-$가법성과 $\sigma-$반가법성, 가법성, 그리고 반가법성을 모두 가진다고 정리할 수 있다.

Theorem 1.21 (Inclusion-exclusion principle) 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하자. 그렇다면 임의의 $A_1, \cdots, A_l \in \mathscr{A}$에 대해 $\bigcup_{i=1}^{l} A_i \in \mathscr{A}$이고 $\rho(\bigcup_{i=1}^{l} A_i) < \infty$이면

$$\rho\left(\bigcup_{i=1}^{l} A_i\right) = \sum_{i=1}^{l} (-1)^{i-1} \sum_{1\le j_1<\cdots<j_i\le l} \rho\left(\bigcap_{k=1}^{i} A_{j_k}\right)$$

이다.

Proof 우선, 임의의 $j_1 < \cdots < j_i \le l$에 대해 $\rho(\bigcap_{k=1}^{i} A_{j_k}) \le \rho(\bigcup_{i=1}^{l} A_i) < \infty$임은 분명하다. 그렇다면 $l = 1$인 경우는 자명하고, $l = 2$인 경우에는

$$\rho(A_1 \cup A_2) = \rho(A_1 \setminus A_2) + \rho(A_2 \setminus A_1) + \rho(A_1 \cap A_2)$$

$$\rho(A_1) = \rho(A_1 \setminus A_2) + \rho(A_1 \cap A_2)$$
$$\rho(A_2) = \rho(A_2 \setminus A_1) + \rho(A_1 \cap A_2)$$

와 $\rho(A_1 \setminus A_2), \rho(A_2 \setminus A_1) \leq \rho(A_1 \cup A_2) < \infty$에서 $\rho(A_1 \cup A_2) = \rho(A_1) + \rho(A_2) - \rho(A_1 \cap A_2)$이다. 이제 $l \geq 3$인 경우에 대해서는 수학적 귀납법을 사용하면 된다. 귀납가정으로서 $l \in \mathbb{N}$에 대해 위의 명제가 성립한다고 하면 $\rho(\bigcup_{i=1}^{l} A_i), \rho(\bigcup_{i=1}^{l}(A_i \cap A_{l+1})) \leq \rho(\bigcup_{i=1}^{l+1} A_i) < \infty$이므로

$$\begin{aligned}
\rho\left(\bigcup_{i=1}^{l+1} A_i\right) &= \rho\left(\bigcup_{i=1}^{l} A_i\right) + \rho(A_{l+1}) - \rho\left(\bigcup_{i=1}^{l}(A_i \cap A_{l+1})\right) \\
&= \sum_{i=1}^{l}(-1)^{i-1} \sum_{1 \leq j_1 < \cdots < j_i \leq l} \rho\left(\bigcap_{k=1}^{i} A_{j_k}\right) + \rho(A_{l+1}) \\
&\qquad - \sum_{i=1}^{l}(-1)^{i-1} \sum_{1 \leq j_1 < \cdots < j_i \leq l} \rho\left(\bigcap_{k=1}^{i} A_{j_k} \cap A_{l+1}\right) \\
&= \sum_{i=1}^{l+1}(-1)^{i-1} \sum_{1 \leq j_1 < \cdots < j_i \leq l+1} \rho\left(\bigcap_{k=1}^{i} A_{j_k}\right)
\end{aligned}$$

에서 $l+1$에 대해서도 위의 명제가 성립함을 알 수 있다. □

다음은 집합열의 극한에 관련된 성질들이다.

Definition 1.22 집합열 $\{A_i\}$에 대해 이의 **하극한(limit infimum)**과 **상극한(limit supremum)**을 각각 $\liminf_{i\to\infty} A_i, \limsup_{i\to\infty} A_i$로 쓰고

$$\liminf_{i\to\infty} A_i = \bigcup_{j=1}^{\infty}\bigcap_{i=j}^{\infty} A_i, \qquad \limsup_{i\to\infty} A_i = \bigcap_{j=1}^{\infty}\bigcup_{i=j}^{\infty} A_i$$

로 정의한다. 나아가, 만약 $\liminf_{i\to\infty} A_i = \limsup_{i\to\infty} A_i$이면 이때 $\{A_i\}$가 **수렴(converge)**한다고 하며, 이때의 공통의 값을 $\lim_{i\to\infty} A_i$로 쓴다.

Proposition 1.23 집합열 $\{A_i\}$에 대해 다음이 성립한다.

i. $\liminf_{i\to\infty} A_i \subseteq \limsup_{i\to\infty} A_i$.
ii. 만약 $\{A_i\}$가 증가하면(혹은 감소하면) $A_i \uparrow \bigcup_{i=1}^{\infty} A_i$(혹은 $A_i \downarrow \bigcap_{i=1}^{\infty} A_i$)이다.
iii. 가측공간 $(X, \mathscr{A})$가 $\{A_i\}$를 포함하면 $\liminf_{i\to\infty} A_i, \limsup_{i\to\infty} A_i \in \mathscr{A}$이다.

PROOF i. 임의의 $x \in \liminf_{i\to\infty} A_i = \bigcup_{j=1}^{\infty}\bigcap_{i=j}^{\infty} A_i$에 대해 적당한 $i_0 \in \mathbb{N}$이 존재하여 $i \geq i_0$이면 $x \in A_i$이므로 임의의 $j \in \mathbb{N}$에 대해 $i = \max\{i_0, j\}$로 두면 $x \in A_i$가 되어 $x \in \bigcap_{j=1}^{\infty}\bigcup_{i=j}^{\infty} A_i = \limsup_{i\to\infty} A_i$이고, 곧 $\liminf_{i\to\infty} A_i \subseteq \limsup_{i\to\infty} A_i$임을 안다.

ii. 간결한 논의를 위해 $\{A_i\}$가 증가하는 경우만 생각하자. (반대의 경우도 비슷하게 하면 된다.) 집합열 $\{A_i\}$가 증가하므로 $\liminf_{i\to\infty} A_i = \bigcup_{i=1}^{\infty} A_i$이고 임의의 $j \in \mathbb{N}$에 대해 $\bigcup_{i=j}^{\infty} A_i = \bigcup_{i=1}^{\infty} A_i$이므로 $\limsup_{i\to\infty} A_i = \bigcap_{j=1}^{\infty}\bigcup_{i=j}^{\infty} A_i = \bigcup_{i=1}^{\infty} A_i$이 되어 정의로부터 $A_i \uparrow \bigcup_{i=1}^{\infty} A_i$이다.

iii. 이는 정의로부터 자명하다. □

Theorem 1.24 측도공간 $(X, \mathscr{A}, \mu)$에 속하는 임의의 집합열 $\{A_i\}$에 대해 다음이 성립한다.

i. 만약 $A_i \uparrow A$이면 $\mu(A_i) \uparrow \mu(A)$이다.
ii. 만약 $A_i \downarrow A$이고 $\mu(A_1) < \infty$이면 $\mu(A_i) \downarrow \mu(A)$이다.

Proof i. 만약 어떤 $i \in \mathbb{N}$에 대해 $\mu(A_i) = \infty$이면 $\mu(A_i) \le \mu(A) = \infty$가 되어 명제가 자명하므로 모든 $i \in \mathbb{N}$에 대해 $\mu(A_i) < \infty$라 하자. 이제 집합열 $\{B_i\}$를 $B_i := A_i \setminus A_{i-1}$ $(B_1 := A_1)$로 정의하면 이는 $\mathscr{A}$에 속하는 서로소인 집합열이며 $A = \bigcup_{i=1}^{\infty} A_i = \bigsqcup_{i=1}^{\infty} B_i$이므로

$$
\begin{aligned}
\mu(A) &= \mu\left(\bigsqcup_{i=1}^{\infty} B_i\right) \\
&= \sum_{i=1}^{\infty} \mu(B_i) \\
&= \mu(A_1) + \sum_{i=2}^{\infty} \mu(A_i \setminus A_{i-1}) \\
&= \lim_{j\to\infty} \left\{\mu(A_1) + \sum_{i=1}^{j} [\mu(A_i) - \mu(A_{i-1})]\right\} \\
&= \lim_{j\to\infty} \mu(A_j).
\end{aligned}
$$

이다. 이제 수열 $\{\mu(A_i)\}$가 증가함이 분명하므로 명제가 성립한다.

ii. 집합열 $\{B_i\}$를 $B_i := A_i \setminus A_{i+1}$로 정의하면 이는 $\mathscr{A}$에 속하는 서로소인 집합열이며 $A_1 = \bigsqcup_{i=1}^{\infty} B_i \sqcup A$이다. 한편, 가정으로부터 모든 $i \in \mathbb{N}$에 대해 $\mu(A_i) \le \mu(A_1) < \infty$이고 $\mu(A) \le \mu(A_1) < \infty$이므로

$$
\begin{aligned}
\mu(A_1) &= \mu\left(\bigsqcup_{i=1}^{\infty} B_i \sqcup A\right) \\
&= \sum_{i=1}^{\infty} \mu(B_i) + \mu(A) \\
&= \sum_{i=1}^{\infty} \mu(A_i \setminus A_{i+1}) + \mu(A) \\
&= \lim_{j\to\infty} \sum_{i=1}^{j} [\mu(A_i) - \mu(A_{i+1})] + \mu(A)
\end{aligned}
$$

$$= \mu(A_1) - \lim_{j\to\infty} \mu(A_j) + \mu(A)$$

에서 $\mu(A_i) \to \mu(A)$임을 안다. 이제 수열 $\{\mu(A_i)\}$가 감소함이 분명하므로 명제가 성립한다. □

Definition 1.25 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하자. 만약 $\rho(X) < \infty$이면 ρ가 **유한(finite)**하다고 한다. 만약 $\mathscr{A}$에 속하는 적당한 집합열 $\{A_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $\rho(A_i) < \infty$이고 $\bigcup_{i=1}^{\infty} A_i = X$이면 ρ가 **σ−유한(σ−finite)**하다고 한다. 특별히, $\mathscr{A}$의 생성자 $\mathscr{C}$에 대해 방금의 조건을 만족하는 집합열이 $\mathscr{C}$에 속하면 ρ가 $\mathscr{C}$ **위에서 σ−유한(σ−finite on $\mathscr{C}$)**하다고 한다.

Theorem 1.26 유한 측도공간 $(X, \mathscr{A}, \mu)$와 $\mathscr{A}$에 속하는 임의의 집합열 $\{A_i\}$에 대해

$$\mu(\liminf_{i\to\infty} A_i) \le \liminf_{i\to\infty} \mu(A_i) \le \limsup_{i\to\infty} \mu(A_i) \le \mu(\limsup_{i\to\infty} A_i)$$

가 성립한다. 특별히, $A_i \to A$이면 $\mu(A_i) \to \mu(A)$이다.

Proof 위의 식에서 $\liminf_{i\to\infty} \mu(A_i) \le \limsup_{i\to\infty} \mu(A_i)$ 부분은 자명하므로 첫 번째와 세 번째 부등식만 보이면 되는데, 간결한 논의를 위해 여기서는 첫 번째 부등식만 보이자. (세 번째 부등식도 비슷하게 하면 된다.) 먼저 $B = \liminf_{i\to\infty} A_i$라 하고 집합열 $\{B_i\}$를 $B_i := \bigcap_{j=i}^{\infty} A_j$로 두면 이는 $\mathscr{A}$에 속하는 증가하는 집합열로서 $B_i \uparrow B$이다. 따라서 정리 ? ? 의 i로부터 $\mu(B) = \lim_{i\to\infty} \mu(B_i) \le \sup_{i\in\mathbb{N}} \mu(\bigcap_{j=i}^{\infty} A_j) \le \sup_{i\in\mathbb{N}} \inf_{j\ge i} \mu(A_j) = \liminf_{i\to\infty} \mu(A_i)$를 얻는다. 한편, $A_i \to A$인 경우에는 방금 보인 부등식으로부터 $\mu(A_i) \to \mu(A)$임이 자명하다. □

나중에 사용할 정리를 하나 증명하는 것으로 이번 절을 끝마친다.

Definition 1.27 측도공간 $(X, \mathscr{A}, \mu)$와 $A \in \mathscr{A}$에 대해 $\mu_A : B \mapsto \mu(A \cap B)$로 정의된 함수 $\mu_A : \mathscr{A} \to \overline{\mathbb{R}}_0^+$를 μ의 A로의 **제한(restriction)**이라 한다.

Proposition 1.28 측도공간 $(X, \mathscr{A}, \mu)$와 집합 $A \in \mathscr{A}$에 대해 μ_A는 $\mathscr{A}$ 위의 측도이다.

Proof 거의 자명하다. 우선 $\mu_A(\emptyset) = \mu(\emptyset) = 0$임은 분명하고 $\mathscr{A}$에 속하는 임의의 서로소인 집합열 $\{B_i\}$에 대해 $\mu_A(\bigsqcup_{i=1}^{\infty} B_i) = \mu(\bigsqcup_{i=1}^{\infty} (A \cap B_i)) = \sum_{i=1}^{\infty} \mu(A \cap B_i) = \sum_{i=1}^{\infty} \mu_A(B_i)$가 성립하므로 μ_A는 $\mathscr{A}$ 위의 측도가 된다. □

Theorem 1.29 가측공간 $(X, \mathscr{A})$와 $\mathscr{A}$의 생성자이자 π−system인 $\mathscr{P}$에 대해 μ, ν를 $\mathscr{A}$ 위의 측도로서 $\mathscr{P}$ 위에서 σ−유한하다고 하자. 만약 μ, ν가 $\mathscr{P}$ 위에서 일치하면 $\mu = \nu$이다.

PROOF 가정으로부터 $\mathscr{P}$에 속하는 집합열 $\{A_i\}$가 존재하여 $\mu(A_i) < \infty$이고 $\bigcup_{i=1}^{\infty} A_i = X$이다. 여기서 집합열 $\{B_i\}$를 $B_i := A_i \setminus \bigcup_{j=1}^{i-1} A_j$로 두면 이는 $\mathscr{A}$에 속하는 서로소인 집합열로 $\bigsqcup_{i=1}^{\infty} B_i = X$이며 포함배제의 원리로부터 모든 $i \in \mathbb{N}$에 대해

$$\begin{aligned}
\mu(B_i) &= \mu\left(A_i \setminus \bigcup_{j=1}^{i-1} A_j\right) \\
&= \mu(A_i) - \mu\left(\bigcup_{j=1}^{i-1} (A_i \cap A_j)\right) \\
&= \mu(A_i) - \sum_{j=1}^{i-1} (-1)^{j-1} \sum_{1 \le k_1 < \cdots < k_j \le i-1} \mu\left(\bigcap_{l=1}^{j} (A_i \cap A_{k_l})\right) \\
&= \nu(A_i) - \sum_{j=1}^{i-1} (-1)^{j-1} \sum_{1 \le k_1 < \cdots < k_j \le i-1} \nu\left(\bigcap_{l=1}^{j} (A_i \cap A_{k_l})\right) \\
&= \nu(A_i) - \nu\left(\bigcup_{j=1}^{i-1} (A_i \cap A_j)\right) \\
&= \nu\left(A_i \setminus \bigcup_{j=1}^{i-1} A_j\right) \\
&= \nu(B_i)
\end{aligned}$$

가 성립한다. 이를 이용하여 우선 μ, ν가 유한한 경우에 대해 위의 정리를 증명하자. 일단 $\mu(X) = \mu(\bigsqcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i) = \sum_{i=1}^{\infty} \nu(B_i) = \nu(\bigsqcup_{i=1}^{\infty} B_i) = \nu(X)$가 성립하므로 집합족 $\mathscr{L} = \{A \in \mathscr{A} : \mu(A) = \nu(A)\}$를 생각하면 이가 $\mathscr{P}$를 포함하는 $\lambda-$system임이 거의 자명하다. 그렇다면 Dynkin의 $\pi-\lambda$ 정리로부터 $\mathscr{A} = \sigma(\mathscr{P}) = \lambda(\mathscr{P}) \subseteq \mathscr{L}$이 되어 증명이 끝난다.

이제 일반적인 경우를 증명하자. 이를 위해 각 $i \in \mathbb{N}$에 대해 측도 μ_{B_i}, ν_{B_i}를 생각하고 앞서 포함배제의 원리를 사용한 것과 비슷하게 하면 μ_{B_i}, ν_{B_i}가 $\mathscr{A}$ 위의 유한 측도이며 $\mathscr{P}$ 위에서 일치함을 보일 수 있다. 따라서 앞선 결론으로부터 각 $i \in \mathbb{N}$에 대해 $\mu_{B_i} = \nu_{B_i}$임을 안다. 그렇다면 임의의 $A \in \mathscr{A}$에 대해 $\mu(A) = \mu(\bigsqcup_{i=1}^{\infty}(A \cap B_i)) = \sum_{i=1}^{\infty} \mu(A \cap B_i) = \sum_{i=1}^{\infty} \mu_{B_i}(A) = \sum_{i=1}^{\infty} \nu_{B_i}(A) = \nu(\bigsqcup_{i=1}^{\infty}(A \cap B_i)) = \nu(A)$이므로 $\mu = \nu$임을 안다. □

위의 정리는 $\sigma-$유한인 측도는 그 측도가 정의된 $\sigma-$대수의 $\pi-$system인 생성자 위에서의 값만으로 특정됨을 함의한다. 일반적으로 $\sigma-$대수의 생성자는 그가 생성하는 $\sigma-$대수보다 그 크기가 작고, 대부분의 경우 그 원소를 명시적으로 표현할 수 있기에 위의 정리는 상당히 유용하고 강력한 정리이다. 한편, 이가 사실상 Dynkin의 $\pi-\lambda$ 정리의 따름정리 격이라는 점도 눈여겨 볼 만하다.

1.3 Extension Theorems

이번 절에서는 Lebesgue 측도의 구성의 핵심이 되는 확장정리들에 대해 자세히 알아본다. 앞서 잠시 언급했듯이 확장정리들은 semi-algebra와 같은 약한 구조에서 정의된 premeasure를 대수나 $\sigma-$대수와 같은 보다 강한 구조를 가진 곳에서의 premeasure나 측도로 확장시킬 수 있는 방법을 제공한다. 여기에서는 크게 두 가지의 확장정리를 소개할텐데, 첫 번째 확장정리에 비해 두 번째 확장정리는 조금 깊은 이해를 필요로 한다. 마음의 준비를 하도록 하자.

Theorem 1.30 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하면 이의 $A(\mathscr{A})$에서의 premeasure로의 확장이 유일하게 존재한다.

Proof 정리 ? ? 로부터 임의의 $A \in A(\mathscr{A})$에 대해 적당한 서로소인 집합 $A_1, \cdots, A_k \in \mathscr{A}$가 존재하여 $A = \bigsqcup_{i=1}^k A_i$이다. 이제 함수 $\widetilde{\rho} : A(\mathscr{A}) \to \overline{\mathbb{R}}_0^+$를 위와 같이 표현되는 A에 대해 $\widetilde{\rho} : A \mapsto \sum_{i=1}^k \rho(A_i)$인 함수로 정의하자. 그렇다면 이는 명백히 ρ의 $A(\mathscr{A})$로의 확장이다. 한편, 이의 well-definedness를 담보하기 위해 위의 A를 또다른 서로소인 $B_1, \cdots, B_l \in \mathscr{A}$에 대해 $A = \bigsqcup_{j=1}^l B_j$로도 쓸 수 있다고 하자. 그렇다면 $A_1 \cap B_1, \cdots, A_k \cap B_l \in \mathscr{A}$이 모두 서로소이므로 $\sum_{i=1}^k \rho(A_i) = \sum_{i=1}^k \rho(\bigsqcup_{j=1}^l (A_i \cap B_j)) = \sum_{i=1}^k \sum_{j=1}^l \rho(A_i \cap B_j) = \sum_{j=1}^l \rho(\bigsqcup_{i=1}^k (A_i \cap B_j)) = \sum_{j=1}^l \rho(B_j)$가 되어 $\widetilde{\rho}$가 well-define됨을 안다.

이제 $\widetilde{\rho}$가 premeasure임을 보이자. 정의로부터 $\widetilde{\rho}(\emptyset) = \rho(\emptyset) = 0$임은 분명하므로 이의 $\sigma-$가법성을 보이는 것으로 충분하다. 이를 위해, $A(\mathscr{A})$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 $A = \bigsqcup_{i=1}^\infty A_i \in A(\mathscr{A})$라 하자. 그렇다면 다시 정리 ? ? 로부터 적당한 서로소인 집합 $B_1, \cdots, B_l \in \mathscr{A}$가 존재하여 $A = \bigsqcup_{j=1}^l B_j$이고 각 $i \in \mathbb{N}$에 대해서도 적당한 서로소인 집합 $C_{i1}, \cdots, C_{im_i} \in \mathscr{A}$가 존재하여 $A_i = \bigsqcup_{k=1}^{m_i} C_{ik}$이다. 이로부터

$$
\begin{aligned}
\widetilde{\rho}(A) &= \widetilde{\rho}\left(\bigsqcup_{j=1}^{l} B_j\right) \\
&= \sum_{j=1}^{l} \rho(B_j) \\
&= \sum_{j=1}^{l} \rho\left(\bigsqcup_{i=1}^{\infty} \bigsqcup_{k=1}^{m_i} (B_j \cap C_{ik})\right) \\
&= \sum_{j=1}^{l} \sum_{i=1}^{\infty} \sum_{k=1}^{m_i} \rho(B_j \cap C_{ik}) \\
&= \sum_{i=1}^{\infty} \widetilde{\rho}\left(\bigsqcup_{k=1}^{m_i} \bigsqcup_{j=1}^{l} (B_j \cap C_{ik})\right)
\end{aligned}
$$

$$= \sum_{i=1}^{\infty} \widetilde{\rho}\left(\bigsqcup_{k=1}^{m_i} C_{ik}\right)$$
$$= \sum_{i=1}^{\infty} \widetilde{\rho}(A_i)$$

가 되어 $\widetilde{\rho}$가 σ−가법성을 가짐을 안다.

이제 이러한 확장의 유일성만 보이면 되는데, ρ의 $A(\mathscr{A})$에서의 premeasure로의 어떠한 확장이든 그 확장은 σ−가법성을 가지고, 곧 가법성을 가지므로 앞서 정의한 $\widetilde{\rho}$와 일치할 수 밖에 없다. 따라서 유일성은 자명하다. □

위의 확장정리는 semi-algebra 위에서의 premeasure를 그 semi-algebra가 생성하는 대수 위의 premeasure로 확장시키는 방법을 제공한다. 그리고 이를 σ−대수 위의 측도로 끌어올리는 확장정리가 바로 두 번째로 소개할 확장정리이다. 다만, 두 번째 확장정리는 첫 번째보다 조금 더 복잡해서, 대수 위의 premeasure를 σ−대수 위의 측도로 곧바로 확장시키는 것이 아니라, 우선 이를 전체 공간으로 확장시킨 다음 다시 적당한 σ−대수로 축소하여 측도로의 확장을 얻는다. 당연히 이때 전체 공간으로 확장시킨 함수는 측도가 요구하는 강한 조건을 만족시키지 못하므로 측도가 아니며, 이를 외측도라 부른다.

Definition 1.31 공집합이 아닌 집합 X에 대해 함수 $\mu^* : \mathscr{P}(X) \to \overline{\mathbb{R}}_0^+$가

i. $\mu^*(\emptyset) = 0$.
ii. (단조성) 임의의 $A, B \subseteq X$에 대해 $A \subseteq B$이면 $\mu^*(A) \leq \mu^*(B)$이다.
iii. (σ−반가법성) 멱집합 $\mathscr{P}(X)$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $\mu^*(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$이다.

를 만족하면 이때의 μ^*를 $\mathscr{P}(X)$ 위의 **외측도(outer measure)**라 한다.

보통 외측도는 그리스 문자에 asterisk $*$을 위첨자로 붙여 $\mu^*, \nu^*, \cdots$로 표기한다. Premeasure가 σ−대수보다 약한 구조에서 정의된 일반적인 측도라면, 외측도는 측도의 강력한 조건인 σ−가법성을 과감히 포기하고 이를 단조성과 σ−반가법성으로 대체하여 얻는 전체 공간에서의 일반적인 측도로 볼 수 있다. 위의 정의의 세 조건을 만족하는 함수라면 모두 외측도라 할 수 있지만, 우리는 외측도를 확장정리의 중간 과정으로만 이용할 계획이므로 지금부터는 premeasure로부터 유도되는 외측도에 집중하도록 하자.

Definition 1.32 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하면 **ρ로부터 유도되는 외측도(outer measure induced by ρ)**를 ρ^*로 쓰고

$$\rho^* : A \mapsto \inf\left\{ \sum_{i=1}^{\infty} \rho(A_i) \in \overline{\mathbb{R}}_0^+ : \{A_i\}\text{는 } \mathscr{A}\text{에 속하는 } A\text{의 가산 덮개이다.} \right\}$$

로 정의한다.

Proposition 1.33 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하자. 그렇다면 이로부터 유도되는 외측도 ρ^*에 대해 다음이 성립한다.

i. 만약 ρ가 유한하면 임의의 $A \subseteq X$에 대해 $\rho^*(A) < \infty$이다.
ii. 함수 ρ^*와 ρ는 $\mathscr{A}$ 위에서 일치한다. 따라서 ρ^*는 ρ의 확장이다.
iii. 함수 ρ^*는 외측도이다.

Proof i. 이는 $\{X\}$가 임의의 $A \subseteq X$의 $\mathscr{A}$에 속하는 가산 덮개라는 점에서 자명하다.

ii. 임의의 $A \in \mathscr{A}$를 택하자. 우선 $\{A\}$가 그 자체로 $\mathscr{A}$에 속하는 A의 가산 덮개이므로 $\rho^*(A) \le \rho(A)$이다. 역으로, $\mathscr{A}$에 속하는 A의 임의의 가산 덮개 $\{A_i\}$에 대해 WLOG, 필요하다면 이에 공집합을 추가하여 이가 집합열을 이룬다고 하면 $\sigma-$반가법성으로부터 $\rho(A) \le \sum_{i=1}^{\infty} \rho(A_i)$이므로 $\rho(A) \le \rho^*(A)$이다.

iii. ii로부터 $\rho^*(\emptyset) = \rho(\emptyset) = 0$임은 분명하다. 또한 임의의 두 $A, B \subseteq X$에 대해 $A \subseteq B$라면 B의 모든 가산 덮개는 곧 A의 가산 덮개이기도 하므로 $\rho^*(A) \le \rho^*(B)$에서 ρ^*는 단조성도 가진다. 곧 ρ^*의 $\sigma-$반가법성을 보이면 충분하므로 이를 위해 $\mathscr{P}(X)$에 속하는 임의의 집합열 $\{A_i\}$를 택하자. 만약 $\rho^*(A_{i_0}) = \infty$인 $i_0 \in \mathbb{N}$가 존재하면 $\rho^*(\bigcup_{i=1}^{\infty} A_i) \ge \rho^*(A_{i_0}) = \infty$에서 $\sigma-$반가법성이 자명하므로 모든 $i \in \mathbb{N}$에 대해 $\rho^*(A_i) < \infty$라 가정하자. 이제 임의의 $\varepsilon > 0$을 택하면 각 $i \in \mathbb{N}$에 대해 $\mathscr{A}$에 속하는 A_i의 적당한 가산 덮개 $\{B_{ij}\}_j$가 존재하여 WLOG, 필요하다면 이에 공집합을 추가하여 이가 집합열을 이룬다고 하면 $\sum_{j=1}^{\infty} \rho(B_{ij}) < \rho^*(A_i) + \varepsilon/2^i$이다. 한편, 이러한 덮개를 모두 합한 $\{B_{ij}\}_{ij}$가 $\mathscr{A}$에 속하는 $\bigcup_{i=1}^{\infty} A_i$의 가산 덮개이므로 $\rho^*(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \rho(B_{ij}) \le \sum_{i=1}^{\infty} [\rho^*(A_i) + \varepsilon/2^i] = \sum_{i=1}^{\infty} \rho^*(A_i) + \varepsilon$에서 $\rho^*(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \rho^*(A_i)$가 되어 ρ^*가 $\sigma-$반가법성을 가짐을 안다. □

이제부터 다소 혼란스러운 부분이 시작된다. 앞서 1절에서 이미 가측집합을 $\sigma-$대수의 원소로 정의한 바 있는데, 우리는 여기서 가측집합을 다시 정의한다.

Definition 1.34 공집합이 아닌 집합 X의 멱집합 $\mathscr{P}(X)$ 위의 외측도 μ^*에 대해 만약 $A \subseteq X$가 임의의 $S \subseteq X$에 대해 $\mu^*(S \cap A) + \mu^*(S \cap A^c) = \mu^*(S)$를 만족하면 이때의 A를 **(μ^*-)가측집합((μ^*-)measurable set)**이라 한다.

외측도의 $\sigma-$반가법성을 이용하면 가측일 조건을 다음과 같이 쓸 수도 있다.

Proposition 1.35 공집합이 아닌 집합 X의 멱집합 $\mathscr{P}(X)$ 위의 외측도 μ^*에 대해 집합 $A \subseteq X$가 가측일 필요충분조건은 임의의 $S \subseteq X$에 대해 $\mu^*(S \cap A) + \mu^*(S \cap A^c) \le \mu^*(S)$가 성립하는 것이다.

Proof 외측도 μ^*의 $\sigma-$반가법성으로부터 $\mu^*(S) = \mu^*((S \cap A) \sqcup (S \setminus A)) \le \mu^*(S \cap A) + \mu^*(S \cap A^c)$가 항상 성립하므로 이는 자명하다. □

더 이상 진행하기 전에, 가측집합에 대한 정의로 인한 혼란을 조금 해결하고 가는 것이 좋을 것 같다. 혼란의 원인은 앞서 제시된 두 개의 정의가 서로 전혀 동등하게 보이지 않고, 실제로도 동등하지 않다는 점이다. 하지만, 다행히 다음 정리에 의하면 두 번째 가측집합의 정의는 첫 번째 가측집합의 정의에 부합하여 위의 두 정의는, 비록 서로 같지는 않지만, 양립 가능하다. (사실 첫 번째 정의가 두 번째 정의보다 더 일반적인 정의인 격이다.) 이러한 혼란에도 불구하고, 전체 공간으로 확장된 외측도를 다시 적당한 σ−대수로 축소하여 측도를 얻기 위해서 두 번째 가측집합의 정의는 반드시 필요하다.

Lemma 1.36 공집합이 아닌 집합 X의 멱집합 $\mathscr{P}(X)$ 위의 외측도 μ^*에 대해 $\mathscr{M}$을 모든 가측집합들의 모임이라 하면 μ^*는 $\mathscr{M}$ 위에서 가법성을 가진다.

Proof 임의의 서로소인 $A,B\in\mathscr{M}$에 대해 $\mu^*(A\sqcup B)=\mu^*((A\sqcup B)\cap A)+\mu^*((A\sqcup B)\cap A^c)=\mu^*(A)+\mu^*(B)$가 성립하고 여기에 수학적 귀납법을 적용하면 이 보조정리는 쉽게 보일 수 있다. □

Theorem 1.37 공집합이 아닌 집합 X의 멱집합 $\mathscr{P}(X)$ 위의 외측도 μ^*에 대해 $\mathscr{M}$을 모든 가측집합들의 모임이라 하면 이는 σ−대수이다.

Proof 먼저 $\mathscr{M}$이 대수임을 보이자. 우선 $\emptyset$이 가측이므로 $\mathscr{M}$은 비어있지 않다. 또한 임의의 $A\in\mathscr{M}$에 대해 $A^c\in\mathscr{M}$이므로 $\mathscr{M}$은 여집합에 대해 닫혀있다. 다음으로, 임의의 $A,B\in\mathscr{M}$를 택하면 임의의 $S\subseteq X$에 대해 $\mu^*(S)=\mu^*(S\cap A)+\mu^*(S\cap A^c)=\mu^*(S\cap A\cap B)+\mu^*(S\cap A\cap B^c)+\mu^*(S\cap A^c)$이고 $(S\cap A\cap B^c)\cup(S\cap A^c)=S\cap(A\cap B)^c$이므로 $\mu^*(S)\geq\mu^*(S\cap(A\cap B))+\mu^*(S\cap(A\cap B)^c)$가 되어 명제 ? ? 로부터 $A\cap B\in\mathscr{M}$이고, 곧 $\mathscr{M}$이 교집합에 대해서도 닫혀있음을 안다. 그렇다면 임의의 $A,B\in\mathscr{M}$에 대해 $A\cup B=(A^c\cap B^c)^c\in\mathscr{M}$이므로 $\mathscr{M}$은 합집합에 대해 닫혀있고, 이로써 $\mathscr{M}$은 대수임을 안다.

이제 $\mathscr{M}$이 가산 합집합에 대해 닫혀있다는 것만 보이면 된다. 이를 위해 $\mathscr{M}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 WLOG, 필요하다면 각 항을 $B_i:=A_i\setminus\bigcup_{j=1}^{i-1}A_j$로 바꾸어 $\{A_i\}$가 처음부터 서로소인 집합열이라고 해도 된다. 그렇다면 방금 $\mathscr{M}$이 대수임을 보였으므로 각 $k\in\mathbb{N}$에 대해 $\bigsqcup_{i=1}^{k}A_i\in\mathscr{M}$이고 위의 보조정리로부터 각 $k\in\mathbb{N}$와 임의의 $S\subseteq X$에 대해

$$\begin{aligned}
\mu^*(S)&=\mu^*\left(S\cap\bigsqcup_{i=1}^{k}A_i\right)+\mu^*\left(S\cap\left(\bigsqcup_{i=1}^{k}A_i\right)^c\right)\\
&=\mu^*\left(\bigsqcup_{i=1}^{k}(S\cap A_i)\right)+\mu^*\left(S\cap\left(\bigsqcup_{i=1}^{k}A_i\right)^c\right)\\
&\geq\sum_{i=1}^{k}\mu^*(S\cap A_i)+\mu^*\left(S\cap\left(\bigsqcup_{i=1}^{\infty}A_i\right)^c\right)
\end{aligned}$$

가 성립한다. 이로부터 임의의 $S\subseteq X$에 대해

$$\begin{aligned}\mu^*(S) &\geq \sum_{i=1}^{\infty}\mu^*(S\cap A_i)+\mu^*\left(S\cap\left(\bigsqcup_{i=1}^{\infty}A_i\right)^c\right)\\ &\geq \mu^*\left(\bigsqcup_{i=1}^{\infty}(S\cap A_i)\right)+\mu^*\left(S\cap\left(\bigsqcup_{i=1}^{\infty}A_i\right)^c\right)\\ &= \mu^*\left(S\cap\bigsqcup_{i=1}^{\infty}A_i\right)+\mu^*\left(S\cap\left(\bigsqcup_{i=1}^{\infty}A_i\right)^c\right)\end{aligned}$$

가 성립하여 명제 ? ? 으로부터 $\bigsqcup_{i=1}^{\infty}A_i\in\mathscr{M}$이고, 곧 증명이 끝난다. □

자, 정의 ? ? 에서의 가측집합을 생각해보자. 어떤 집합 A가 μ^*−가측이라 함은 곧 모든 μ^*−가측집합들의 모임 $\mathscr{M}$에 대해 $A\in\mathscr{M}$임을 뜻하고, 따라서 정의 ? ? 에 따르더라도 A는 $\mathscr{M}$−가측집합이 되어 앞서 말한 바와 같이 비록 두 정의가 서로 동등하지는 않지만, 이 두 정의는 서로 양립가능하다. 이제 두 번째 확장정리인 Carathéodory의 확장정리를 보자.

Lemma 1.38 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 함수 $f:\mathscr{A}\to\overline{\mathbb{R}}_0^+$를 $f(\emptyset)=0$인 가법성을 갖는 함수라 하면 f가 σ−가법성을 가질 필요충분조건은 f가 σ−반가법성을 갖는 것이다.

Proof 만약 f가 σ−가법성을 가지면 이는 premeasure이므로 명제 ? ? 의 ii로부터 f가 σ−반가법성을 가지게 되어 충분조건임은 자명하다. 필요조건임을 보이기 위해 $\mathscr{A}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$를 생각하여 $\bigsqcup_{i=1}^{\infty}A_i\in\mathscr{A}$라 하면 가정으로부터 $f(\bigsqcup_{i=1}^{\infty}A_i)\leq\sum_{i=1}^{\infty}f(A_i)$가 성립하고 보조정리 ? ? 의 i로부터 임의의 $k\in\mathbb{N}$에 대해 $\sum_{i=1}^{k}f(A_i)=f(\bigsqcup_{i=1}^{k}A_i)\leq f(\bigsqcup_{i=1}^{\infty}A_i)$가 되어 $f(\bigsqcup_{i=1}^{\infty}A_i)\geq\sum_{i=1}^{\infty}f(A_i)$도 성립한다. 이로부터 $f(\bigsqcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}f(A_i)$가 되어 f가 σ−가법성을 가짐을 안다. □

Theorem 1.39 (Carathéodory's extension theorem) 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하고 $\mathscr{M}$을 모든 ρ^*−가측집합들의 모임이라 하자. 그렇다면 $\mu:=\rho^*|_{\mathscr{M}}$는 ρ의 $\mathscr{M}$에서의 측도로의 확장이고, 따라서 $(X,\mathscr{M},\mu)$는 측도공간이 된다.

Proof 우선 $\mu(\emptyset)=0$이고 μ가 정의로부터 σ−반가법성을 가지며 보조정리 ? ? 로부터 가법성도 가지므로 위의 보조정리에 의해 이는 σ−가법성을 가지고, 곧 μ는 $\mathscr{M}$ 위의 측도가 된다. 이제 μ가 ρ의 확장임을 보이기만 하면 되는데, 만약 $\mathscr{M}$이 $\mathscr{A}$를 포함한다는 것을 보인다면 명제 ? ? 의 ii로부터 μ가 ρ의 확장임이 분명하다. 이를 위해 임의의 $A\in\mathscr{A}$와 임의의 $S\subseteq X$ 그리고 임의의 $\varepsilon>0$을 택하면 $\mathscr{A}$에 속하는 적당한 가산 덮개 $\{S_i\}$가 존재하여 WLOG, 필요하다면 공집합을 추가하여 이가 집합열을 이룬다고 하면 $\sum_{i=1}^{\infty}\rho(S_i)\leq\rho^*(S)+\varepsilon$이다. 여기서 $\{S_i\cap A\},\{S_i\cap A^c\}$가 각각 $\mathscr{A}$에 속하는 $S\cap A, S\cap A^c$의 가산 덮개이므로 $\rho^*(S\cap A)+\rho^*(S\cap A^c)\leq\sum_{i=1}^{\infty}[\rho(S_i\cap A)+\rho(S_i\cap A^c)]=\sum_{i=1}^{\infty}\rho(S_i)\leq\rho^*(S)+\varepsilon$이 되어 곧 $\rho^*(S\cap A)+\rho^*(S\cap A^c)\leq\rho^*(S)$이다. 따라서 명제 ? ? 로부터 $A\in\mathscr{M}$이고, 곧 $\mathscr{A}\subseteq\mathscr{M}$이므로 증명이 끝난다. □

이와 같이 어떤 premeasure ρ에 Carathéodory의 확장정리를 적용함으로써 얻는 측도를 **ρ의 Carathéodory 확장(− extension of ρ)**이라 한다. 다만, 일반적으로 어떤 premeasure ρ를 모든 ρ^*−가측집합들의 모임 $\mathscr{M}$으로 확장하는 방법은 유일하지 않다. 즉, ρ의 $\mathscr{M}$에서의 측도로의 확장은 ρ의 Carathéodory 확장 말고도 얼마든지 있을 수 있다.[3]하지만 만약 ρ가 σ−유한하다면 그러한 확장은 Carathéodory 확장으로 유일함을 보일 수 있다.

Theorem 1.40 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 σ−유한 premeasure라 하고 $\mathscr{M}$을 모든 ρ^*−가측집합들의 모임이라 하자. 또한 μ를 ρ의 Carathéodory 확장이라 하자. 그렇다면 μ도 σ−유한하고, 임의의 측도공간 $(X, \mathscr{N}, \nu)$에 대해 만약 ν가 ρ의 확장이라면 μ와 ν는 $\mathscr{M} \cap \mathscr{N}$에서 일치한다.

Proof 우선 μ가 σ−유한함은 거의 자명하다. 가정으로부터 ρ가 σ−유한하므로 $\mathscr{A}$에 속하는 적당한 집합열 $\{A_i\}$가 존재하여 모든 $i \in \mathbb{N}$에 대해 $\mu(A_i) = \rho(A_i) < \infty$이고 $\bigcup_{i=1}^{\infty} A_i = X$이며, 따라서 μ도 σ−유한하다.

다음으로, $\mathscr{M} \cap \mathscr{N}$에서 μ와 ν가 일치함을 보이자. 이를 위해 임의의 $A \in \mathscr{M} \cap \mathscr{N}$를 택하면 $\mathscr{A}$에 속하는 A의 임의의 가산 덮개 $\{A_i\}$에 대해 WLOG, 필요하다면 공집합을 추가하여 이가 집합열을 이룬다고 하면 $\nu(A) \le \nu(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \nu(A_i) = \sum_{i=1}^{\infty} \rho(A_i)$이므로 $\nu(A) \le \rho^*(A) = \mu(A)$이다.

이제 임의의 $A \in \mathscr{M} \cap \mathscr{N}$를 고정하고 $\mu(A) < \infty$인 경우를 생각해보자. 이 경우에는 $\mathscr{A}$에 속하는 적당한 A의 가산 덮개 $\{A_i\}$가 존재하여 $\sum_{i=1}^{\infty} \rho(A_i) < \infty$이고, WLOG, 필요하다면 각 항을 $B_i := A_i \setminus \bigcup_{j=1}^{i-1} A_j$로 교체하여 $\{A_i\}$가 처음부터 서로소인 집합열이라 해도 된다. 여기서 $B = \bigsqcup_{i=1}^{\infty} A_i$라 하면 $B \setminus A \in \mathscr{M} \cap \mathscr{N}$이고, 곧 앞서 결론으로부터 $\nu(A) \le \mu(A)$, $\nu(B \setminus A) \le \mu(B \setminus A)$이다. 또한, $\mu(B) = \sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \rho(A_i) = \sum_{i=1}^{\infty} \nu(A_i) = \nu(B) < \infty$에서 $\nu(A) + \nu(B \setminus A) = \nu(B) = \mu(B) = \mu(A) + \mu(B \setminus A)$이다. 지금까지 얻은 결론들을 정리하면

i. $\nu(A) \le \mu(A)$.
ii. $\nu(B \setminus A) \le \mu(B \setminus A)$.
iii. $\nu(B) = \mu(B) < \infty$.
iv. $\nu(A) + \nu(B \setminus A) = \mu(A) + \mu(B \setminus A)$.

인데, 이를 종합하면 $\mu(A) = \nu(A)$임을 쉽게 알 수 있다.

마지막으로, 이번에는 $\mu(A) = \infty$인 경우를 생각해보자. 앞서 보인 바와 같이 μ가 σ−유한하므로 $\mathscr{A}$에 속하는 적당한 집합열 $\{B_i\}$가 존재하여 모든 $i \in \mathbb{N}$에 대해 $\mu(B_i) < \infty$이고 $\bigcup_{i=1}^{\infty} B_i = X$이며, WLOG, $\mu(A) < \infty$인 경우에서와 비슷하게 하여 $\{B_i\}$가 처음부터 서로소인 집합열이라 해도 된다. 그렇다면 모든 $i \in \mathbb{N}$에 대해 $\mu(A \cap B_i) \le \mu(B_i) < \infty$이므로 앞선 결론으로부터 $\mu(A) = \mu(\bigsqcup_{i=1}^{\infty}(A \cap B_i)) = \sum_{i=1}^{\infty} \mu(A \cap B_i) = \sum_{i=1}^{\infty} \nu(A \cap B_i) = \nu(\bigsqcup_{i=1}^{\infty}(A \cap B_i)) = \nu(A)$가 되어 증명이 끝난다. $\square$

나중에 사용할 정리 하나를 증명하는 것으로 이번 절을 마무리하자. 다음 정리는 측도의 근사에 대한 정리로, semi-algebra $\mathscr{A}$가 생성하는 $\sigma-$대수에서 정의된 $\sigma-$유한한 측도 μ는 $\mathscr{A}$에서의 정보로 충분히 근사될 수 있음을 함의한다. 이는 μ도 결국에는 $\mathscr{A}$에서의 적당한 premeasure에 조금씩 정보를 더해나가며 확장시켜 얻은 것이라는 점을 생각하면 자연스러운 결과이다.

Theorem 1.41 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$에 대해 μ를 $\sigma(\mathscr{A})$ 위의 $\sigma-$유한 측도라 하자. 그렇다면 임의의 $A \in \sigma(\mathscr{A})$와 임의의 $\varepsilon > 0$에 대해 다음이 성립한다.

i. Semi-algebra $\mathscr{A}$에 속하는 적당한 서로소인 집합열 $\{A_i\}$가 존재하여 $A \subseteq \bigsqcup_{i=1}^{\infty} A_i$이고 $\mu(\bigsqcup_{i=1}^{\infty} A_i \setminus A) < \varepsilon$이다.
ii. 만약 $\mu(A) < \infty$이면 적당한 서로소인 $A_1, \cdots, A_k \in \mathscr{A}$가 존재하여 $\mu(A \triangle \bigsqcup_{i=1}^{k} A_i) < \varepsilon$이다.

Proof i. 먼저 $\mu(A) < \infty$인 경우부터 생각하자. Carathéodory 확장의 유일성으로부터 $(\mu|_{A(\mathscr{A})})^*|_{\sigma(\mathscr{A})} = \mu$이므로 $A(\mathscr{A})$에 속하는 적당한 A의 가산 덮개 $\{A_i\}$가 존재하여, WLOG, 필요하다면 공집합을 추가하여 이가 집합열을 이룬다고 하면 $\mu(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \mu(A_i) < \mu(A) + \varepsilon$이다. 이제 집합열 $\{B_i\}$를 $B_i := A_i \setminus \bigcup_{j=1}^{i-1} A_j$로 두면 이는 $A(\mathscr{A})$에 속하는 서로소인 집합열이고 $\bigcup_{i=1}^{\infty} A_i = \bigsqcup_{i=1}^{\infty} B_i$이므로 WLOG, 필요하다면 $\{A_i\}$를 $\{B_i\}$로 바꾸어 처음부터 $\{A_i\}$가 서로소인 집합열이라 해도 된다. 한편, 각 $i \in \mathbb{N}$에 대해 $A_i \in A(\mathscr{A})$이므로 정리 ? ? 로부터 서로소인 $A_{i1}, \cdots, A_{ik_i} \in \mathscr{A}$가 존재하여 $A_i = \bigsqcup_{j=1}^{k_i} A_{ij}$이고, 곧 $\mu(\bigsqcup_{i=1}^{\infty} \bigsqcup_{j=1}^{k_i} A_{ij}) = \sum_{i=1}^{\infty} \mu(\bigsqcup_{j=1}^{k_i} A_{ij}) = \sum_{i=1}^{\infty} \mu(A_i) < \mu(A) + \varepsilon$에서 $\mu(\bigsqcup_{i=1}^{\infty} \bigsqcup_{j=1}^{k_i} A_{ij} \setminus A) = \mu(\bigsqcup_{i=1}^{\infty} \bigsqcup_{j=1}^{k_i} A_{ij}) - \mu(A) < \varepsilon$이 되어 정리가 성립한다.

다음으로 $\mu(A) = \infty$인 경우를 생각하면, μ가 $\sigma-$유한하므로 $\mathscr{A}$에 속하는 적당한 집합열 $\{B_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $\mu(B_i) < \infty$이고 $\bigcup_{i=1}^{\infty} B_i = X$이다. 그렇다면 앞선 결과로부터 임의의 $\varepsilon > 0$과 각 $i \in \mathbb{N}$에 대해 $\mathscr{A}$에 속하는 적당한 서로소인 집합열 $\{A_{ij}\}_j$가 존재하여 $A \cap B_i \subseteq \bigsqcup_{j=1}^{\infty} A_{ij}$이고 $\mu(\bigsqcup_{j=1}^{\infty} A_{ij} \setminus (A \cap B_i)) < \varepsilon/2^i$이다. 이상으로부터 $A = \bigcup_{i=1}^{\infty}(A \cap B_i) \subseteq \bigcup_{i=1}^{\infty} \bigsqcup_{j=1}^{\infty} A_{ij}$이고 $\mu(\bigcup_{i=1}^{\infty} \bigsqcup_{j=1}^{\infty} A_{ij} \setminus A) \le \mu(\bigcup_{i=1}^{\infty} [\bigsqcup_{j=1}^{\infty} A_{ij} \setminus (A \cap B_i)]) \le \sum_{i=1}^{\infty} \mu(\bigsqcup_{j=1}^{\infty} A_{ij} \setminus (A \cap B_i)) < \sum_{i=1}^{\infty} \varepsilon/2^i = \varepsilon$인데, 이전과 같은 방법으로 WLOG, $\{A_{ij}\}_{ij}$가 서로소인 집합열이라 가정할 수 있으므로 곧 모든 경우에 대해 정리가 성립한다.

ii. i의 $\mu(A) < \infty$인 경우에 대한 증명에서와 같이 $A(\mathscr{A})$에 속하는 $\{A_i\}$를 택하여 $A \subseteq \bigsqcup_{i=1}^{\infty} A_i$이고 $\mu(\bigsqcup_{i=1}^{\infty} A_i) < \mu(A) + \varepsilon/2$이도록 할 수 있다. 따라서 집합열 $\{B_i\}$를 $B_i := \bigsqcup_{j=1}^{i} A_j$로 두면 $B_i \uparrow \bigsqcup_{j=1}^{\infty} A_j =: B$에서 정리 ? ? 로부터 $\mu(B_i) \uparrow \mu(B)$이고, 곧 충분히 큰 $i_0 \in \mathbb{N}$를 택하면 $\mu(B \setminus B_{i_0}) < \varepsilon/2$가 되어 $\mu(A \triangle \bigsqcup_{i=1}^{i_0} A_i) = \mu(A \setminus \bigsqcup_{i=1}^{i_0} A_i) + \mu(\bigsqcup_{i=1}^{i_0} A_i \setminus A) \le \mu(B \setminus B_{i_0}) + \mu(\bigsqcup_{i=1}^{\infty} A_i \setminus A) < \varepsilon$이다. 이제 i의 증명에서와 같이 정리 ? ? 를 사용하면 원하는 결론을 얻는다. □

1.4 Construction of The Lebesgue Measure I

지금까지 Lebesgue 측도를 구성하기 위한 준비를 부단히 해 왔으니, 이제 때가 되었다. 이번 절에서는 전에 언급한 바와 같이 확장정리를 사용하여 Lebesgue 측도를 구성해보도록 하자.

Definition 1.42 구간 $I \subseteq \mathbb{R}$가 적당한 $x, y \in \mathbb{R}$에 대해 $(x, y]$, $(-\infty, y]$, (x, ∞), $(-\infty, \infty)$ 중 하나의 형태로 나타내어지면 이때의 I를 **반열린구간(semi-open interval)**이라 한다. 나아가, 만약 집합 $B \in \mathbb{R}^n$가 반열린구간들의 Cartesian 곱으로 표현되면 이때의 B를 **semi-open box**라 한다.

이제부터는 $\mathbb{R}^n$의 모든 semi-open box의 모임을 $\mathscr{S}_n$로 쓰도록 하겠다.

Proposition 1.43 집합족 $\mathscr{S}_n$은 semi-algebra이다.

Proof 간결한 논의를 위해 $n = 1$인 경우만 생각하도록 하자. (일반적인 경우에도 이와 비슷하게 보일 수 있다.) 우선 $\mathscr{S}_1$이 비어있지 않음은 분명하다. 이제 임의의 $B, B' \in \mathscr{S}_1$에 대해 정의로부터 각각이 $(x, y]$, $(-\infty, y]$, (x, ∞), $(-\infty, \infty)$ 중 하나의 형태로 나타내어지므로, 총 16가지의 가능성이 있다. 이 16가지 경우 각각에 대해 잠시 생각해보면, 어느 경우이든 $B \cap B'$이 다시 반열린구간임을 쉽게 알 수 있고, 따라서 $B \cap B' \in \mathscr{S}_1$이다. 한편, $\mathscr{S}_1$이 semi-algebra가 되기 위한 다른 조건들도 비슷하게 따져볼 수 있다. □

Proposition 1.44 함수 $\rho_n : \mathscr{S}_n \to \overline{\mathbb{R}}_0^+$를

$$\rho_n : B \mapsto \begin{cases} \displaystyle\prod_{i=1}^{n}(y_i - x_i) & \text{적당한 } x, y \in \mathbb{R}\text{에 대해 } B = (x, y]\text{인 경우} \\ \infty & \text{ow. 즉, 집합 } B\text{가 유계가 아닌 경우} \end{cases}$$

로 정의하면 이는 $\mathscr{S}_n$ 위의 premeasure이다.

Proof 우선 $\rho_n(\emptyset) = 0$임은 분명하므로, ρ_n이 σ−가법성을 가짐을 보이는 것으로 충분하다. 이를 위해 $\mathscr{S}_n$에 속하는 임의의 서로소인 집합열 $\{B_i\}$를 생각하여 $B := \bigsqcup_{i=1}^{\infty} B_i \in \mathscr{S}_n$이라 하자. 여기서 만약 B_{i_0}가 유계가 아닌 $i_0 \in \mathbb{N}$가 존재한다면 이 경우에 σ−가법성은 자명하므로 각 $i \in \mathbb{N}$에 대해 B_i가 유계라 가정하고, 따라서 적당한 $x_i, y_i \in \mathbb{R}^n$에 대해 $B_i = (x_i, y_i]$로 쓸 수 있다. 이제 경우를 나누어 $n = 1$인 경우를 먼저 생각하자.

$n = 1$인 경우에 B는 적당한 $x, y \in \mathbb{R}$에 대해 $(x, y]$, $(-\infty, y]$, (x, ∞), $(-\infty, \infty)$ 중 하나의 형태로 나타낼 수 있다. 첫 번째의 경우, 반드시 어떤 $i_0 \in \mathbb{N}$가 존재하여 $y_{i_0} = y$이므로 B_{i_0}부터 모든 $i \in \mathbb{N}$에 대해 $x_i = y_{i+1}$이고 $x_i \downarrow x$가 되도록 집합열 $\{B_i\}$를 재배열할 수 있다. 그렇다면 $\sum_{i=1}^{\infty} \rho_1(B_i) = \lim_{k\to\infty} \sum_{i=1}^{k}(y_i - x_i) = y - \lim_{k\to\infty} x_k = y - x = \rho_1(B)$가 성립한다. 두 번째의

경우도 첫 번째와 동일하게 재배열하여 같은 결론을 얻을 수 있다. (단, 이때는 $x_i \to -\infty$가 되도록 한다.) 세 번째 경우, 우선 (x,∞)를 $I=(x,x_1]$과 $J=(x_1,\infty)$의 두 개의 구간으로 나누고, J에 속하는 $\{B_i\}$의 항들만 추려내어 부분열 $\{B_{i_j}\}_j$를 구성한다. 그렇다면, B_1은 분명 그 부분열에 포함되어 있을 것이므로 B_1부터 모든 $j\in\mathbb{N}$에 대해 $y_{i_j}=x_{i_{j+1}}$이고 $y_{i_j}\to\infty$가 되도록 집합열 $\{B_{i_j}\}_j$를 재배열할 수 있다. 그렇다면 $\sum_{j=1}^{\infty}\rho_1(B_{i_j})=\lim_{k\to\infty}\sum_{j=1}^{k}(y_{i_j}-x_{i_j})=\lim_{k\to\infty}y_{i_k}-x_1=\infty$이고, 따라서 $\sum_{i=1}^{\infty}\rho_1(B_i)\geq\sum_{j=1}^{\infty}\rho_1(B_{i_j})=\infty=\rho_1(B)$가 성립한다. 마지막 경우도 세 번째 경우와 동일하게 재배열하여 같은 결론을 얻을 수 있다. (단, 이때는 B를 $I=(-\infty,x_1]$과 $J=(x_1,\infty)$의 두 개의 구간으로 나눈다.) 이상으로부터 어떠한 경우에도 ρ_1이 $\sigma-$가법성을 가짐을 안다.

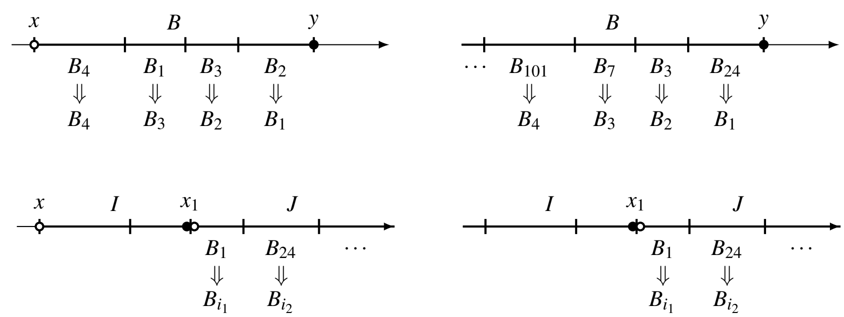


Figure 1.1 명제 ? ? 의 증명에서 $n=1$인 경우 B의 형태로 가능한 네 가지 경우에 대한 $\{B_i\}$의 재배열 과정. 각각 $B=(x,y]$ (왼쪽 위), $B=(-\infty,y]$ (오른쪽 위), $B=(x,\infty)$ (왼쪽 아래), $B=(-\infty,\infty)$ (오른쪽 아래)인 경우를 나타낸다.

이제 일반적인 $n\in\mathbb{N}$의 경우를 생각하면 적당한 반열린구간 $I_1,\cdots,I_n\subseteq\mathbb{R}$에 대해 $B=\prod_{j=1}^{n}I_j$로 쓸 수 있다. 우선 $\{B_i\}$가 B의 'regular tilling' 을 이루는 경우를 먼저 생각해보자. 여기서 regular tilling을 이룬다고 함은 $\{B_i\}$가 B를 딱 맞아떨어지게 채워넣는 경우를 이르는 것으로, 보다 엄밀하게, 각 $j\leq n$에 대해 I_j를 가산개의 반열린구간 $I_{j1},I_{j2},\cdots$로 분할할 수 있어서 이러한 분할들이 이루는 semi-open box들이 모두 정확히 하나의 B_i와 일치하는 경우를 이른다.

이 경우, 만약 B가 유계라면 곧 모든 $I_1,\cdots,I_n$도 유계이고, 앞선 결과로부터

$$\begin{aligned}\rho_n(B)&=\prod_{j=1}^{n}\rho_1(I_j)\\&=\prod_{j=1}^{n}\sum_{k=1}^{l_j}\rho_1(I_{jk})\end{aligned}$$

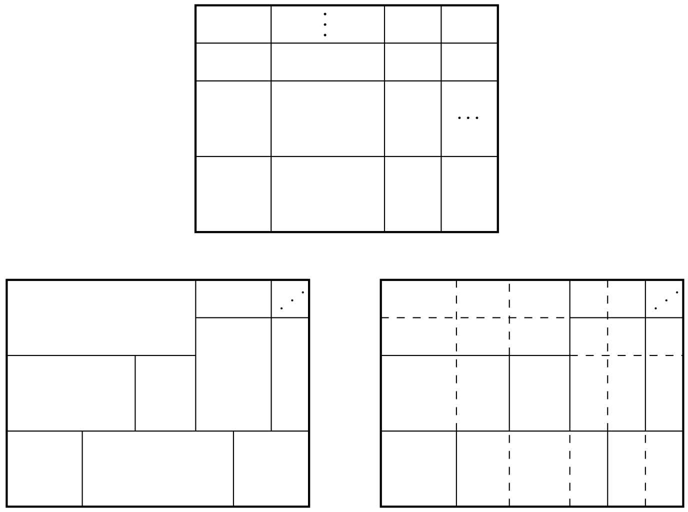


Figure 1.2 명제 ? ? 의 증명에서 집합열 $\{B_i\}$가 B의 regular tilling을 이루는 경우. (위) 집합열 $\{B_i\}$가 regular tilling을 이루지 않는 경우에 이의 각 면을 연장하여 regular tilling을 이루도록 하는 과정. (아래)

$$\begin{aligned}
&= \sum_{k_1=1}^{l_1} \cdots \sum_{k_n=1}^{l_n} \prod_{j=1}^{n} \rho_1(I_{jk_j}) \\
&= \sum_{k_1=1}^{l_1} \cdots \sum_{k_n=1}^{l_n} \rho_n\left(\prod_{j=1}^{n} I_{jk_j}\right) \\
&= \sum_{i=1}^{\infty} \rho_n(B_i)
\end{aligned}$$

임을 안다. (여기서 $l_1, \cdots, l_n$은 유한할 수도 있고, ∞일 수도 있다.) 만약 B가 유계가 아니라면 WLOG, I_1이 유계가 아니라 해도 되고, 이번에도 앞선 결과로부터

$$\begin{aligned}
\sum_{k=1}^{l_1} \rho_n\left(I_{1k} \times \prod_{j=2}^{n} I_{j1}\right) &= \left[\sum_{k=1}^{l_1} \rho_1(I_{1k})\right] \prod_{j=2}^{n} \rho_1(I_{j1}) \\
&= \rho_1(I_1) \prod_{j=2}^{n} \rho_1(I_{j1}) \\
&= \infty
\end{aligned}$$

가 되어 $\sum_{i=1}^{\infty} \rho_n(B_i) \geq \sum_{k=1}^{l_1} \rho_n(I_{1k} \times \prod_{j=2}^{n} I_{j1}) = \infty = \rho(B)$를 얻는다. 다음으로, $\{B_i\}$가 B의 regular tilling을 이루지 않는 경우를 생각해보자. 이 경우에는 각 B_i의 면을 연장하여 B를 가산개의 semi-open box $C_1, C_2, \cdots$로 분할할 수 있으며 WLOG, 필요하다면 공집합

을 추가하여 이가 집합열 $\{C_j\}$을 이룬다고 해도 된다. (공집합은 $\emptyset = \prod_{j=1}^{n}(0,0]$에서 명백히 semi-open box이다.) 그렇다면 모든 $j \in \mathbb{N}$에 대해 $C_j \subseteq B_i$인 $i \in \mathbb{N}$가 유일하게 존재하며, $\{C_j\}$는 모든 B_i와 B의 regular tilling을 이룬다. 따라서 $\rho_n(B) = \sum_{j=1}^{\infty} \rho_n(C_j) = \sum_{i=1}^{\infty} \sum_{j:C_j \subseteq B_i} \rho_n(C_j) = \sum_{i=1}^{\infty} \rho_n(B_i)$이고, 곧 어떠한 경우에도 ρ_n이 $\sigma-$가법성을 가짐을 알 수 있으므로 증명이 끝난다. □

이와 같이 정의된 $\mathscr{S}_n$ 위의 premeasure ρ_n이 우리의 출발점이다. 여기에 정리 ? ? 를 사용하면 $\mathscr{S}_n$ 위의 premeasure ρ_n을 $A(\mathscr{S}_n)$에서의 premeasure로 유일하게 확장할 수 있고, 이제부터는 표기를 남용하여 이러한 유일한 확장을 ρ_n이라 쓰도록 하겠다. 이러한 확장 ρ_n이 $\sigma-$유한함은 쉽게 확인할 수 있다.

Proposition 1.45 대수 $A(\mathscr{S}_n)$ 위의 premeasure ρ_n은 $\sigma-$유한하다.

Proof $\bigsqcup_{x \in \mathbb{Z}^n}(x, x+\mathbf{1}] = \mathbb{R}^n$이고 모든 $x \in \mathbb{Z}^n$에 대해 $\rho_n((x, x+\mathbf{1}]) = 1$이므로 이는 자명하다. □

마지막 단계로, premeasure ρ_n으로부터 유도되는 $\mathscr{P}(\mathbb{R}^n)$ 위의 외측도 ρ_n^*을 생각하고 $\mathscr{M}_n$을 모든 ρ_n^*-가측집합의 모임이라 하자. 그렇다면 Carathéodory의 확장정리와 정리 ? ? 로부터 $A(\mathscr{S}_n)$위의 premeasure ρ_n을 $\mathscr{M}_n$ 위의 측도로 유일하게 확장할 수 있고, 이가 우리가 그토록 원하던 Lebesgue 측도이다.

Definition 1.46 위에서와 같이 확장된 $A(\mathscr{S}_n)$위의 premeasure ρ_n에 대해 이의 Carathéodory 확장을 **Lebesgue 측도(– measure)**라 하고 λ_n으로 쓴다. 또한, 이때 λ_n의 정의역인 모든 ρ_n^*-가측집합들의 모임을 $\mathscr{M}_n$으로 쓰고, triple $(\mathbb{R}^n, \mathscr{M}_n, \lambda_n)$을 **Lebesgue 측도공간(– measure space)**이라 한다.

보통 특별히 정하지 않는 이상 $\mathbb{R}^n$에는 Lebesgue 측도 λ_n이 주어진 것으로 보고, 이는 이 책에서도 마찬가지이다.

이로써 우리의 첫 번째 Lebesgue 측도의 구성이 일단락되었다. 그러나 아직 몇몇 답하기 어려운 질문들이 남아있다. 그 중 하나로, $\sigma-$대수 $\mathscr{M}_n$을 한 번 살펴보자. 과연 이는 어떤 집합들로 이루어져 있을까? Lebesgue 측도의 구성 과정에서 $A(\mathscr{S}_n)$의 구조까지는 쉽게 상상이 가지만 Carathéodory의 확장정리의 다소 현란한 확장 방식 덕에 $\mathscr{M}_n$의 구조는 쉽게 손에 잡히지 않는다. 우리가 앞으로 기본적인 측도공간으로 삼을 Lebesgue 측도공간의 구조를 모른다는 것은 그다지 유쾌한 일이 아니며, 이는 곧 Lebesgue 측도의 구성은 마쳤지만 $\mathbb{R}^n$의 부분집합 중에서 Lebesgue 측도로써 잴 수 있는 집합과 그렇지 않은 집합을 구별할 수 있는 기준은 여전히 모호하다는 뜻이다.

사실, $\mathscr{M}_n$은 생각보다 굉장히 큰 $\sigma-$대수이다. 우리가 상식적인 선에서 생각해 낼 수 있는 거의 대부분의 집합, 예컨대 (표준 위상에 대해) 열린집합, 닫힌집합, 가산 집합, box나

ball, 혹은 $\mathbb{N}^n, \mathbb{Z}^n, \mathbb{Q}^n$ 따위의 집합들은 $\mathscr{M}_n$의 극히 일부분을 차지할 뿐이다. 집합족 $\mathscr{M}_n$의 대부분은 정말 상상조차 하기 힘든 괴상한 집합들로 이루어져 있어 $\mathscr{M}_n$은 그다지 경제적인 $\sigma-$대수는 못된다.[4]어쩌면 1절에서 지적한 것과 같이 측도를 부여할 집합으로 너무 많은 것들을 택한 것일 수도 있다. 그럼에도 불구하고 (이후에 자세히 보겠지만) 이 거대한 $\mathscr{M}_n$ 위에서 Lebesgue 측도는 놀라우리만치 바람직한 성질들을 가지기에 단순히 $\mathscr{M}_n$이 너무 크다는 이유로 힘들게 구성해낸 Lebesgue 측도공간을 포기하기는 뭔가 아쉽다. 이 상황을 타개할 현명한 해결책은 바로 $\mathscr{M}_n$ 중에서 이상한 집합들을 모두 걷어내고 '실용적' 인 집합들만 골라 $\mathscr{M}_n$ 보다 작은 $\sigma-$대수를 적당히 만든 다음 Lebesgue 측도를 그 $\sigma-$대수로 제한시키는 것이다. 이렇게 하면 이 새로운 측도공간은 Lebesgue 측도공간의 바람직한 성질들을 그대로 상속받으면서 훨씬 더 가볍고 경제적인 $\sigma-$대수로 구성되어 그 구조도 선명히 파악할 수 있을 것이다. 언듯 보기에 이 과정 또한 Lebesgue 측도의 구성 못지않게 꽤나 복잡할 것 같지만, 사실 $\mathbb{R}^n$의 열린집합을 모으는 것만으로 충분하다.

Definition 1.47 집합족 $\mathscr{U}$를 $\mathbb{R}^n$의 모든 열린집합의 모임이라 하자. 이때 $\mathscr{U}$가 생성하는 $\sigma-$대수를 **Borel $\sigma-$대수(– $\sigma-$algebra)**라 하고 $\mathscr{B}_n$으로 쓴다. 나아가, $\mathscr{B}_n$에 속하는 임의의 집합을 **Borel 집합(– set)**이라 한다.

Proposition 1.48 집합 $\mathbb{R}^n$의 열린집합이나 닫힌집합에 교집합, 합집합, 여집합, 차집합 혹은 이들을 가산번 조합하여 정의한 연산을 취하여 얻는 집합은 Borel이다.

Proof 임의의 닫힌집합 $F \subseteq \mathbb{R}^n$에 대해 그 정의로부터 F^c가 열린집합이므로 F는 Borel이다. 이제 $\mathscr{B}_n$이 $\sigma-$대수로서 교집합, 합집합, 여집합, 차집합 그리고 이들을 가산번 조합한 연산에 대해 모두 닫혀있음을 생각하면 이 명제는 자명하다. □

단순히 $\mathbb{R}^n$의 열린집합을 모아 이들로 $\sigma-$대수를 생성한 것 뿐인데 우리가 상식적으로 떠올릴 수 있는 거의 대부분이 집합이 Borel $\sigma-$대수에 속한다는 사실을 알 수 있다. 따라서 $\mathscr{B}_n \subseteq \mathscr{M}_n$만 성립한다면 모든 문제가 해결된다. 이는 $\mathscr{B}_n$의 생성자를 조금 더 조사해보면 바로 알 수 있다.

Theorem 1.49 집합족 $\mathscr{C}$를 $\mathbb{R}^n$의 모든 유계인 열린 box의 모임이라 하면 이는 $\mathscr{B}_n$의 생성자이다.

Proof 집합족 $\mathscr{U}$를 $\mathbb{R}^n$의 모든 열린집합의 모임이라 하면 $\mathscr{C} \subseteq \mathscr{U}$이므로 $\sigma(\mathscr{C}) \subseteq \mathscr{B}_n$임은 분명하다. 이제 $\mathbb{R}^n$의 모든 열린집합은 유계인 열린 box의 가산 합집합으로 표현된다는 점[5]을 생각해보면 $\mathscr{U} \subseteq \sigma(\mathscr{C})$이고, 곧 $\mathscr{B}_n \subseteq \sigma(\mathscr{C})$가 되어 증명이 끝난다. □

Theorem 1.50 집합족 $\mathscr{S}_n$은 $\mathscr{B}_n$의 생성자이다.

PROOF 간결한 논의를 위해 $n = 1$인 경우만 생각하자. (일반적인 경우에도 이와 비슷하게 보일 수 있다.) 또한 $\mathscr{C}$를 $\mathbb{R}$의 모든 유계인 열린 box의 모임이라 하자. 그렇다면 임의의 $B \in \mathscr{C}$에 대해 적당한 $x, y \in \mathbb{R}$가 존재하여 $B = (x, y)$이고, $(x, y) = \bigcup_{i=1}^{\infty}(x, y - 1/i]$에서 $\mathscr{C} \subseteq \sigma(\mathscr{S}_1)$이므로 곧 $\sigma(\mathscr{C}) \subseteq \sigma(\mathscr{S}_1)$이다. 이제 역의 포함관계를 보이기 위해 임의의 $I \in \mathscr{S}_1$를 택하면 정의로부터 I는 적당한 $x, y \in \mathbb{R}$에 대해 $(x, y], (-\infty, y], (x, \infty), (-\infty, \infty)$ 중 하나의 형태로 나타낼 수 있다. 이번에도 간결한 논의를 위해 첫 번째 경우만 생각해보자. (나머지 경우들도 비슷하게 따져보면 동일한 결론을 얻는다.) $I = (x, y]$라 하면 $I = (x, \infty) \setminus (y, \infty) = \bigcup_{i=1}^{\infty}(x, x+i) \setminus \bigcup_{i=1}^{\infty}(y, y+i)$이므로 $I \in \sigma(\mathscr{C})$이고, 곧 $\mathscr{S}_1 \subseteq \sigma(\mathscr{C})$가 되어 $\sigma(\mathscr{S}_1) \subseteq \sigma(\mathscr{C})$임을 안다. 이제 정리 ? ? 로부터 $\sigma(\mathscr{S}_1) = \mathscr{B}_1$이 되어 정리가 증명된다. □

여기서 $\mathscr{M}_n$이 $\mathscr{S}_n$을 포함한다는 사실을 떠올린다면, $\mathscr{B}_n \subseteq \mathscr{M}_n$임은 당연하다.[6]이제 우리는 다음을 정의할 수 있다.

Definition 1.51 Lebesgue 측도 λ_n에 대해 이의 $\mathscr{B}_n$으로의 제한 $\lambda_n|_{\mathscr{B}_n}$을 **Borel 측도(– measure)**라 하고 μ_n으로 쓴다. 나아가, triple $(\mathbb{R}^n, \mathscr{B}_n, \mu_n)$을 **Borel 측도공간(– measure space)**이라 한다.

나중에 사용하기 위해 $\mathscr{B}_n$의 생성자를 하나만 더 소개한다.

Theorem 1.52 집합족 $\mathscr{C}$를 어떤 $x \in \mathbb{R}^n$에 대해 $(-\infty, x]$로 표현되는 $\mathbb{R}^n$의 모든 부분집합의 모임이라 하면 이는 $\mathscr{B}_n$의 생성자이다.

PROOF 간결한 논의를 위해 $n = 1$인 경우만 생각하자. (일반적인 경우에도 이와 비슷하게 보일 수 있다.) 우선, $\mathscr{C} \subseteq \mathscr{S}_1$이므로 $\sigma(\mathscr{C}) \subseteq \sigma(\mathscr{S}_1)$임은 분명하다. 이제 역의 포함관계를 보이기 위해 임의의 $I \in \mathscr{S}_1$를 택하면 정의로부터 I는 적당한 $x, y \in \mathbb{R}$에 대해 $(x, y], (-\infty, y], (x, \infty), (-\infty, \infty)$ 중 하나의 형태로 나타낼 수 있다. 이번에도 간결한 논의를 위해 첫 번째 경우만 생각해보자. (나머지 경우들도 비슷하게 따져보면 동일한 결론을 얻는다.) $I = (x, y]$라 하면 $I = (-\infty, y] \setminus (-\infty, x]$이므로 $I \in \sigma(\mathscr{C})$이고, 곧 $\mathscr{S}_1 \subseteq \sigma(\mathscr{C})$가 되어 $\sigma(\mathscr{S}_1) \subseteq \sigma(\mathscr{C})$임을 안다. 이제 정리 ? ? 로부터 $\sigma(\mathscr{C}) = \mathscr{B}_1$이 되어 정리가 증명된다. □

Lebesgue 측도의 중요한 성질 몇 가지를 알아보자.

Theorem 1.53 임의의 유계인 $A \in \mathscr{M}_n$에 대해 $\lambda_n(A) < \infty$이다.

PROOF 집합 A가 유계이므로 적당한 $M > 0$에 대해 $A \subseteq (-M, M]^n$이고, 곧 $\lambda_n(A) \leq \lambda_n((-M, M]^n) = (2M)^n < \infty$에서 이 정리는 자명하다. □

Theorem 1.54 Lebesgue 측도 λ_n는 regular하다. 즉, 다음 두 가지 성질이 성립한다.

i. 임의의 $A \in \mathscr{M}_n$에 대해 $\lambda_n(A) = \inf\{\lambda_n(U) \in \overline{\mathbb{R}}_0^+ : A \subseteq U \subseteq \mathbb{R}^n, U$는 열린집합$\}$이다.

ii. 임의의 $A \in \mathscr{M}_n$에 대해 $\lambda_n(A) = \sup\{\lambda_n(K) \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K$는 compact 집합$\}$이다.

Proof i. 만약 $\lambda_n(A) = \infty$이면 어떠한 열린집합 $U \supseteq A$에 대해서도 $\lambda_n(U) \geq \lambda_n(A) = \infty$가 되어 정리가 자명하므로 $\lambda_n(A) < \infty$라 가정하자. 한편, 임의의 열린집합 $U \supseteq A$에 대해 $\lambda_n(U) \geq \lambda_n(A)$임이 분명하므로 임의의 $\varepsilon > 0$을 택하고 적당한 열린집합 $U \supseteq A$가 존재하여 $\lambda_n(U) < \lambda_n(A) + \varepsilon$임을 보이면 충분하다. 이를 위해, $\lambda_n(A)$의 정의가 $\lambda_n(A) = \inf\{\sum_{i=1}^{\infty} \rho_n(A_i) \in \overline{\mathbb{R}}_0^+ : \{A_i\}$는 $A(\mathscr{S}_n)$에 속하는 A의 가산 덮개이다.$\}$라는 점을 상기한다면 $A(\mathscr{S}_n)$에 속하는 적당한 A의 가산 덮개 $\{A_i\}$가 존재하여 WLOG, 필요하다면 공집합을 추가하여 이가 집합열을 이룬다고 하고 $\sum_{i=1}^{\infty} \rho_n(A_i) < \lambda_n(A) + \varepsilon$이다. 한편, 정리 ? ? 로부터 각 $i \in \mathbb{N}$에 대해 적당한 서로소인 $B_{i1}, \cdots, B_{ik_i} \in \mathscr{S}_n$이 존재하여 $A_i = \bigsqcup_{j=1}^{k_i} B_{ij}$이고, 따라서 WLOG, 필요하다면 A_i를 $B_{i1}, \cdots, B_{ik_i}$로 바꾸어 $\{A_i\}$가 처음부터 $\mathscr{S}_n$에 속한다고 해도 된다. 또한, $\lambda_n(A) < \infty$이므로 각 A_i도 유계이고 WLOG, 필요하다면 각 A_i를 조금씩 크게 하여 $A \subseteq \bigcup_{i=1}^{\infty} A_i^\circ := U$라 해도 된다.[7] 그렇다면 U가 열려있으므로 이는 가측이고, 곧 $\lambda_n(U) \leq \lambda_n(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \rho_n(A_i) \leq \lambda_n(A) + \varepsilon$이 되어 증명이 끝난다.

ii. 먼저 A가 유계인 경우를 생각하면 적당한 $r > 0$에 대해 $A \subseteq B(r)$이다. 한편, 임의의 compact한 $K \subseteq A$에 대해 $\lambda_n(K) \leq \lambda_n(A)$이므로 임의의 $\varepsilon > 0$을 택하고 적당한 compact 집합 $K \subseteq A$가 존재하여 $\lambda_n(K) > \lambda_n(A) - \varepsilon$임을 보이면 충분하다. 이를 위해 앞서 보인 i을 이용하면 적당한 열린집합 $U \supseteq \overline{B(r)} \setminus A$가 존재하여 $\lambda_n(U) < \lambda(\overline{B(r)} \setminus A) + \varepsilon$이다. 여기서 $K = \overline{B(r)} \setminus U$라 하면 이는 compact하고 $\overline{B(r)} \subseteq U \sqcup K$에서 $\lambda_n(U) + \lambda_n(K) = \lambda_n(U \sqcup K) \geq \lambda_n(\overline{B(r)})$인데, 이는 곧 $\lambda_n(K) \geq \lambda_n(\overline{B(r)}) - \lambda_n(U) > \lambda_n(\overline{B(r)}) - \lambda_n(\overline{B(r)} \setminus A) - \varepsilon = \lambda_n(A) - \varepsilon$임을 뜻하므로 이 경우에 대해서는 증명이 끝난다. 이제 일반적인 A에 대해 유계인 집합열 $\{A \cap B(i)\}_i$를 생각하면 $A \cap B(i) \uparrow A$이고 따라서 앞선 결론으로부터

$$\begin{aligned}
\lambda_n(A) &= \lim_{i \to \infty} \lambda_n(A \cap B(i)) \\
&= \sup_{r > 0} \lambda_n(A \cap B(r)) \\
&= \sup_{r > 0} \sup\{\lambda_n(K) \in \overline{\mathbb{R}}_0^+ : K \subseteq A \cap B(i), K\text{는 compact 집합}\} \\
&= \sup\{\lambda_n(K) \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K\text{는 compact 집합}\}
\end{aligned}$$

이 성립하여 일반적인 경우에 대해서도 증명이 끝난다. □

Theorem 1.55 Lebesgue 측도 λ_n는 이동 불변성을 갖는다. 즉, 임의의 $A \in \mathscr{M}_n$와 임의의 $x \in \mathbb{R}^n$에 대해 $A + x \in \mathscr{M}_n$이고 $\lambda_n(A + x) = \lambda_n(A)$이다.

Proof 만약 $A \in A(\mathscr{S}_n)$이라면 이 정리는 자명하다. 이제 임의의 $B \subseteq \mathbb{R}^n$에 대해 $\{B_i\}$를 $A(\mathscr{S}_n)$에 속하는 B의 가산 덮개라 하면 $\{B_i + x\}$는 $A(\mathscr{S}_n)$에 속하는 $B + x$의 가산 덮개가 되

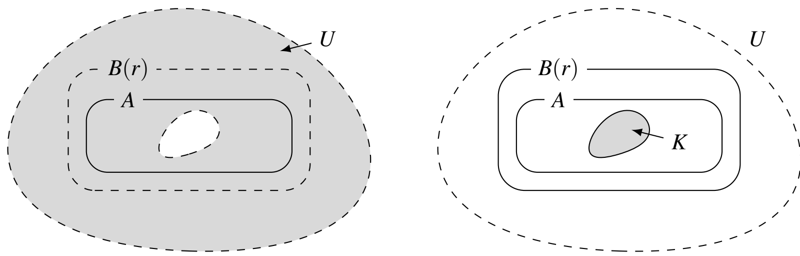


Figure 1.3 정리 ? ? 의 ii의 증명에서 집합 U (왼쪽)와 집합 K (오른쪽).

고, 이의 역도 성립하므로 $\rho_n^*(B) = \rho_n^*(B+x)$임을 안다. 이러한 결론이 임의의 $B \subseteq \mathbb{R}^n$에 대해 성립한다는 사실을 상기한다면, $A+x$가 가측이라는 사실만 보이는 것으로 충분함을 알 수 있다. 이를 위해 임의의 $S \subseteq \mathbb{R}^n$를 생각하면 A가 가측이라는 사실과 앞선 결과로부터 $\rho_n^*(S\cap(A+x)) + \rho_n^*(S\cap(A+x)^c) = \rho_n^*((S-x)\cap A) + \rho_n^*((S-x)\cap A^c) = \rho_n^*(S-x) = \rho_n^*(S)$가 되어 $A+x$가 가측임을 알고, 곧 증명이 끝난다. □

마지막으로 $\mathscr{B}_n$에 관한 놀라운 결과 하나를 소개하는 것으로 이번 절을 마치도록 하자. 이러한 결과는 $\mathscr{B}_n$이 $\mathscr{M}_n$보다 작은 σ−대수로 그 구조가 비교적 명확하다는 점에 크게 기인한다. ($\mathscr{B}_n$의 생성자는 서너개 알고 있는 반면, $\mathscr{M}_n$의 생성자나 구조는 여전히 오리무중이다.)

Definition 1.56 집합 $A \subseteq \mathbb{R}^n$에서 정의된 함수 $f : A \to \mathbb{R}$와 유계인 semi-open box $B \subseteq A$에 대해 B가 적당한 $x, y \in \mathbb{R}^n$에 대해 $B = (x, y]$와 같이 표현된다고 하자. 이때,

$$\sum_{0\le\alpha\le 1} (-1)^{|\alpha|} f(x_1^{1-\alpha_1} y_1^{\alpha_1}, \cdots, x_n^{1-\alpha_n} y_n^{\alpha_n})$$

를 f의 **B에서의 difference**라 하고 $\Delta_B f$로 쓴다.

Theorem 1.57 유한 측도공간 $(\mathbb{R}^n, \mathscr{B}_n, \mu)$에 대해 함수 $F_\mu : \mathbb{R}^n \to \mathbb{R}$를 $F_\mu : x \mapsto \mu((-\infty, x])$로 정의하면 다음이 성립한다.

i. 함수 F_μ는 $\mu(\mathbb{R}^n)$에 의해 위로 유계이다.
ii. 임의의 유계인 semi-open box $B \subseteq \mathbb{R}^n$에 대해 $\Delta_B F_\mu = \mu(B) \ge 0$이다.
iii. 함수 F_μ는 각 변수에 대해 증가한다.
iv. 함수 F_μ는 오른쪽 연속이다.
v. 함수 F_μ는 각 변수에 대해 오른쪽 연속이다.
vi. 각 $i \le n$에 대해 $\lim_{x_i\to-\infty} F_\mu(x) = 0$이다.
vii. $\lim_{x_1,\cdots,x_n\to\infty} F_\mu(x) = \mu(\mathbb{R}^n)$.

PROOF i. 임의의 $x \in \mathbb{R}^n$에 대해 $F_\mu(x) = \mu((-\infty, x]) \le \mu(\mathbb{R}^n) < \infty$이므로 이는 자명하다.

ii. 간결한 논의를 위해 $n = 2$인 경우만 생각하도록 하자. ($n = 1$인 경우를 포함한 일반적인 경우도 이와 비슷하게 보일 수 있다.) 그렇다면 적당한 $x, y \in \mathbb{R}^2$에 대해 $B = (x_1, y_1] \times (x_2, y_2]$로 쓸 수 있으므로

$$\begin{aligned}
\Delta_B F_\mu &= F_\mu(y_1, y_2) - F_\mu(x_1, y_2) - F_\mu(y_1, x_2) + F_\mu(x_1, x_2) \\
&= \mu((-\infty, y_1] \times (-\infty, y_2]) - \mu((-\infty, x_1] \times (-\infty, y_2]) \\
&\qquad\qquad - \mu((-\infty, y_1] \times (-\infty, x_2] + \mu((-\infty, x_1] \times (-\infty, x_2]) \\
&= \mu((-\infty, y_1] \times (-\infty, y_2] \setminus (-\infty, x_1] \times (-\infty, y_2]) \\
&\qquad\qquad - \mu((-\infty, y_1] \times (-\infty, x_2] \setminus (-\infty, x_1] \times (-\infty, x_2]) \\
&= \mu((x_1, y_1] \times (-\infty, y_2]) - \mu((x_1, y_1] \times (-\infty, x_2]) \\
&= \mu((x_1, y_1] \times (-\infty, y_2] \setminus (x_1, y_1] \times (-\infty, x_2]) \\
&= \mu((x_1, y_1] \times (x_2, y_2]) \\
&= \mu(B)
\end{aligned}$$

가 성립한다.

iii. 간결한 논의를 위해 F_μ가 첫 번째 변수에 대해 증가한다는 것만 보이자. (다른 변수에 대해서도 이와 비슷하게 보일 수 있다.) 이제 임의의 $x \in \mathbb{R}^n$와 임의의 $h \ge 0$에 대해 $B = (-\infty, x], B' = (-\infty, x_1 + h] \times \prod_{i=2}^n (-\infty, x_i]$라 하면 $F_\mu(x_1 + h, \cdots) - F_\mu(x_1, \cdots) = \mu(B') - \mu(B) = \mu(B' \setminus B) \ge 0$에서 증명이 끝난다.

iv. 임의의 $x \in \mathbb{R}^n$를 택하여 $B = (-\infty, x]$라 하고, 집합열 $\{B_j\}$를 $B_j := (-\infty, x + \mathbf{1}/j]$로 두면 이는 B로 수렴하는 $\mathscr{S}_n$에 속하는 감소하는 집합열이므로 정리 ? ? 의 ii에서 $\mu(B_j) \downarrow \mu(B) = F_\mu(x)$이고, 곧 임의의 $\varepsilon > 0$을 택하면 적당한 $j_0 \in \mathbb{N}$가 존재하여 $\mu(B_{j_0}) - F_\mu(x) < \varepsilon$이다. 이제 $\delta = 1/j_0$라 하면 $||x - y|| < \delta$이고 $x \le y$인 모든 $y \in \mathbb{R}^n$에 대해 $B \subseteq (-\infty, y] \subseteq B_{j_0}$에서 $F_\mu(x) = \mu(B) \le F_\mu(y) = \mu((-\infty, y]) \le \mu(B_{j_0}) < F_\mu(x) + \varepsilon$이므로 곧 $|F_\mu(x) - F_\mu(y)| < \varepsilon$이 되어 증명이 끝난다.

v. 이는 iv로부터 자명하다.

vi. 간결한 논의를 위해 $i = 1$인 경우에 대해서만 보이도록 하자. (다른 경우에도 이와 비슷하게 보일 수 있다.) 이제 임의의 $x_2, \cdots, x_n \in \mathbb{R}$을 택하고 집합열 $\{B_j\}$를 $B_j := (-\infty, -j] \times \prod_{i=2}^n (-\infty, x_i]$로 두면 이는 $\emptyset$으로 수렴하는 $\mathscr{S}_n$에 속하는 감소하는 집합열이므로 정리 ? ? 의 ii에서 $\mu(B_j) \downarrow \mu(\emptyset) = 0$이고, 곧 임의의 $\varepsilon > 0$을 택하면 적당한 $j_0 \in \mathbb{N}$가 존재하여 $\mu(B_{j_0}) < \varepsilon$이다. 이제 $x_1 \le -j_0$인 임의의 $x_1 \in \mathbb{R}$에 대해 $F_\mu(x_1, x_2, \cdots) \le F_\mu(-j_0, x_2, \cdots) = \mu(B_{j_0}) < \varepsilon$이 되어 증명이 끝난다.

vii. 집합열 $\{B_j\}$를 $B_j := (-\infty, j]^n$으로 두면 이는 $\mathbb{R}^n$으로 수렴하는 $\mathscr{S}_n$에 속하는 증가하는 집합열이므로 정리 ? ? 의 i에서 $\mu(B_j) \uparrow \mu(\mathbb{R}^n)$이고, 곧 임의의 $\varepsilon > 0$을 택하면 적당한 $j_0 \in \mathbb{N}$가 존재하여 $\mu(\mathbb{R}^n) - \mu(B_{j_0}) < \varepsilon$이다. 이제 $x \geq j_0\mathbf{1}$인 임의의 $x \in \mathbb{R}^n$에 대해 $B_{j_0} \subseteq (-\infty, x]$인데, 이와 i로부터 $\mu(\mathbb{R}^n) \geq F_\mu(x) = \mu((-\infty, x]) \geq F_\mu(j_0\mathbf{1}) = \mu(B_{j_0}) > \mu(\mathbb{R}^n) - \varepsilon$이므로 곧 $|F_\mu(x) - \mu(\mathbb{R}^n)| < \varepsilon$이 되어 증명이 끝난다. □

Theorem 1.58 함수 $F : \mathbb{R}^n \to \mathbb{R}$에 대해 이가

i. 함수 F는 오른쪽 연속이고 각 변수에 대해 증가한다.
ii. 유계인 semi-open box $B \subseteq \mathbb{R}^n$에 대해 $\Delta_B F \geq 0$이다.
iii. 각 $i \leq n$에 대해 $\lim_{x_i \to -\infty} F(x) = 0$이다.

를 만족하면 $F(x) = \mu((-\infty, x])$인 $\mathscr{B}_n$ 위의 측도 μ가 유일하게 존재한다.

Proof 집합족 $\mathscr{C}$를 어떤 $x \in \mathbb{R}^n$에 대해 $(-\infty, x]$로 표현되는 $\mathbb{R}^n$의 모든 부분집합의 모임이라 하자. 일단 정리의 조건을 만족하는 측도가 존재하기만 한다면 정리 ? ? 로부터 $\mathscr{C}$가 $\mathscr{B}_n$의 생성자이고 이는 $\pi-$system이므로 정리 ? ? 에 의해 그 측도의 유일성은 자명하다. 따라서 존재성을 보이는 것으로 증명은 충분하다. 이를 위해 우리는 Lebesgue 측도를 구성할 때 썼던 방법을 비슷하게 사용하려고 한다. 우선 $\mathscr{S}_n$ 위의 적당한 premeasure ρ를 정의하고, 이에 정리 ? ? 와 Carathéodory의 확장정리를 연달아 사용하여 ρ를 측도로 확장한 다음, 이를 다시 $\mathscr{B}_n$으로 제한하는 것이다. 또한, 간결한 논의를 위해 $n = 2$인 경우만 생각하도록 하자. ($n = 1$인 경우를 포함한 일반적인 경우에도 이와 비슷하게 보일 수 있다.)

먼저 함수 $\rho : \mathscr{S}_2 \to \overline{\mathbb{R}}_0^+$를

$$\rho : B \mapsto \begin{cases} \Delta_B F & \text{집합 } B\text{가 유계인 경우} \\ \infty & \text{ow.} \end{cases}$$

로 정의하면 $\rho(\emptyset) = 0$임은 분명하다. 이제 이가 $\sigma-$가법성을 가진다는 것만 보이면 ρ가 premeasure임을 알 수 있는데, 이를 위해서는 보조정리 ? ? 로부터 ρ가 가법성과 $\sigma-$반가법성을 가짐을 보이는 것으로 충분하다. 먼저 가법성을 보이기 위해 임의의 서로소인 $B_1, \cdots, B_l \in \mathscr{S}_2$를 생각하여 $B := \bigsqcup_{i=1}^l B_i \in \mathscr{S}_2$이라 하면 적당한 반열린구간 $I_1, I_2 \subseteq \mathbb{R}$에 대해 $B = I_1 \times I_2$로 쓸 수 있다. 만약 $B_1, \cdots, B_l$ 중에서 유계가 아닌 집합이 있다면 당연히 B도 유계가 아니게 되어 이 경우에는 가법성이 자명하므로 $B_1, \cdots, B_l$이 모두 유계라 하자. 이제 경우를 나누어 $B_1, \cdots, B_l$이 B의 regular tilling을 이루는 경우를 먼저 생각해보자. (Regular tilling을 이룬다는 것의 의미는 명제 ? ? 의 증명을 참조하기 바란다.) 이 경우에는 $\sum_{i=1}^l \rho(B_i) = \sum_{i=1}^l \Delta_{B_i} F = \Delta_B F = \rho(B)$가 되어 ρ가 가법성을 가진다. 이제 $B_1, \cdots, B_l$이 regular tilling을 이루지 않는 경우를 생각해보자. 이 경우에는 $B_1, \cdots, B_l$들의 각 면을 연장하여 B를 유한개의 semi-open box $C_1, \cdots, C_{l'}$로 분할할 수 있으며, 모든 $k \leq l'$

에 대해 $C_k \subseteq B_i$인 $i \le l$가 유일하게 존재하여 $C_1, \cdots, C_{l'}$은 모든 $B_1, \cdots, B_l$과 B의 regular tilling을 이룬다. 따라서 앞선 결과로부터 $\rho(B) = \Delta_B F = \sum_{k=1}^{l'} \Delta_{C_k} F = \sum_{i=1}^{l} \sum_{k:C_k \subseteq B_i} \Delta_{C_k} F = \sum_{i=1}^{l} \Delta_{B_i} F = \sum_{i=1}^{l} \rho(B_i)$가 되어 어느 경우에나 ρ는 가법성을 가짐을 안다.

다음으로 ρ가 $\sigma-$반가법성을 가짐을 보이기 위해 $\mathscr{S}_2$에 속하는 임의의 집합열 $\{B_i\}$를 생각하여 $B := \bigcup_{i=1}^{\infty} B_i \in \mathscr{S}_2$이라 하자. 한편, 만약 B_{i_0}가 유계가 아닌 $i_0 \in \mathbb{N}$가 존재한다면 당연히 B도 유계가 아니게 되어 이 경우에는 $\sigma-$반가법성이 자명하므로 모든 $i \in \mathbb{N}$에 대해 B_i가 유계라고 하자. 비슷하게, B가 유계가 아닌 경우에도 $\sigma-$반가법성이 자명하므로 B도 유계라 하자. 그렇다면 각 $i \in \mathbb{N}$에 대해 적당한 $x_i, y_i \in \mathbb{R}^2$가 존재하여 $B_i = (x_i^1, y_i^1] \times (x_i^2, y_i^2]$로 쓸 수 있고, 비슷하게 적당한 $x, y \in \mathbb{R}^2$가 존재하여 $B = (x^1, y^1] \times (x^2, y^2]$로 쓸 수 있다. 한편, F가 오른쪽 연속이므로 임의의 $\varepsilon > 0$을 택하면 각 $i \in \mathbb{N}$에 대해 적당한 $\delta_i > 0$가 존재하여

$$
\begin{aligned}
&|F(y_i^1 + \delta_i, y_i^2 + \delta_i) - F(y_i^1, y_i^2)| < \varepsilon/3 \cdot 2^{i+1}, \\
&\quad |F(x_i^1, y_i^2 + \delta_i) - F(x_i^1, y_i^2)| < \varepsilon/3 \cdot 2^{i+1}, \\
&\quad |F(y_i^1 + \delta_i, x_i^2) - F(y_i^1, x_i^2)| < \varepsilon/3 \cdot 2^{i+1}
\end{aligned}
$$

이고, 곧 $B_i' := (x_i^1, y_i^1 + \delta_i] \times (x_i^2, y_i^2 + \delta_i] \supseteq B_i$로 두면

$$
\begin{aligned}
|\rho(B_i') - \rho(B_i)| &= |\Delta_{B_i'} F - \Delta_{B_i} F| \\
&= |[F(y_i^1 + \delta_i, y_i^2 + \delta_i) - F(x_i^1, y_i^2 + \delta_i) - F(y_i^1 + \delta_i, x_i^2) + F(x_i^1, x_i^2)] \\
&\qquad - [F(y_i^1, y_i^2) - F(x_i^1, y_i^2) - F(y_i^1, x_i^2) + F(x_i^1, x_i^2)]| \\
&\le |F(y_i^1 + \delta_i, y_i^2 + \delta_i) - F(y_i^1, y_i^2)| + |F(x_i^1, y_i^2 + \delta_i) - F(x_i^1, y_i^2)| \\
&\qquad + |F(y_i^1 + \delta_i, x_i^2) - F(y_i^1, x_i^2)| \\
&< \varepsilon/2^{i+1}.
\end{aligned}
$$

이 성립한다. 비슷하게,

$$
\begin{aligned}
&|F(x^1 + \delta, x^2 + \delta) - F(x^1, x^2)| < \varepsilon/6, \\
&\quad |F(x^1 + \delta, y^2) - F(x^1, y^2)| < \varepsilon/6, \\
&\quad |F(y^1, x^2 + \delta) - F(y^1, x^2)| < \varepsilon/6
\end{aligned}
$$

인 $\delta > 0$도 존재하여 $B' := (x^1 + \delta, y^1] \times (x^2 + \delta, y^2] \subseteq B$로 두면 $|\rho(B') - \rho(B)| < \varepsilon/2$이 성립한다. 여기서 $B' \subseteq \overline{B'} \subseteq B \subseteq \bigcup_{i=1}^{\infty} (B_i')^\circ$이므로 $\{(B_i')^\circ\}$는 compact한 $\overline{B'}$의 열린 덮개가 되어 이의 유한 부분덮개가 존재하고, WLOG, 필요하다면 relabel하여 그 유한 부분덮개를 $\{(B_1')^\circ, \cdots, (B_k')^\circ\}$라 할 수 있다. 이제 $B' \subseteq \bigcup_{i=1}^{k} (B_i')^\circ \subseteq \bigcup_{i=1}^{k} B_i'$에 보조정리 ? ?

의 ii를 적용하면 $\rho(B) < \rho(B') + \varepsilon/2 \le \sum_{i=1}^{k} \rho(B_i') + \varepsilon/2 < \sum_{i=1}^{\infty} [\rho(B_i) + \varepsilon/2^{i+1}] + \varepsilon/2 = \sum_{i=1}^{\infty} \rho(B_i) + \varepsilon$이 되어 ρ가 $\sigma-$반가법성을 가짐을 안다.

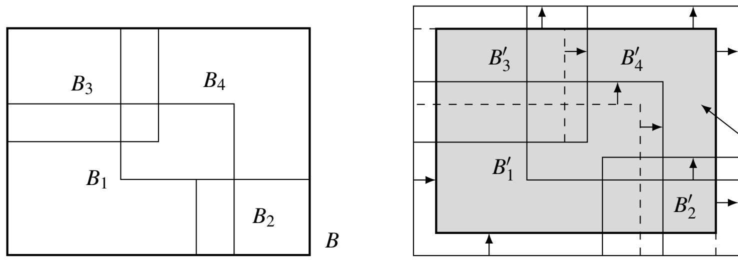


Figure 1.4 정리 ? ? 의 증명에서의 집합 B, B_i (왼쪽)와 집합 B', B_i' (오른쪽).

이제 ρ가 $\mathscr{S}_2$ 위의 premeasure임을 알았으므로 앞서 설명한 바와 같이 이를 $\mathscr{B}_2$ 위의 측도로 확장하는 과정은 자연스럽게 이루어진다. 남은 일은 과연 이러한 확장을 통해 얻은 측도 μ가 정리의 조건을 만족하는지를 확인하는 일인데, 이는 어렵지 않다. 우선 임의의 $x \in \mathbb{N}$에 대해 집합열 $\{B_i\}$를 $B_i := (x^1 - i, x^1] \times (x^2 - i, x^2]$로 두면 이는 $B_i \uparrow (-\infty, x^1] \times (-\infty, x^2]$인 $\mathscr{S}_2$에 속하는 증가하는 집합열임이 분명하고 μ의 구성 과정으로부터 $\mu(B_i) = \Delta_{B_i} F = F(x^1, x^2) - F(x^1 - i, x^2) - F(x^1, x^2 - i) + F(x^1 - i, x^2 - i)$이다. 그런데 가정으로부터 $F(x^1 - i, x^2), F(x^1, x^2 - i) \to 0$이고 F의 함숫값은 항상 0보다 크거나 같으므로(만약 어떤 $x \in \mathbb{R}^2$에 대해 $F(x) < 0$이라면 F가 첫 번째 변수에 대해 감소한다는 사실로부터 $0 = \lim_{i\to\infty} F(x^1 - i, x^2) \le F(x) < 0$의 모순이 발생하므로 모든 $x \in \mathbb{R}^2$에 대해 $F(x) \ge 0$이다.) $F(x^1 - i, x^2 - i) \le F(x^1 - i, x^2) \to 0$에서 $F(x^1 - i, x^2 - i) \to 0$이 되어 $\mu((-\infty, x^1] \times (-\infty, x^2]) = \lim_{i\to\infty} \mu(B_i) = F(x)$임을 안다. □

이상의 결과는 가측공간 $(\mathbb{R}^n, \mathscr{B}_n)$ 위의 유한 측도 μ에 대한 정보는 정리 ? ? 에서와 같이 정의된 함수 F_μ에 모두 담겨있음을 뜻한다. 따라서 F_μ만 알고 있다면 원래의 측도 μ를 완벽히 복원해낼 수 있다. 이 놀라운 결과는 한참 후에 CDF를 소개하며 다시 등장할 것이다.

1.5 Completion of Measures

측도가 0인 집합을 생각해보자. 상식적으로, 이 집합의 모든 부분집합은 0의 측도를 가질 것 같지만, 아쉽게도 이가 항상 성립하는 것은 아니다. 어떤 부분집합은 가측이 아닐수도 있기 때문이다. 이렇게 측도가 0인 집합의 부분집합이 가측이 아닐 수도 있다는 사실은 측

도론의 전개에 있어 상당히 신경쓰이는 부분이다. 이번 절에서는 Lebesgue 측도의 구성이라는 우리의 목표에서 잠시 옆으로 빠져 이런 문제를 해결하는 방법에 대해 알아본다. 그리고 놀랍게도, 이를 공부함으로써 우리에게 숙제로 남겨진 $\mathscr{M}_n$의 구조에 대한 질문에 어느 정도 만족스러운 답을 할 수 있게 될 것이다.

Definition 1.59 측도공간 $(X, \mathscr{A}, \mu)$와 부분집합 $N \subseteq X$에 대해 만약 $\mu(A) = 0$이고 $N \subseteq A$인 $A \in \mathscr{A}$가 존재하면 이때의 N을 **$(\mu-)$영집합($(\mu-)$null set)**이라 한다. 또한, 만약 모든 영집합이 $\mathscr{A}-$가측이면 이때의 측도공간 $(X, \mathscr{A}, \mu)$를 **완비측도공간(complete measure space)**이라 한다.

어떤 측도공간을 완비측도공간으로 만드는 체계적인 방법 즉, 완비화 과정을 찾는 것이 이번 절에서의 우리의 목표이다. 우선 영집합에 대한 몇몇 기본적인 성질들을 보자.

Proposition 1.60 측도공간 $(X, \mathscr{A}, \mu)$에 대해 다음이 성립한다.

i. 영집합 N의 임의의 부분집합 $M \subseteq N$도 영집합이다.
ii. 영집합으로 구성된 집합열 $\{N_i\}$에 대해 $\bigcup_{i=1}^{\infty} N_i$도 영집합이다.

Proof i. 정의로부터 적당한 $A \in \mathscr{A}$가 존재하여 $M \subseteq N \subseteq A$이고 $\mu(A) = 0$이므로 이 명제는 자명하다.

ii. 이번에도 정의로부터 각 $i \in \mathbb{N}$에 대해 적당한 $A_i \in \mathscr{A}$가 존재하여 $N_i \subseteq A_i$이고 $\mu(A_i) = 0$이므로 곧 $\bigcup_{i=1}^{\infty} N_i \subseteq \bigcup_{i=1}^{\infty} A_i$이고 $\mu(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i) = 0$이 되어 $\bigcup_{i=1}^{\infty} N_i$도 영집합이다. □

영집합이라는 용어에서 알 수 있듯이 영집합은 '거의' 공집합과 같다. 물론, 아예 아무런 원소를 포함하지 않는 빈 집합인 공집합과 원소를 포함할 수는 있지만 이들이 마치 여기저기 흩어진 점과 같아 의미있는 측도를 이루지 못하는 영집합이 엄밀하게는 서로 다르지만, 이러한 차이점은 많은 경우에 흐려져 영집합을 마치 공집합과 같이 사용하는 경우가 측도론에서는 흔하다. 특히 어떤 성질을 만족하지 않는 경우가 영집합을 이루는 경우, 마치 이런 예외가 아예 없는 것과 같이 생각하곤 한다. 이는 Lebesgue 적분론을 배우기 시작하면 충분히 느낄 수 있으니, 일단은 앞서 설명한 것과 같이 완비화 과정을 정립하는 것에 집중하자.

Proposition 1.61 측도공간 $(X, \mathscr{A}, \mu)$와 모든 영집합들의 모임 $\mathscr{N}$에 대해 $\sigma(\mathscr{A} \cup \mathscr{N}) = \{A \cup N \subseteq X : A \in \mathscr{A}, N \in \mathscr{N}\}$이다.

Proof 표기의 편의를 위해 $\overline{\mathscr{A}} = \{A \cup N \subseteq X : A \in \mathscr{A}, N \in \mathscr{N}\}$라 하자. 그렇다면 $\emptyset \in \mathscr{A} \cap \mathscr{N}$에서 $\mathscr{A} \cup \mathscr{N} \subseteq \overline{\mathscr{A}}$임은 분명하다. 또한, $\mathscr{A} \cup \mathscr{N}$을 포함하는 모든 $\sigma-$대수가 $\overline{\mathscr{A}}$를 포함해야 함도 분명하므로 $\overline{\mathscr{A}}$가 $\sigma-$대수임을 보이는 것으로 증명은 충분하다. 우선 $\mathscr{A}$가 비

어있지 않으므로 $\overline{\mathscr{A}}$도 비어있지 않다. 한편, 임의의 $B \in \overline{\mathscr{A}}$에 대해 적당한 $A \in \mathscr{A}, N \in \mathscr{N}$이 존재하여 $B = A \cup N$인데, WLOG, 필요하다면 N을 $N \setminus A$로 바꾸어 A와 N이 처음부터 서로소라고 해도 된다. 또한, 정의로부터 적당한 $C \in \mathscr{A}$가 존재하여 $\mu(C) = 0$이고 $N \subseteq C$인데, $\mu(C \setminus A) \leq \mu(C) = 0$이고 $N \subseteq C \setminus A$이므로 WLOG, 필요하다면 C를 $C \setminus A$로 바꾸어 A와 C가 처음부터 서로소라고 해도 된다. 그렇다면 $B^c = (A \sqcup N)^c = (A \sqcup C)^c \cup (C \setminus N)$에서 $B^c \in \overline{\mathscr{A}}$가 되어 $\overline{\mathscr{A}}$가 여집합에 대해 닫혀있음을 안다. 마지막으로 $\overline{\mathscr{A}}$에 속하는 임의의 집합열 $\{B_i\}$에 대해 이번에도 정의로부터 각 $i \in \mathbb{N}$에 대해 적당한 $A_i \in \mathscr{A}, N_i \in \mathscr{N}$가 존재하여 $B_i = A_i \cup N_i$로 쓸 수 있다. 그렇다면 $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i \cup \bigcup_{i=1}^{\infty} N_i$에서 명제 ? ? 의 ii로부터 $\bigcup_{i=1}^{\infty} N_i$도 영집합이므로 $\bigcup_{i=1}^{\infty} B_i \in \overline{\mathscr{A}}$가 되어 $\overline{\mathscr{A}}$가 가산 합집합에 대해서도 닫혀있음을 알고, 따라서 이는 $\sigma-$대수이다. □

Proposition 1.62 측도공간 $(X, \mathscr{A}, \mu)$에 대해 $A, B \in \mathscr{A}$이고 $N, M \subseteq X$이 영집합이라 하자. 만약 $A \cup N = B \cup M$이면 $\mu(A) = \mu(B)$이다.

Proof 집합 M이 영집합이므로 적당한 $C \in \mathscr{A}$가 존재하여 $\mu(C) = 0$이고 $M \subseteq C$이다. 이는 곧 $A \subseteq A \cup N = B \cup M \subseteq B \cup C$에서 $\mu(A) \leq \mu(B \cup C) \leq \mu(B) + \mu(C) = \mu(B)$임을 의미하는데, 이와 비슷하게 $\mu(A) \geq \mu(B)$도 보일 수 있으므로 증명이 끝난다. □

Theorem 1.63 측도공간 $(X, \mathscr{A}, \mu)$와 모든 $\mu-$영집합들의 모임 $\mathscr{N}$에 대해 $\overline{\mathscr{A}} = \{A \cup N \subseteq X : A \in \mathscr{A}, N \in \mathscr{N}\}$라 하고 함수 $\overline{\mu} : \overline{\mathscr{A}} \to \overline{\mathbb{R}}_0^+$를 $\overline{\mu} : A \cup N \mapsto \mu(A)$로 정의하자. (여기서 $A \in \mathscr{A}, N \in \mathscr{N}$이다.) 그렇다면 $\overline{\mu}$는 μ의 $\overline{\mathscr{A}}$에서의 측도로의 유일한 확장이고 $(X, \overline{\mathscr{A}}, \overline{\mu})$는 완비측도공간이 된다.

Proof 명제 ? ? , ? ? 로부터 $\overline{\mathscr{A}}$는 $\mathscr{A}$를 포함하는 $\sigma-$대수이며 $\overline{\mu}$는 well-define된다. 이제 $\overline{\mu}$가 $\overline{\mathscr{A}}$ 위의 측도임을 보이자. 이를 위해 $\overline{\mathscr{A}}$에 속하는 임의의 서로소인 집합열 $\{B_i\}$를 생각하면 각 $i \in \mathbb{N}$에 대해 적당한 $A_i \in \mathscr{A}, N_i \in \mathscr{N}$가 존재하여 $B_i = A_i \cup N_i$이고 $\{A_i\}$와 $\{N_i\}$가 명백히 서로소이므로, $\overline{\mu}(\bigsqcup_{i=1}^{\infty} B_i) = \overline{\mu}(\bigsqcup_{i=1}^{\infty} A_i \cup \bigsqcup_{i=1}^{\infty} N_i) = \mu(\bigsqcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \overline{\mu}(B_i)$이다. 한편, $\overline{\mu}(\emptyset) = 0$임은 분명하므로 $\overline{\mu}$는 명백히 $\overline{\mathscr{A}}$ 위의 측도이다. 그렇다면 $\emptyset$이 $\mu-$영집합이라는 사실로부터 $\overline{\mu}$가 μ의 확장임은 자명하다.

다음으로, 측도공간 $(X, \overline{\mathscr{A}}, \overline{\mu})$의 완비성을 보이자. 이를 위해 임의의 $B \in \overline{\mathscr{A}}$를 택하여 $\overline{\mu}(B) = 0$이라 하면, 적당한 $A \in \mathscr{A}, N \in \mathscr{N}$이 존재하여 $B = A \cup N$이고 $\mu(A) = \overline{\mu}(B) = 0$이다. 이는 $A \in \mathscr{N}$임을 뜻하므로 임의의 $M \subseteq B$에 대해 $M \in \mathscr{N} \subseteq \overline{\mathscr{A}}$이고, 곧 B의 임의의 부분집합이 $\overline{\mathscr{A}}-$가측이 되어 $(X, \overline{\mathscr{A}}, \overline{\mu})$가 완비측도공간임을 안다.

마지막으로 유일성을 보이기 위해 ν를 μ의 $\overline{\mathscr{A}}$에서의 측도로의 확장이라 하고 $B \in \overline{\mathscr{A}}$를 임의로 하나 택하자. 그렇다면 적당한 $A \in \mathscr{A}$와 $N \in \mathscr{N}$이 존재하여 $B = A \cup N$이고, 다시 적당한 $C \in \mathscr{A}$가 존재하여 $N \subseteq C$이고 $\mu(C) = 0$이다. 이로부터 $\nu(B) = \nu(A \cup N) \geq \nu(A) = \mu(A) = \overline{\mu}(B)$이고 $\nu(B) = \nu(A \cup N) \leq \nu(A) + \nu(N) \leq \nu(A) + \nu(C) = \mu(A) = \overline{\mu}(B)$에서 $\nu = \overline{\mu}$가 되어 원하던 유일성을 얻는다. □

이상으로부터 다음과 같은 well-define된 완비화를 얻을 수 있다.

Definition 1.64 측도공간 $(X, \mathscr{A}, \mu)$와 모든 $\mu-$영집합들의 모임 $\mathscr{N}$에 대해 $\overline{\mathscr{A}} = \{A \cup N \subseteq X : A \in \mathscr{A}, N \in \mathscr{N}\}$라 하고 함수 $\overline{\mu} : \overline{\mathscr{A}} \to \overline{\mathbb{R}}_0^+$를 $\overline{\mu} : A \cup N \mapsto \mu(A)$로 정의하자. (여기서 $A \in \mathscr{A}, N \in \mathscr{N}$이다.) 이때의 완비측도공간 $(X, \overline{\mathscr{A}}, \overline{\mu})$를 측도공간 $(X, \mathscr{A}, \mu)$의 **완비화 (completion)**라 한다.

위의 정의에서와 같이 일반적으로 측도공간의 완비화를 통해 얻는 $\sigma-$대수와 측도는 각각 원래의 $\sigma-$대수와 측도에 overline을 그어 표기한다. 나중을 위해 정리 하나를 소개한다.

Theorem 1.65 측도공간 $(X, \mathscr{A}, \mu)$와 부분집합 $A \subseteq X$에 대해 $A \in \overline{\mathscr{A}}$일 필요충분조건은 적당한 $B, C \in \mathscr{A}$가 존재하여 $B \subseteq A \subseteq C$이고 $\mu(C \setminus B) = 0$인 것이다.

PROOF 충분조건임을 보이기 위해 $A \in \overline{\mathscr{A}}$라 하면 적당한 $B \in \mathscr{A}$와 영집합 $N \subseteq X$이 존재하여 $A = B \cup N$이며, 다시 적당한 $D \in \mathscr{A}$가 존재하여 $N \subseteq D$이고 $\mu(D) = 0$이다. 이제 $C = B \cup D$라 하면 $B \subseteq A \subseteq C$이고 $\mu(C \setminus B) \leq \mu(D) = 0$이 되어 충분조건임이 보여진다. 다음으로, 필요조건임을 보이기 위해 적당한 $B, C \in \mathscr{A}$가 존재하여 $B \subseteq A \subseteq C$이고 $\mu(C \setminus B) = 0$이라 하자. 그렇다면 $A = (A \setminus B) \sqcup B$이고 $A \setminus B \subseteq C \setminus B$에서 $A \setminus B$가 영집합이므로 $A \in \overline{\mathscr{A}}$가 되어 필요조건임도 보여진다. □

이번 절의 남은 부분에서는 앞서 예고했던 바와 같이 이러한 완비화가 $\mathscr{M}_n$의 구조에 대한 질문으로 어떻게 연결되는지 살펴보려고 한다. 우선 Carathéodory의 확장정리로 잠시 돌아가자.

Theorem 1.66 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하고, $\mathscr{M}$을 모든 ρ^*-가측집합들의 모임이라 하자. 또한 μ를 ρ의 Carathéodory 확장이라 하면 $(X, \mathscr{M}, \mu)$는 완비측도공간이다.

PROOF 집합 $N \subseteq X$이 영집합이라 하면 정의로부터 적당한 $A \subseteq \mathscr{M}$가 존재하여 $\mu(A) = 0$이고 $N \subseteq A$이므로 임의의 $S \subseteq X$에 대해 $\rho^*(S \cap N) + \rho^*(S \cap N^c) \leq \rho^*(A) + \rho^*(S \cap N^c) = \mu(A) + \rho^*(S \cap N^c) = \rho^*(S \cap N^c) \leq \rho^*(S)$가 성립하여 $N \in \mathscr{M}$이고, 곧 $(X, \mathscr{M}, \mu)$는 완비측도공간임을 안다. □

따라서 Carathéodory의 확장정리를 통해 얻는 모든 측도공간은 완비성을 가지고, 특히 Lebesgue 측도공간은 완비측도공간이다. 그렇다면 조심스럽게, 어쩌면 Lebesgue 측도공간이 Borel 측도공간의 완비화가 아닐까하는 추측을 해 볼 수도 있다. 이 추측의 결론은 다음과 같다.

Theorem 1.67 Lebesgue 측도공간 $(\mathbb{R}^n, \mathscr{M}_n, \lambda_n)$은 Borel 측도공간 $(\mathbb{R}^n, \mathscr{B}_n, \mu_n)$의 완비화이다. 즉, $(\mathbb{R}^n, \mathscr{M}_n, \lambda_n) = (\mathbb{R}^n, \overline{\mathscr{B}_n}, \overline{\mu_n})$.

Proof 정리 ? ? 로부터 완비화를 통해 얻는 σ−대수로의 측도의 확장은 유일하므로 $\overline{\mathscr{B}_n} = \mathscr{M}_n$을 보이는 것으로 충분하다. 우선 $\overline{\mathscr{B}_n} \subseteq \mathscr{M}_n$부터 보이자. 임의의 집합 $A \in \overline{\mathscr{B}_n}$를 택하면 적당한 $B \in \mathscr{B}_n$와 μ_n−영집합 $N \subseteq \mathbb{R}^n$이 존재하여 $A = B \cup N$이고, 다시 적당한 $C \in \mathscr{B}_n$가 존재하여 $N \subseteq C$이고 $\mu_n(C) = 0$이다. 이제 Lebesgue 측도공간은 정리 ? ? 로부터 완비측도공간이므로 $N \in \mathscr{M}_n$이 되어 $A = B \cup N \in \mathscr{M}_n$에서 곧 $\overline{\mathscr{B}_n} \subseteq \mathscr{M}_n$이다.

역의 포함 관계를 보이는 것은 조금 더 복잡하다. 임의의 $A \in \mathscr{M}_n$를 택하고 먼저 $\lambda_n(A) < \infty$인 경우를 생각하자. 이 경우, 정리 ? ? 로부터 각 $i \in \mathbb{N}$에 대해 compact한 $K_i \in \mathscr{B}_n$와 열린 $U_i \in \mathscr{B}_n$가 존재하여 $K_i \subseteq A \subseteq U_i$이고 $\lambda_n(A) - 1/i < \lambda_n(K_i) \leq \lambda_n(A) \leq \lambda_n(U_i) < \lambda_n(A) + 1/i$이므로 이들로써 집합열 $\{K_i\}, \{U_i\}$를 구성할 수 있으며 WLOG, 필요하다면 각각의 항을 $F_i := \bigcup_{j=1}^{i} K_j, V_i := \bigcap_{j=1}^{i} U_j$로 바꾸어 $\{K_i\}, \{U_i\}$ 각각을 증가하고 감소하는 집합열이라 해도 된다. 그렇다면 그 구성으로부터 $\{K_i\}, \{U_i\}$는 각각 $K := \bigcup_{i=1}^{\infty} K_i, U := \bigcap_{i=1}^{\infty} U_i$로 수렴하고, 곧 $K \subseteq A \subseteq U$이며 $\lambda_n(K) = \lambda_n(A) = \lambda_n(U)$이다. 따라서 $\mu_n(U \setminus K) = \lambda_n(U) - \lambda_n(K) = 0$이 되어 정리 ? ? 로부터 $A \in \overline{\mathscr{B}_n}$임을 안다. 한편, $\lambda_n(A) = \infty$인 경우에는 각 $i \in \mathbb{N}$에 대해 $A \cap B(i)$가 유계이므로 정리 ? ? 로부터 이는 유한한 Lebesgue 측도를 가지고, 곧 앞선 결과로부터 $A \cap B(i) \in \overline{\mathscr{B}_n}$이 되어 적당한 $B_i \in \mathscr{B}_n$과 μ_n−영집합 $N_i \subseteq \mathbb{R}^n$가 존재하여 $A \cap B(i) = B_i \cup N_i$이다. 이로부터 $A = \bigcup_{i=1}^{\infty} [A \cap B(i)] = \bigcup_{i=1}^{\infty} B_i \cup \bigcup_{i=1}^{\infty} N_i$인데 명제 ? ? 의 ii에서 $\bigcup_{i=1}^{\infty} N_i$도 μ_n−영집합이므로 $A \in \overline{\mathscr{B}_n}$임을 안다. 곧 어느 경우에나 $A \in \overline{\mathscr{B}_n}$이므로 $\overline{\mathscr{B}_n} \supseteq \mathscr{M}_n$이 되어 증명이 끝난다. □

실로 놀라운 결과가 아닐 수 없다! 우리는 $\mathscr{M}_n$의 구조를 전혀 모르는 상태에서 단지 이가 너무 크다는 생각에 이 중에서 쓸만하다고 생각되는 극히 일부의 집합만 뽑아내어 $\mathscr{B}_n$을 구성했는데, $\mathscr{B}_n$은 정말로 $\mathscr{M}_n$에서 중요한 원소들로만 이루어진 알짜배기 σ−대수였던 것이다.

1.6 Measurable Functions

이제 Lebesgue 측도를 다른 방법으로 다시 구성하는 것으로 본 장의 후반부를 시작한다. Lebesgue 측도의 두 번째 구성 또한 기본적으로는 확장을 이용한 방법으로, 이번에는 마치 n개의 $\mathbb{R}$에 Cartesian 곱을 취하여 $\mathbb{R}^n$을 얻는 것과 비슷하게 n개의 Lebesgue 측도공간 $(\mathbb{R}, \mathscr{M}_1, \lambda_1)$에 적당한 곱연산을 취하여 $(\mathbb{R}^n, \mathscr{M}_n, \lambda_n)$을 얻어보려고 한다. 다만, 이번에도 $\mathscr{M}_n$이 너무 거대한 탓에 곱연산을 취한다고 해서 곧바로 $\mathscr{M}_n$을 얻지는 못하는데, 다행히 이는 완비화를 통해 해결할 수 있다. 이러한 방법은 Carathéodory의 확장정리를 비롯한 확장정리들을 사용한 방법보다 훨씬 더 직관적인 방법인데, 그럼에도 불구하고 이 방법을 두 번째로 소개하는 까닭은 이때 사용할 '적당한 곱연산' 을 정의하기가 꽤나 어렵기 때문이다.

마치 배보다 배꼽이 더 큰 격이다. 특히, 이 곱연산을 정의하려면 Lebesgue 적분론의 전개가 반드시 필요하므로 일단은 Lebesgue 적분론에 집중하도록 하고, 이 과정이 끝나면 다시 원래의 목표로 돌아와 Lebesgue 측도의 두 번째 구성을 마무리하도록 하겠다.

앞서 초록에서 지적한 바와 같이 Lebesgue 적분론은 우리가 측도론을 배우는 주된 이유 중 하나이다. 물론, Lebesgue 적분론에서 극한과 적분의 상호작용이 Riemann 적분론에서보다 자연스럽다는 것도 Lebesgue 적분론의 장점 중 하나이지만, Lebesgue 적분론의 가장 큰 수학적 의의는 측도가 정의되는 임의의 공간에서 적분론을 전개할 수 있도록 적분의 개념을 확장했다는 데 있다. 우리가 $\mathbb{R}^2$나 $\mathbb{R}^3$에 한정되었던 넓이나 부피의 개념을 일반적인 공간으로 일반화하여 측도의 개념을 정립하였으니 어찌보면 이는 당연한 수순이다.

Lebesgue 적분론을 정립하기 위한 준비로서, 과연 어떤 함수를 적분의 대상으로 삼아야 하는지에 대해 생각해보자. 당연히, 더 이상 함수의 정의역이 $\mathbb{R}^n$의 부분집합일 필요가 없다. 정의역에서 적당한 측도가 정의되기만 하면 충분하다. 한편, Riemann 적분론에서 properly 적분가능한 함수는 반드시 유계여야 했는데, 이 또한 왠지 불필요하게 과도한 제한이라 생각된다. 결론부터 말하자면, Lebesgue 적분론에서 적분의 대상이 되는 함수는 가측함수들이다. (가측함수라고 해서 모두 적분가능하다는 뜻이 아니다. 마치 Riemann 적분론에서 유계함수라고 해서 모두 적분가능하지는 않은 것처럼 Lebesgue 적분론에서도 적분불가능한 가측함수가 얼마든지 있다.) 이번 절에서는 이러한 가측함수에 대해 알아보도록 하겠다.

Definition 1.68 가측공간 $(X,\mathscr{A})$, $(Y,\mathscr{B})$와 함수 $f:X\to Y$를 생각하자. 만약 임의의 $A\in\mathscr{B}$에 대해 $f^{-1}(A)\in\mathscr{A}$이면 이때의 f를 **$(\mathscr{A}/\mathscr{B}-)$가측함수($(\mathscr{A}/\mathscr{B}-)$measurable function)**라 한다. 특히, 만약 $\mathscr{A}$가 Borel $\sigma-$대수이면 이때의 f를 **Borel 함수(− function)**라 한다.

Theorem 1.69 가측공간 $(X,\mathscr{A})$, $(Y,\mathscr{B})$와 함수 $f:X\to Y$에 대해 $\mathscr{C}$를 $\mathscr{B}$의 생성자라 하면 f가 가측일 필요충분조건은 임의의 $A\in\mathscr{C}$에 대해 $f^{-1}(A)\in\mathscr{A}$인 것이다.

Proof 충분조건임은 정의로부터 자명하므로 필요조건임만 보이면 된다. 이를 위해 $\mathscr{D}=\{A\subseteq Y: f^{-1}(A)\in\mathscr{A}\}$를 생각하면 가정으로부터 $\mathscr{C}\subseteq\mathscr{D}$에서 $\mathscr{D}$는 비어있지 않다. 또한, 임의의 $A\in\mathscr{D}$에 대해 $f^{-1}(A^c)=[f^{-1}(A)]^{-1}\in\mathscr{A}$에서 $A^c\in\mathscr{D}$가 되어 $\mathscr{D}$는 여집합에 대해 닫혀있다. 마지막으로 $\mathscr{D}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $f^{-1}(\bigcup_{i=1}^{\infty}A_i)=\bigcup_{i=1}^{\infty}f^{-1}(A_i)\in\mathscr{A}$에서 $\bigcup_{i=1}^{\infty}A_i\in\mathscr{D}$이므로 $\mathscr{D}$는 가산 합집합에 대해서도 닫혀있어서 곧 이는 $\sigma-$대수이고 $\mathscr{B}\subseteq\mathscr{D}$가 되어 증명이 끝난다. □

적분론에서는 이론 전개상의 편의를 위해 실수체계에 $\pm\infty$를 추가하는 것이 편하다. 이렇게 실수에 $\pm\infty$를 추가한 수 체계를 **확장된 실수(extended real number)**라 하며 $\overline{\mathbb{R}}$로 쓴다. 우리가 $\mathbb{R}$에서 성립하는 것으로 알고있는 거의 대부분의 성질들, 예컨대 LUBP, GLBP, MSP 따위의 성질들은 $\overline{\mathbb{R}}$에서도 그대로 (사실은 더 잘) 성립한다. 다만, $\overline{\mathbb{R}}$에서는 $\mathbb{R}$에서의

표준 거리함수가 더 이상 제 역할을 하지 못하므로 (당장 $d(-\infty,\infty)$를 뭐라고 해야 할지부터가 대략 난감하다.) $\overline{\mathbb{R}}$에서의 열림과 닫힘을 정의하기 위해서는 새로운 아이디어가 필요하다. 여기에서는 위상수학의 내용을 피하기 위해 결론만 받아들이도록 하자. 우리는 집합 $U \subseteq \overline{\mathbb{R}}$가 적당한 $x, y \in \mathbb{R}$에 대해 $(x, y), (x, \infty], [-\infty, y)$ 중 하나로 표현되는 구간들의 가산 합집합으로 표현되면 이때의 U를 열린집합이라 정의한다. 이제 닫힌집합의 정의는 $\mathbb{R}$에서와 마찬가지이다. 한편, $\mathbb{R}$에서의 열린집합이 유계인 열린구간의 가산 합집합으로 표현된다는 점을 상기하면(각주 ? ? 참조.), 이는 $\overline{\mathbb{R}}$에서도 여전히 열려있음을 알 수 있고, 곧 $\mathscr{U}, \mathscr{U}'$를 각각 $\mathbb{R}$과 $\overline{\mathbb{R}}$에서의 모든 열린집합의 모임이라 하면 $\mathscr{U} \subseteq \mathscr{U}'$이 성립한다.[8]

이제 $\mathscr{B}_1$을 정의한 것과 비슷하게 $\overline{\mathbb{R}}$ 위의 σ-대수 $\mathscr{B}_1'$을 $\mathscr{U}'$이 생성하는 σ-대수로 정의하자.[9]그렇다면 $\mathscr{U} \subseteq \mathscr{U}'$에서 $\mathscr{B}_1 \subseteq \mathscr{B}_1'$이 성립한다. 사실, 대부분의 경우 WLOG, 실함수의 공역을 $\overline{\mathbb{R}}$라 해도 되므로 $\mathscr{B}_1$과 $\mathscr{B}_1'$을 엄밀하게 구분해야 할 필요가 그렇게 많지 않고, 따라서 표기를 남용하여 $\mathscr{B}_1'$을 그냥 $\mathscr{B}_1$로 쓴다. 나아가 정리 ? ? , ? ? , ? ? 의 증명을 비슷하게 반복하면 $\overline{\mathbb{R}}$ 위의 $\mathscr{B}_1$에 대해 다음 결과를 얻는다.

Theorem ? ? ° 집합족 $\mathscr{C}$를 어떤 $x \in \mathbb{R}$에 대해 $[-\infty, x]$로 표현되는 모든 구간의 모임이라 하면 이는 $\mathscr{B}_1$의 생성자이다.

한편, 앞서 가측함수를 굉장히 일반적으로 정의했지만, 적분론의 전개를 목표로 하는 우리는 함수의 공역이 $\overline{\mathbb{R}}$이거나 $\mathbb{R}^n$인 두 가지 경우에 특히 관심이 있다. 그리고 예외적으로, $\overline{\mathbb{R}}$이나 $\mathbb{R}^n$이 함수의 공역을 쓰인 경우, 여기에는 Lebesgue 측도가 아닌 Borel 측도가 주어진 것으로 가정한다. 따라서 다음 보조정리가 자주 쓰인다.

Corollary 1.70 가측공간 $(X, \mathscr{A})$에서 정의된 함수 $f : X \to \mathbb{R}^n$(혹은 $f : X \to \overline{\mathbb{R}}$)에 대해 f가 가측일 필요충분조건은 임의의 $x \in \mathbb{R}^n$(혹은 $x \in \mathbb{R}$)에 대해 $f^{-1}(\prod_{i=1}^n(-\infty, x_i]) \in \mathscr{A}$(혹은 $f^{-1}([-\infty, x]) \in \mathscr{A}$)인 것이다.

Proof 이는 정리 ? ? °, ? ? 로부터 자명하다. □

정의만 가지고는 어떤 함수가 가측인지의 여부를 직접 판단하는 것은 쉬운 일이 아니므로 이에 도움이 될 만한 정리들을 다수 소개한다.

Proposition 1.71 가측공간 $(X, \mathscr{A})$에 대해 집합 $A \subseteq X$가 가측일 필요충분조건은 $\mathbf{1}_A$가 가측인 것이다.

Proof 정의로부터 $\mathbf{1}_A$가 가측인 것과 임의의 $x \in \mathbb{R}$에 대해 집합

$$\mathbf{1}_A^{-1}([-\infty, x]) = \begin{cases} X & x \geq 1\text{인 경우} \\ A^c & 0 \leq x < 1\text{인 경우} \\ \emptyset & \text{ow.} \end{cases}$$

이 가측인 것은 동치이고, 이는 다시 A가 가측이라는 것과 동치이므로 이 명제는 자명하다. □

Theorem 1.72 가측공간 $(X, \mathscr{A})$에서 정의된 가측함수 $f : X \to \mathbb{R}^m$와 Borel 함수 $g : \mathbb{R}^m \to \mathbb{R}^n$(혹은 $g : \mathbb{R}^m \to \overline{\mathbb{R}}$)에 대해 합성 $g \circ f$는 가측이다.

Proof 임의의 $x \in \mathbb{R}^n$를 택하면 가정으로부터 $g^{-1}((-\infty, x])$가 Borel이고, 따라서 $(g \circ f)^{-1}((-\infty, x]) = f^{-1}(g^{-1}((-\infty, x]))$가 가측이 되어 $g \circ f$가 가측임을 안다.

한편, $g : \mathbb{R}^m \to \overline{\mathbb{R}}$의 경우에도 이와 비슷하게 보일 수 있다. □

위의 정리에서 함수 g가 Borel이라는 조건은 필수적이다. 만약 g가 단지 가측함수에 그친다면, f, g는 모두 가측이지만 그 합성이 가측이 아닌 반례가 존재한다. [10]

Theorem 1.73 모든 연속함수 $f : \mathbb{R}^m \to \mathbb{R}^n$는 Borel이다.

Proof 임의의 $x \in \mathbb{R}^n$에 대해 $(-\infty, x]$가 닫혀있으므로 그 역상 $f^{-1}((-\infty, x])$도 닫혀있고, 곧 Borel이므로 f는 Borel이다. □

Theorem 1.74 가측공간 $(X, \mathscr{A})$에서 정의된 함수 $f : X \to \mathbb{R}^n$에 대해 f가 가측일 필요충분조건은 f의 각 성분 $f_1, \cdots, f_n : X \to \mathbb{R}$이 가측인 것이다.

Proof 충분조건임을 보이기 위해 임의의 $x \in \mathbb{R}$를 택하면 각 $i \le n$에 대해 $\mathbb{R}^{i-1} \times (-\infty, x] \times \mathbb{R}^{n-i}$이 닫혀있으므로 곧 Borel이 되어 가정으로부터 $f_i^{-1}((-\infty, x]) = f^{-1}(\mathbb{R}^{i-1} \times (-\infty, x] \times \mathbb{R}^{n-i})$가 가측이고, 따라서 f_i가 가측이 되어 충분조건임이 보여진다. 반대로, 필요조건임을 보이기 위해 임의의 $x \in \mathbb{R}^n$를 택하면 가정으로부터 모든 $i \le n$에 대해 $f_i^{-1}((-\infty, x_i])$가 가측이므로 $f^{-1}((-\infty, x]) = \bigcap_{i=1}^n f_i^{-1}((-\infty, x_i])$가 가측이고, 따라서 f가 가측이 되어 필요조건임이 보여진다. □

Theorem 1.75 가측공간 $(X, \mathscr{A})$에서 정의된 가측함수 $f, g : X \to \mathbb{R}^n$(혹은 $f, g : X \to \overline{\mathbb{R}}$)와 가측함수 $h : X \to \mathbb{R}$(혹은 $h : X \to \overline{\mathbb{R}}$)에 대해 $f \pm g, hf, f/h$는 모두 가측이다. 단, 연산의 과정에서 0으로 나누게 되거나 $\infty - \infty, \infty/\infty, 0 \cdot \infty$와 같은 부정형이 등장하여 그 결과가 well-define되지 않는 경우 이를 $c \in \mathbb{R}$와 같은 dummy value로 두는 관례를 전제하자.

Proof 각 $i \le n$에 대해 정리 ? ? 로부터 f_i, g_i는 가측이고, 함수 $+, -, \cdot : \mathbb{R}^2 \to \mathbb{R}$가 모두 연속함수이므로 정리 ? ? 로부터 이들은 Borel이다. 그렇다면 정리 ? ? 로부터 $f_i \pm g_i, hf_i$가 모두 가측이며, 다시 정리 ? ? 에서 $f \pm g, hf$가 가측임을 안다. 마지막으로, f/h가 가측임을 보이기 위해서는 $1/h$가 가측임을 보이는 것으로 충분한데, $h^{-1}(0) = \emptyset$인 경우에는

$$(1/h)^{-1}((-\infty, x]) = \begin{cases} h^{-1}((-\infty, 0)) \cup h^{-1}([1/x, \infty)) & x > 0\text{인 경우} \\ h^{-1}((-\infty, 0)) & x = 0\text{인 경우} \\ h^{-1}([1/x, 0)) & \text{ow.} \end{cases}$$

이므로 $1/h$는 가측이다. 이제 일반적인 h에 대해서도 $A = h^{-1}(0)$에 대해 이가 닫힌집합이므로 곧 Borel이고 $1/h = (1/h)\mathbf{1}_{X\setminus A} + c\mathbf{1}_A$에서 $1/h$가 가측이 되어 증명이 끝난다. 한편, $f, g, h : X \to \overline{\mathbb{R}}$인 경우에도 비슷하게 하면 된다. □

Theorem 1.76 가측공간 $(X, \mathscr{A})$에서 정의된 가측함수 $f_i : X \to \overline{\mathbb{R}}$의 열 $\{f_i\}$에 대해 다음이 성립한다.

i. 함수 $\sup_{i\in\mathbb{N}} f_i, \inf_{i\in\mathbb{N}} f_i, \limsup_{i\to\infty} f_i, \liminf_{i\to\infty} f_i$는 모두 가측이다.
ii. 만약 모든 $x \in X$에 대해 $\lim_{i\to\infty} f_i(x)$가 존재하면 이는 가측이다.
iii. 집합 $\{x \in X : \{f_i(x)\}$가 수렴$\}$이 가측이다.
iv. 임의의 가측인 $f : X \to \overline{\mathbb{R}}$에 대해 집합 $\{x \in X : f_i(x) \to f(x)\}$가 가측이다.

한편, 가측함수 $f_i : X \to \mathbb{R}^n$의 열 $\{f_i\}$에 대해서도 i을 제외한 나머지 성질들이 성립한다. 다만, iv의 경우 가측함수 $f : X \to \mathbb{R}^n$에 대해 성립한다

Proof i. 임의의 $x \in \mathbb{R}$에 대해 $\{y \in X : \sup_{i\in\mathbb{N}} f_i(y) \le x\} = \bigcap_{i=1}^{\infty} f_i^{-1}([-\infty, x])$가 가측이므로 곧 $\sup_{i\in\mathbb{N}} f_i$가 가측이고, 비슷하게 $\inf_{i\in\mathbb{N}} f_i$도 가측임을 보일 수 있다. 한편, $\limsup_{i\to\infty} f_i = \inf_{j\in\mathbb{N}} \sup_{i\ge j} f_i$이고 $\liminf_{i\to\infty} f_i = \sup_{j\in\mathbb{N}} \inf_{i\ge j} f_i$이므로 앞선 결과로부터 이들 또한 가측이다.

ii. 이는 $\lim_{i\to\infty} f_i = \limsup_{i\to\infty} f_i = \liminf_{i\to\infty} f_i$에서 자명하다.

iii. 정리에서 주어진 집합이 곧 $\{x \in X : \limsup_{i\to\infty} f_i(x) = \liminf_{i\to\infty} f_i(x)\}$이므로 자명하다.

iv. 정리에서 주어진 집합이 곧 $\{x \in X : \limsup_{i\to\infty} f_i(x) = \liminf_{i\to\infty} f_i(x) = f(x)\}$이므로 자명하다.

한편, 가측함수 $f_i : X \to \mathbb{R}^n$의 열 $\{f_i\}$와 가측함수 $f : X \to \mathbb{R}^n$의 경우에는 앞선 결과를 성분별로 적용하면 된다. □

이전 절에서 공집합과 영집합 간의 차이가 흐려져 마치 공집합과 영집합을 같은 것처럼 생각하는 경우가 많다고 하였는데, 이를 조금 더 명확히 하기 위해 '거의 어디서나' 라는 개념을 소개한다. 이 뭔가 굉장히 수학스럽지 않은 개념은 측도론의 발전이 낳은 사고방식 중 하나이다.

Definition 1.77 측도공간 $(X, \mathscr{A}, \mu)$에서 적당한 영집합 $N \subseteq X$이 존재하여 어떤 성질이 모든 $x \in X \setminus N$에 대해 성립한다면 그 성질은 X의 **(μ−)거의 어디서나((**μ**−)almost everywhere)** 혹은 **(μ−)거의 대부분의** $x \in X$**에 대해((**μ**−)almost every** $x \in X$**)** 성립한다고 한다.

어떤 성질 P가 거의 어디서나 성립한다는 것을 때로는 P (ae.)라 줄여 간단히 표기할 때도 있다. 한편, 거의 어디서나 성립하는 성질 중 가운데 특히 자주 쓰이는 것이 바로 **거의 어디서나 같다(almost everywhere equality)**는 개념이다.

Proposition 1.78 거의 어디서나 같다는 관계는 동치관계이다. 즉, 측도공간 $(X, \mathscr{A}, \mu)$와 집합 Y, 함수 $f, g, h : X \to Y$에 대해 다음이 성립한다.

i. (반사성) $f = f$ (ae.)
ii. (대칭성) $f = g$ (ae.)이면 $g = f$ (ae.)이다.
iii. (전이성) $f = g$ (ae.)이고 $g = h$ (ae.)이면 $f = h$ (ae.)이다.

Proof 반사성과 대칭성은 자명하므로 전이성만 보이면 되는데, $\{x \in X : f(x) \neq h(x)\} \subseteq \{x \in X : f(x) \neq g(x)\} \cup \{x \in X : g(x) \neq h(x)\}$에서 우변의 두 집합이 모두 영집합이므로 좌변의 집합도 영집합이 되어 전이성이 성립함을 알고, 증명이 끝난다. □

Theorem 1.79 완비측도공간 $(X, \mathscr{A}, \mu)$, 가측공간 $(Y, \mathscr{B})$와 함수 $f, g : X \to Y$에 대해 만약 f가 가측이고 $f = g$ (ae.)이면 g도 가측이다.

Proof 임의의 $A \in \mathscr{B}$를 택하고 $N = f^{-1}(A) \triangle g^{-1}(A)$라 하면 $N \subseteq \{x \in X : f(x) \neq g(x)\}$이므로 이는 가측이고, 따라서 $g^{-1}(A) = f^{-1}(A) \triangle N$도 가측이 되어 g가 가측임을 안다. □

Theorem 1.80 완비측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f_i : X \to \mathbb{R}^n$(혹은 $f_i : X \to \overline{\mathbb{R}}$)의 열 $\{f_i\}$, 함수 $f : X \to \mathbb{R}^n$(혹은 $f : X \to \overline{\mathbb{R}}$)에 대해 만약 $f_i \to f$ (ae.)이면 f도 가측이다.

Proof 집합 $N = \{x \in X : f_i(x) \not\to f(x)\}$를 생각하여 함수 $g_i : X \to \mathbb{R}^n$(혹은 $g_i : X \to \overline{\mathbb{R}}$)의 열 $\{g_i\}$를 $g_i = f_i \mathbf{1}_{X \setminus N}$로 두고 비슷하게 함수 $g : X \to \mathbb{R}^n$(혹은 $g : X \to \overline{\mathbb{R}}$)를 $g = f \mathbf{1}_{X \setminus N}$로 두면 모든 $i \in \mathbb{N}$에 대해 $f_i = g_i$ (ae.)이고 $f = g$ (ae.)이며 $g_i \to g$이다. 이제 정리 ? ? 로부터 $\{g_i\}$는 가측함수열이며, 정리 ? ? 의 ii로부터 g는 가측이고, 다시 정리 ? ? 로부터 f도 가측이다. □

정리 ? ? , ? ? 에서 측도공간의 완비성은 필수적인 조건이다. 완비성을 가지는 Lebesgue 측도공간과는 달리 우리가 주로 다루게 될 측도공간인 확률공간은 완비성을 가질 필요가 없으므로 이와 같이 완비성을 조건으로 하는 정리는 사용하기 전에 완비성 조건을 반드시 확인해줘야 한다. (확률공간은 다음 장에서 소개될 것이다.) 다음으로 소개할 내용은 가측함수의 근사에 관한 내용이다.

Definition 1.81 가측공간 $(X, \mathscr{A})$에서 정의된 가측함수 $f : X \to \mathbb{R}^n$에 대해 이의 치역 $f(X)$가 유한하면 이때의 f를 **단순함수(simple function)**라 한다.

임의의 단순함수 $f : X \to \mathbb{R}^n$에 대해 이가 갖는 유한개의 서로다른 함숫값을 $a_1, \cdots, a_k \in \mathbb{R}^n$라 하고 각 $i \leq k$에 대해 $A_i = f^{-1}(a_i)$라 하면 $A_1, \cdots, A_k$는 가측이고 서로소이며 $\bigsqcup_{i=1}^k A_i = X$이다. 따라서 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$와 같이 지시함수를 사용하여 이를 나타낼 수 있는데, 이러한 형태를 f의 **표준형(standard expression)**이라 한다.

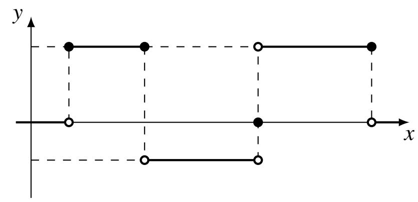


Figure 1.5 단순함수의 예시.

Theorem 1.82 가측공간 $(X, \mathscr{A})$에서 정의된 가측함수 $f : X \to \overline{\mathbb{R}}$에 대해 적당한 단순함수 $f_i : X \to \mathbb{R}$의 열 $\{f_i\}$가 존재하여 다음을 만족한다.

i. $f(x) \geq 0$인 $x \in X$에 대해서는 $0 \leq f_i(x) \uparrow f(x)$이고 $f(x) < 0$인 $x \in X$에 대해서는 $0 \geq f_i(x) \downarrow f(x)$이다.
ii. 임의의 $x \in X$와 모든 $i \in \mathbb{N}$에 대해 $|f(x) - f_i(x)| < 2^{-i}$이거나 $|f_i(x)| = i$이다.

한편, 가측함수 $f : X \to \mathbb{R}^n$에 대해서도 적당한 단순함수 $f_i : X \to \mathbb{R}^n$의 열 $\{f_i\}$가 존재하여 $f_i \to f$이다.

Proof 함수열 $\{f_i\}$를

$$f_i : x \mapsto \begin{cases} -i & f(x) \leq -i\text{인 경우} \\ -(j-1)2^{-i} & \text{적당한 } j \leq i2^i\text{에 대해 } -j2^{-i} < f(x) \leq -(j-1)2^{-i}\text{인 경우} \\ (j-1)2^{-i} & \text{적당한 } j \leq i2^i\text{에 대해 } (j-1)2^{-i} \leq f(x) < j2^{-i}\text{인 경우} \\ i & f(x) \geq i\text{인 경우} \end{cases}$$

로 정의하면 이는 명백히 단순함수이고 정리의 조건 i, ii를 모두 만족함을 쉽게 보일 수 있다.

한편, 가측함수 $f : X \to \mathbb{R}^n$의 경우에는 각 성분이 단순함수인 함수도 단순함수라는 사실로부터 앞선 결과를 성분별로 적용하기만 하면 된다. □

비록 가측함수는 꽤나 다양한 함수를 포괄하는 개념이지만 모든 가측함수는 단순함수의 열로써 근사가 가능하다는 공통된 성질을 가지고, 이는 이후 Lebesgue 적분론의 전개에 있어 매우 중요한 역할을 하게 될 것이다. 나중에 사용할 가측함수에 관한 몇 가지 개념과 성질을 소개하는 것으로 이번 절을 마치도록 하겠다.

Definition 1.83 공집합이 아닌 집합 X, 가측공간 $(Y, \mathscr{B})$에 대해 $\mathscr{F}$를 함수 $f : X \to Y$의 모임이라 할 때, 모든 $f \in \mathscr{F}$가 $\mathscr{A}/\mathscr{B}$–가측이도록 하는 가장 작은 σ–대수 $\mathscr{A}$를 **$\mathscr{F}$가 생성하는 σ–대수(σ–algebra generated by $\mathscr{F}$)**라 하고 $\sigma(\mathscr{F})$로 쓴다.

Proposition 1.84 공집합이 아닌 집합 X, 가측공간 $(Y,\mathscr{B})$에 대해 $\mathscr{F}$를 함수 $f:X\to Y$의 모임이라 하면 이가 생성하는 σ−대수 $\sigma(\mathscr{F})$는 유일하게 존재한다. 따라서, $\sigma(\mathscr{F})$는 well-define된다.

Proof $\sigma(\mathscr{F})$가 존재하기만 한다면, 이의 유일성은 그 정의로부터 자명하다. 존재성의 경우 Σ를 모든 $f\in\mathscr{F}$가 $\mathscr{A}/\mathscr{B}$−가측이도록 하는 모든 σ−대수 $\mathscr{A}$의 모임이라 두고 명제 ? ? 의 증명을 그대로 따라가면 된다. □

특히, $\mathscr{F}$가 한원소 집합인 경우에는 $\sigma(\mathscr{F})$의 원소들과 $\sigma(\mathscr{F})/\mathscr{B}_n$−가측인 함수들을 명시적으로 표현할 수 있다.

Theorem 1.85 공집합이 아닌 집합 X, 가측공간 $(Y,\mathscr{B})$와 함수 $f:X\to Y$에 대해 $\sigma(f)=\{f^{-1}(A)\subseteq X : A\in\mathscr{B}\}$이다.

Proof 집합족 $\mathscr{C}=\{f^{-1}(A)\subseteq X : A\in\mathscr{B}\}$에 대해 이가 σ−대수라면 f가 $\mathscr{C}/\mathscr{B}$−가측임이 분명하므로 $\sigma(f)\subseteq\mathscr{C}$인 한편, 정의로부터 역의 포함관계는 자명하므로 증명이 끝난다. 따라서 $\mathscr{C}$가 σ−대수임을 보이는 것으로 충분하다. 우선 $\emptyset=f^{-1}(\emptyset)\in\mathscr{C}$에서 이는 비어있지 않다. 또한, 임의의 $B\in\mathscr{C}$에 대해 적당한 $A\in\mathscr{B}$가 존재하여 $B=f^{-1}(A)$이므로 $B^c=[f^{-1}(A)]^c=f^{-1}(A^c)\in\mathscr{C}$가 되어 $\mathscr{C}$는 여집합에 대해 닫혀있다. 마지막으로 $\mathscr{C}$에 속하는 임의의 집합열 $\{B_i\}$를 생각하면 각 $i\in\mathbb{N}$에 대해 적당한 $A_i\in\mathscr{B}$가 존재하여 $B_i=f^{-1}(A_i)$이므로 $\bigcup_{i=1}^\infty B_i=\bigcup_{i=1}^\infty f^{-1}(A_i)=f^{-1}(\bigcup_{i=1}^\infty A_i)\in\mathscr{C}$가 되어 $\mathscr{C}$는 가산 합집합에 대해서도 닫혀 있고, 곧 이는 σ−대수이다. □

Theorem 1.86 공집합이 아닌 집합 X에서 정의된 $f:X\to\mathbb{R}^m$에 대해 함수 $g:X\to\mathbb{R}^n$(혹은 $g:X\to\overline{\mathbb{R}}$)가 $\sigma(f)/\mathscr{B}_n$−가측(혹은 $\sigma(f)/\mathscr{B}_1$−가측)일 필요충분조건은 적당한 Borel 함수 $h:\mathbb{R}^m\to\mathbb{R}^n$(혹은 $h:\mathbb{R}^m\to\overline{\mathbb{R}}$)가 존재하여 $g=h\circ f$인 것이다.

Proof 먼저 충분조건임을 보이기 위해 g가 단순함수인 경우를 생각해보자. 단순함수 g의 표준형을 $\sum_{i=1}^k a_i\mathbf{1}_{A_i}$라 하면 각 $i\le k$에 대해 $A_i\in\sigma(f)$이므로 정리 ? ? 로부터 적당한 $B_i\in\mathscr{B}_m$가 존재하여 $A_i=f^{-1}(B_i)$이다. 이제 $h=\sum_{i=1}^k a_i\mathbf{1}_{B_i}$라 하면 이는 명백히 Borel 함수이다. 또한, 각 $i\in\mathbb{N}$와 임의의 $x\in A_i$에 대해 $f(x)\in B_i$이고 $j\neq i$에 대해서는 $f(x)\in B_j$일 수 없으므로 $g=h\circ f$이다. (만약 $f(x)\in B_j$이면 $A_1,\cdots,A_k$가 서로소라는 표준형의 전제에 모순된다.)

이제 일반적인 $\sigma(f)/\mathscr{B}_n$−가측함수 g를 생각하면 정리 ? ? 로부터 적당한 단순함수 $g_i: X\to\mathbb{R}^n$의 열 $\{g_i\}$가 존재하여 $g_i\to g$이고 각 $i\in\mathbb{N}$에 대해 g_i는 $\sigma(f)/\mathscr{B}_n$−가측이므로 앞선 결과로부터 적당한 Borel 함수 $h_i:\mathbb{R}^m\to\mathbb{R}^n$가 존재하여 $g_i=h_i\circ f$이다. 이제 $A=\{x\in\mathbb{R}^m : \{h_i(x)\}_i$가 수렴$\}$라 두고 함수 h를

$$h : x \mapsto \begin{cases} \lim_{i\to\infty} h_i(x) & x \in A\text{인 경우} \\ 0 & \text{ow.} \end{cases}$$

로 정의하면 이는 Borel이고, 임의의 $x \in X$에 대해 $\lim_{i\to\infty} h_i(f(x)) = \lim_{i\to\infty} g_i(x) = g(x)$에서 $f(x) \in A$이므로 $g = h \circ f$임을 안다.

반대로, 필요조건임을 보이는 것은 쉽다. 임의의 $x \in \mathbb{R}^n$에 대해 $h^{-1}((-\infty, x])$가 Borel이므로 정리 ? ? 로부터 $g^{-1}((-\infty, x]) = f^{-1}(h^{-1}((-\infty, x])) \in \sigma(f)$가 되어 g는 $\sigma(f)/\mathscr{B}_n$–가측이다.

한편, 함수 $g : X \to \overline{\mathbb{R}}$의 경우에도 이와 비슷하게 보일 수 있다. □

다음은 측도공간의 완비화와 관련이 있다. 결론부터 말하자면, 측도공간의 완비화는 함수의 가측성을 보존하며, 완비화된 공간에서 가측인 함수는 원래 공간에서의 가측함수에 의해 근사될 수 있다는 것이 다음 정리의 함의이다.

Theorem 1.87 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 함수 $f : X \to \mathbb{R}^n$ (혹은 $f : X \to \overline{\mathbb{R}}$)에 대해 다음이 성립한다.

i. 만약 f가 $\mathscr{A}/\mathscr{B}_n$–가측(혹은 $\mathscr{A}/\mathscr{B}_1$–가측)이면 이는 $\overline{\mathscr{A}}/\mathscr{B}_n$–가측(혹은 $\overline{\mathscr{A}}/\mathscr{B}_1$–가측)이다.
ii. 만약 f가 $\overline{\mathscr{A}}/\mathscr{B}_n$–가측(혹은 $\overline{\mathscr{A}}/\mathscr{B}_1$–가측)이면 적당한 $\mathscr{A}/\mathscr{B}_n$–가측함수 $g : X \to \mathbb{R}^n$ (혹은 $\mathscr{A}/\mathscr{B}_1$–가측함수 $g : X \to \overline{\mathbb{R}}$)가 존재하여 $f = g$ (ae.)이다.

Proof i. 이는 $\mathscr{A} \subseteq \overline{\mathscr{A}}$에서 자명하다.

ii. 먼저 f가 단순함수인 경우를 생각해보자. 이제 f의 표준형을 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하면 가정으로부터 각 $i \le k$에 대해 적당한 $B_i \in \mathscr{A}$와 영집합 $N_i \subseteq X$가 존재하여 $A_i = B_i \cup N_i$이다. 그렇다면 함수 g를 $g = \sum_{i=1}^k a_i \mathbf{1}_{B_i}$라 하면 이는 명백히 $\mathscr{A}/\mathscr{B}_n$–가측이고, $\{x \in X : f(x) \neq g(x)\} = \bigcup_{i=1}^k (A_i \setminus B_i) \subseteq \bigcup_{i=1}^k N_i$에서 $\bigcup_{i=1}^k N_i$도 영집합이므로 $f = g$ (ae.)이다.

한편, 일반적인 $\overline{\mathscr{A}}/\mathscr{B}_n$–가측함수 f를 생각하면 정리 ? ? 로부터 적당한 단순함수 $f_i : X \to \mathbb{R}^n$의 열 $\{f_i\}$가 존재하여 $f_i \to f$이고, 앞선 결과로부터 각 $i \in \mathbb{N}$에 대해 적당한 $\mathscr{A}/\mathscr{B}_n$–가측함수 $g_i : X \to \mathbb{R}^n$가 존재하여 $f_i = g_i$ (ae.)이다. 이제 각 $i \in \mathbb{N}$에 대해 $N_i = \{x \in X : f_i(x) \neq g_i(x)\}$라 하면 이는 영집합이므로 $N = \bigcup_{i=1}^\infty N_i$도 영집합이어서 적당한 $C \in \mathscr{A}$에 대해 $N \subseteq C$이고 $\mu(C) = 0$이다. 이제 임의의 $x \in X \setminus C$에 대해 $g_i(x) \to f(x)$이므로 함수 $g : X \to \mathbb{R}^n$를

$$g : x \mapsto \begin{cases} \lim_{i\to\infty} g_i(x) & x \in X \setminus C\text{인 경우} \\ 0 & \text{ow.} \end{cases}$$

로 두면 이는 $\mathscr{A}/\mathscr{B}_n$–가측이고, $f = g$ (ae.)임을 안다.

한편, 함수 $f : X \to \overline{\mathbb{R}}$의 경우에도 이와 비슷하게 보일 수 있다. □

위의 정리의 ii에서 임의의 μ−영집합은 $\overline{\mu}$−영집합이고 그 역도 성립하므로 일부러 어떤 측도를 기준으로 거의 어디서나인지를 나타내지 않았다.

1.7 Integration

이번 절의 전반부에서 우리는 Lebesgue 적분을 총 세 단계에 걸쳐 정의한다. 먼저 첫 번째 단계에서는 음이 아닌 단순함수들에 대해 Lebesgue 적분을 정의한다. 이어지는 두 번째 단계에서는 정리 ? ? 를 사용하여 Lebesgue 적분의 정의를 음이 아닌 가측함수로 확장한다. 마지막으로 세 번째 단계에서는 함수를 양의 부분과 음의 부분으로 나눔으로써 Lebesgue 적분의 정의를 일반적인 가측함수로 확장하여 Lebesgue 적분의 가장 일반적인 정의를 얻는다. 이렇게 Lebesgue 적분의 정의를 끝마친 이후에는 앞서 언급한 바와 같이 Lebesgue 적분론에서 적분과 극한의 순서를 바꾸는 데 사용되는 유용한 정리들을 살펴보도록 하겠다.

Definition 1.88 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 단순함수 $f : X \to \mathbb{R}_0^+$에 대해 이의 표준형을 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하면 이때 μ에 대한 f의 **(Lebesgue) 적분(− integral)**을 $\int_X f(x)\, d\mu(x)$ 혹은 간단히 $\int_X f\, d\mu$로 쓰고 $\int_X f\, d\mu = \sum_{i=1}^k a_i \mu(A_i)$로 정의한다.

Proposition 1.89 측도공간 $(X, \mathscr{A}, \mu)$에 대해 $A_1, \cdots, A_k \in \mathscr{A}$를 $X = \bigsqcup_{i=1}^k A_i$인 서로소인 집합이라 하자. 이제 $a_1, \cdots, a_k \geq 0$에 대해 함수 $f : X \to \mathbb{R}_0^+$를 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하면 $\int_X f\, d\mu = \sum_{i=1}^k a_i \mu(A_i)$이다.

Proof 양수 $a_1, \cdots, a_k$ 중에서 서로 같은 값들을 하나만 남기고 제거하여 이를 $b_1, \cdots, b_l \in \mathbb{R}$라 하고 각 $j \leq l$에 대해 $B_j = f^{-1}(b_j) = \bigsqcup_{i: a_i = b_j} A_i$라 하면 이는 가측이고 $f = \sum_{j=1}^l b_j \mathbf{1}_{B_j}$이다. 그렇다면 이는 곧 단순함수 f의 표준형이므로 Lebesgue 적분의 정의로부터 $\int_X f\, d\mu = \sum_{j=1}^l b_j \mu(B_j) = \sum_{j=1}^l b_j \mu(\bigsqcup_{i: a_i = b_j} A_i) = \sum_{j=1}^l b_j \sum_{i: a_i = b_j} \mu(A_i) = \sum_{i=1}^k a_i \mu(A_i)$이다. □

위의 명제에서는 $a_1, \cdots, a_k$가 서로 다르다는 조건이 없으므로 이는 음의 아닌 단순함수에 대한 적분의 정의보다 조금 더 일반적이다. 다음으로 Lebesgue 적분의 기초적인 성질들을 보자.

Proposition 1.90 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 단순함수 $f, g : X \to \mathbb{R}_0^+$에 대해 다음이 성립한다.

i. 임의의 $c \geq 0$에 대해 $\int_X cf\, d\mu = c \int_X f\, d\mu$이다.
ii. $\int_X f + g\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu$.

iii. 만약 $f \le g$이면 $\int_X f\,d\mu \le \int_X g\,d\mu$이다.

Proof 단순함수 f, g의 표준형을 각각 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$, $g = \sum_{j=1}^l b_j \mathbf{1}_{B_j}$라 하자.

i. 함수 $cf = \sum_{i=1}^k ca_i \mathbf{1}_{A_i}$가 음이 아닌 단순함수임이 분명하므로 명제 ? ? 로부터 $\int_X cf\,d\mu = \sum_{i=1}^k ca_i \mu(A_i) = c\sum_{i=1}^k a_i \mu(A_i) = c\int_X f\,d\mu$이다.

ii. 함수 $f+g = \sum_{i=1}^k \sum_{j=1}^l (a_i+b_j)\mathbf{1}_{A_i \cap B_j}$가 음이 아닌 단순함수임이 분명하고 $\bigsqcup_{i=1}^k \bigsqcup_{j=1}^l (A_i \cap B_j) = X$이므로 명제 ? ? 로부터 $\int_X f+g\,d\mu = \sum_{i=1}^k \sum_{j=1}^l (a_i+b_j)\mu(A_i \cap B_j)$이다. 한편, $f = \sum_{i=1}^k \sum_{j=1}^l a_i \mathbf{1}_{A_i \cap B_j}$, $g = \sum_{i=1}^k \sum_{j=1}^l b_j \mathbf{1}_{A_i \cap B_j}$로 쓸 수 있으므로 다시 명제 ? ? 로부터 $\int_X f\,d\mu = \sum_{i=1}^k \sum_{j=1}^l a_i \mu(A_i \cap B_j)$이고 $\int_X g\,d\mu = \sum_{i=1}^k \sum_{j=1}^l b_j \mu(A_i \cap B_j)$가 되어 명제가 성립한다.

iii. 함수 $g-f$가 음이 아닌 단순함수이므로 ii로부터 $\int_X g\,d\mu = \int_X g-f\,d\mu + \int_X f\,d\mu \ge \int_X f\,d\mu$이다. □

Corollary 1.91 측도공간 $(X, \mathscr{A}, \mu)$와 $A_1, \cdots, A_k \in \mathscr{A}$, $a_1, \cdots, a_k \ge 0$에 대해 함수 $f : X \to \mathbb{R}_0^+$를 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하면 $\int_X f\,d\mu = \sum_{i=1}^k a_i \mu(A_i)$이다.

Proof 이는 $\int_X f\,d\mu = \sum_{i=1}^k a_i \int_X \mathbf{1}_{A_i}\,d\mu = \sum_{i=1}^k a_i \mu(A_i)$에서 자명하다. □

위의 따름정리에서는 $A_1, \cdots, A_k$가 서로소일 필요도, X의 덮개일 필요도 없다. 따라서 이는 명제 ? ? 의 결과에서 한 발짝 더 나아간 결과이다. 이제 두 번째 단계로 넘어가자.

Definition 1.92 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 μ에 대한 f의 **(Lebesgue) 적분(− integral)**을 $\int_X f(x)\,d\mu(x)$ 혹은 간단히 $\int_X f\,d\mu$로 쓰고

$$\int_X f\,d\mu = \sup\left\{ \int_X g\,d\mu \in \overline{\mathbb{R}} : \text{함수 } g : X \to \mathbb{R}_0^+\text{는 } g \le f\text{인 음이 아닌 단순함수} \right\}$$

로 정의한다.

정리 ? ? 로부터 임의의 음이 아닌 가측함수는 이로 점별수렴하는 음이 아닌 함수열을 가지므로 위의 정의는 well-define되어 있다. 하지만, 이런 정의를 직접 사용하는 것은 그다지 편리하지 않으므로 이를 대신할 정리를 소개한다.

Lemma 1.93 측도공간 $(X, \mathscr{A}, \mu)$와 $A \in \mathscr{A}$에 대해 음이 아닌 단순함수 $f_i : X \to \mathbb{R}_0^+$의 열 $\{f_i\}$가 증가하고 어떤 $M \ge 0$과 모든 $x \in A$에 대해 $\lim_{i\to\infty} f_i(x) \ge M$을 만족한다고 하면 $\lim_{i\to\infty} \int_X f_i \mathbf{1}_A\,d\mu \ge M\mu(A)$이다.

Proof 만약 $M = 0$이거나 $\mu(A) = 0$이면 보조정리가 자명하므로 $M, \mu(A) > 0$라 하자. 또한, $\{\int_X f_i \mathbf{1}_A\,d\mu\}_i$가 증가하므로 만약 이가 위로 유계가 아니라면 이번에도 보조정리가 자명하여 $\{\int_X f_i \mathbf{1}_A\,d\mu\}$가 위로 유계라 하자. 그렇다면 MSP로부터 이는 수렴한다. 이제 임의

의 양수 $a < M$를 택하고 집합열 $\{A_i\}$를 $A_i = f_i^{-1}([a,\infty)) \cap A$로 두면 이는 가측이고 $A_i \uparrow A$이다. 따라서 정리 ? ? 의 i로부터 $\mu(A_i) \uparrow \mu(A)$이고, 곧 임의의 $b < \mu(A)$를 택하면 적당한 $i_0 \in \mathbb{N}$가 존재하여 임의의 $i \geq i_0$에 대해 $\mu(A_i) > b$이다. 이상의 결과로부터 임의의 $i \geq i_0$에 대해 $\int_X f_i \mathbf{1}_A \, d\mu \geq \int_X f_i \mathbf{1}_{A_i} \, d\mu \geq \int_X a \mathbf{1}_{A_i} \, d\mu = a\mu(A_i) > ab$가 성립하여 $\lim_{i\to\infty} \int_X f_i \mathbf{1}_A \, d\mu \geq M\mu(A)$임을 안다. □

Lemma 1.94 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 단순함수 $f : X \to \mathbb{R}_0^+$에 대해 음이 아닌 단순함수 $g_i : X \to \mathbb{R}_0^+$의 열 $\{g_i\}$가 증가하고 $\lim_{i\to\infty} g_i \geq f$를 만족한다면 $\lim_{i\to\infty} \int_X g_i \, d\mu \geq \int_X f \, d\mu$이다.

Proof 단순함수 f의 표준형을 $f = \sum_{j=1}^{k} a_j \mathbf{1}_{A_j}$라 하면 임의의 $j \leq k$와 임의의 $x \in A_j$에 대해 $\lim_{i\to\infty} g_i(x) \geq a_j$이므로 위의 보조정리로부터 $\lim_{i\to\infty} \int_X g_i \mathbf{1}_{A_j} \, d\mu \geq a_j \mu(A_j)$이다. 한편, 각 $i \in \mathbb{N}$에 대해 $\int_X g_i \, d\mu = \int_X g_i \sum_{j=1}^{k} \mathbf{1}_{A_j} \, d\mu = \sum_{j=1}^{k} \int_X g_i \mathbf{1}_{A_j} \, d\mu$이므로 $\lim_{i\to\infty} \int_X g_i \, d\mu = \sum_{j=1}^{k} \lim_{i\to\infty} \int_X g_i \mathbf{1}_{A_j} \, d\mu \geq \sum_{j=1}^{k} a_j \mu(A_j) = \int_X f \, d\mu$를 얻는다. □

Theorem 1.95 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 $f_i \uparrow f$인 임의의 음이 아닌 단순함수 $f_i : X \to \mathbb{R}_0^+$의 열 $\{f_i\}$를 생각하면 $\int_X f_i \, d\mu \uparrow \int_X f \, d\mu$이다.

Proof 수열 $\{\int_X f_i \, d\mu\}_i$가 증가하므로 이는 수렴하거나 ∞로 발산하는데, 후자의 경우 $\int_X f \, d\mu = \infty$가 되어 정리가 자명하므로 $\{\int_X f_i \, d\mu\}$가 수렴한다고 가정하자. 그렇다면 $\{\int_X f_i \, d\mu\}$는 $\int_X f \, d\mu$에 의해 위로 유계이므로 $\lim_{i\to\infty} \int_X f_i \, d\mu \leq \int_X f \, d\mu$이다. 한편, $g \leq f$인 임의의 음이 아닌 단순함수 $g : X \to \mathbb{R}_0^+$에 대해 $f_i \uparrow f \geq g$이므로 위의 보조정리로부터 $\lim_{i\to\infty} \int_X f_i \, d\mu \geq \int_X g \, d\mu$이고, 곧 $\lim_{i\to\infty} \int_X f_i \, d\mu \geq \int_X f \, d\mu$가 되어 이를 앞선 결과와 합하면 $\int_X f_i \, d\mu \uparrow \int_X f \, d\mu$를 얻는다. □

이제 마지막 단계이다.

Definition 1.96 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 $\int_X f \, d\mu < \infty$이면 이때의 f를 **(μ−Lebesgue) 적분가능(− integrable)**하다고 한다. 나아가, 일반적인 가측함수 $f : X \to \overline{\mathbb{R}}$에 대해 $f_\pm$이 모두 적분가능하면 이때의 f를 **(μ−Lebesgue) 적분가능(− integrable)**하다고 한다. 그리고 이때 μ에 대한 f의 **(Lebesgue) 적분(− integral)**을 $\int_X f(x) \, d\mu(x)$ 혹은 간단히 $\int_X f \, d\mu$로 쓰고 $\int_X f \, d\mu = \int_X f_+ \, d\mu - \int_X f_- \, d\mu$로 정의한다. 여기서 $f_\pm$은 각각 f의 양의 부분과 음의 부분을 의미하는 것으로 각각 $f_\pm = (|f| \pm f)/2$로 정의된다.

Definition 1.97 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f : X \to \overline{\mathbb{R}}$와 $A \in \mathscr{A}$에 대해 함수 $f\mathbf{1}_A$가 적분가능하면 이때의 f를 A 위에서 **(μ−Lebesgue) 적분가능(integrable)**하다

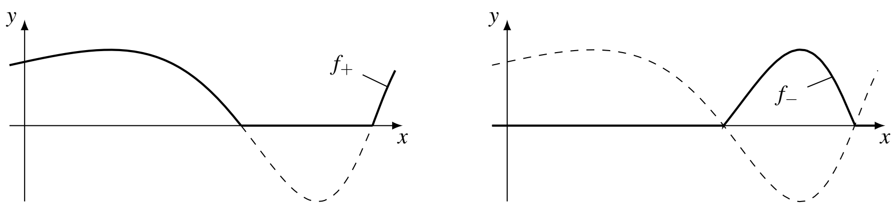


Figure 1.6 함수 f (파선)에 대한 f_+ (왼쪽)와 f_- (오른쪽)의 그래프.

고 한다. 그리고 이때 μ에 대한 f의 A**에서의 (Lebesgue) 적분(– integral over** A)을 $\int_A f(x)\,d\mu(x)$ 혹은 간단히 $\int_A f\,d\mu$로 쓰고 $\int_A f\,d\mu = \int_X f\mathbf{1}_A\,d\mu$로 정의한다.

Lebesgue 적분의 정의가 끝났으니 이제 이의 성질들을 알아보도록 하자. 표기의 편의와 간결한 논의를 위해 대부분의 명제나 정리들에서 적분 영역이 전체 공간 X로 주어져 있지만 임의의 가측집합 $A \subseteq X$에 대해 f 대신 $f\mathbf{1}_A$를 생각함으로써 이들을 임의의 가측집합을 적분 영역으로 갖는 경우로 자명하게 일반화할 수 있다.

Theorem 1.98 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f, g : X \to \overline{\mathbb{R}}_0^+$와 $A, B \in \mathscr{A}$에 대해 다음이 성립한다.

i. 임의의 $c \geq 0$에 대해 $\int_X cf\,d\mu = c\int_X f\,d\mu$이다.
ii. $\int_X f+g\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$.
iii. 만약 $f \leq g$이면 $\int_X f\,d\mu \leq \int_X g\,d\mu$이다.
iv. 만약 $A \subseteq B$이면 $\int_A f\,d\mu \leq \int_B f\,d\mu$이다.

한편, 적분가능한 함수 $f, g : X \to \overline{\mathbb{R}}$에 대해서도 iv를 제외한 나머지 성질들이 성립한다. 다만, i의 경우 더 이상 $c \geq 0$의 조건이 요구되지 않으며, cf와 $f+g$의 적분가능이 보장된다.

i°. 임의의 $c \in \overline{\mathbb{R}}$에 대해 cf가 적분가능하고 $\int_X cf\,d\mu = c\int_X f\,d\mu$이다.
ii°. 함수 $f+g$가 적분가능하고 $\int_X f+g\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$이다. [11]

나아가 이 경우, 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해서는 자명한 다음 성질도 성립한다.

v. $|\int_X f\,d\mu| \leq \int_X |f|\,d\mu$.

Proof 정리 ? ? 로부터 음이 아닌 단순함수 $f_i, g_i : X \to \mathbb{R}_0^+$의 열 $\{f_i\}, \{g_i\}$가 존재하여 $f_i \uparrow f, g_i \uparrow g$이고, 그렇다면 정리 ? ? 로부터 $\int_X f_i\,d\mu \uparrow \int_X f\,d\mu$, $\int_x g_i\,d\mu \uparrow \int_X g\,d\mu$이다.

i. 함수열 $\{cf_i\}$ 음이 아닌 단순함수열로서 $cf_i \uparrow cf$임이 분명하므로 정리 ? ? 로부터 $\int_X cf\,d\mu = \lim_{i\to\infty}\int_X cf_i\,d\mu = c\lim_{i\to\infty}\int_X f_i\,d\mu = c\int_X f\,d\mu$이다.

ii. 함수열 $\{f_i+g_i\}$가 음이 아닌 단순함수열로서 $f_i+g_i \uparrow f+g$임이 분명하므로 정리 ? ? 로부터 $\int_X f+g\,d\mu = \lim_{i\to\infty}\int_X f_i+g_i\,d\mu = \lim_{i\to\infty}\int_X f_i\,d\mu + \lim_{i\to\infty}\int_X g_i\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$이다.

iii. 상수 함수 $h=0$은 $g-f \ge h$인 음이 아닌 단순함수로 볼 수 있으므로 정의로부터 $\int_X g-f\,d\mu \ge 0$이고, 곧 ii로부터 $\int_X g\,d\mu = \int_X g-f\,d\mu + \int_X f\,d\mu \ge \int_X f\,d\mu$이다.

iv. 이는 $f\mathbf{1}_A \le f\mathbf{1}_B$와 iii으로부터 자명하다.

한편, 적분가능한 함수 $f,g : X \to \overline{\mathbb{R}}$의 경우에는 이를 각각 양과 음의 부분으로 나누어 앞선 결과를 적용하면 된다. 특히 i°의 경우 $c \ge 0$인 경우와 $c < 0$인 경우를 나누어 생각하면 된다.

v. 이는 $-|f| \le f \le |f|$와 iii으로부터 $-\int_X |f|\,d\mu \le \int_X f\,d\mu \le \int_X |f|\,d\mu$이므로 자명하다. □

Theorem 1.99 측도공간 $(X,\mathscr{A},\mu)$에서 정의된 가측함수 $f : X \to \overline{\mathbb{R}}$와 $A,B \in \mathscr{A}$에 대해 다음이 성립한다.

i. 함수 f가 적분가능할 필요충분조건은 $|f|$가 적분가능한 것이다.
ii. 만약 f가 적분가능하면 이는 거의 어디서나 유한하다.
iii. 만약 $A \subseteq B$이고 f가 B 위에서 적분가능하면 이는 A 위에서도 적분가능하다.

Proof i. 정의로부터 $|f|$가 적분가능하다는 것은 $\int_X |f|\,d\mu = \int_X f_+\,d\mu + \int_X f_-\,d\mu < \infty$인 것과 동치이고, 이는 다시 f_+, f_-가 적분가능하다는 것과 동치이므로 곧 정의로부터 f가 적분가능하다는 것과 동치이다.

ii. 우선 $\{\pm\infty\}$가 닫혀있으므로 Borel이고, 따라서 $N = f^{-1}(\pm\infty)$이 가측이다. 그런데 만약 $\mu(N) > 0$이라면 $\int_X |f|\,d\mu \ge \int_N |f|\,d\mu = \infty$에서 f가 적분가능하다는 가정에 모순되므로 N은 영집합이고, 곧 명제가 성립한다.

iii. i로부터 $\int_A |f|\,d\mu \le \int_B |f|\,d\mu < \infty$가 되어 이는 자명하다. □

Theorem 1.100 측도공간 $(X,\mathscr{A},\mu)$에서 정의된 가측함수 $f : X \to \overline{\mathbb{R}}$에 대해 다음이 성립한다.

i. 만약 $f = 0$ (ae.)이면 이는 적분가능하고 $\int_X f\,d\mu = 0$이다.
ii. 만약 f가 적분가능하고 거의 어디서나 음이 아니며 $\int_X f\,d\mu = 0$이면 $f = 0$ (ae.)이다.

Proof i. 먼저 f가 음이 아닌 경우를 생각하고 $f \ge g$인 임의의 음이 아닌 단순함수 $g : X \to \mathbb{R}_0^+$를 택하여 이의 표준형을 $g = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하자. 그렇다면 $g = 0$ (ae.)이므로 WLOG, $a_1 = 0$이라 한다면 각 $1 < i \le k$에 대해 $\mu(A_i) = 0$이 되어 $\int_X g\,d\mu = \sum_{i=1}^k a_i \mu(A_i) = 0$에서 $\int_X f\,d\mu = 0$이다. 이제 일반적인 가측함수 f에 대해서 $f_+ = f_- = 0$ (ae.)이므로 앞선 결과로부터 $\int_x f\,d\mu = \int_X f_+\,d\mu - \int_X f_-\,d\mu = 0$이다.

ii. 먼저 f가 음이 아닌 경우를 생각하고 모순을 유도하기 위해 $f = 0$ (ae.)가 아니라 가정하자. 이제 집합열 $\{A_i\}$를 $A_i := f^{-1}((1/i, \infty])$로 두면 이는 가측이고 곧 $\bigcup_{i=1}^{\infty} A_i = f^{-1}(\bigcup_{i=1}^{\infty}(1/i, \infty]) = f^{-1}((0, \infty])$도 가측이 되어 가정으로부터 양의 측도를 가진다. 나아가 $A_i \uparrow f^{-1}((0, \infty])$이므로 정리 ? ? 의 i로부터 $\mu(A_i) \uparrow \mu(f^{-1}((0, \infty])) > 0$이 되어 $\mu(A_{i_0}) > 0$인 적당한 $i_0 \in \mathbb{N}$가 존재한다. 그러나, $g = \mathbf{1}_{A_{i_0}}/i_0$라 하면 이가 $f \geq g$인 음이 아닌 단순함수이므로 $\int_X f\,d\mu \geq \int_X g\,d\mu = \mu(A_{i_0})/i_0 > 0$의 모순이 발생한다. 따라서 $f = 0$ (ae.)이다. 이제 일반적인 가측함수 f에 대해 가정으로부터 $f_- = 0$ (ae.)이므로 i에서 $\int_X f_+\,d\mu = \int_X f_-\,d\mu = 0$이고, 곧 앞선 결론으로부터 $f_+ = 0$ (ae.)가 성립하여 $f = 0$ (ae.)임을 안다. □

Corollary 1.101 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f, g : X \to \overline{\mathbb{R}}_0^+$에 대해 다음이 성립한다.

i. 만약 $N \in \mathscr{A}$이 영집합이면 f는 N 위에서 적분가능하며 $\int_N f\,d\mu = 0$이다.
ii. 만약 $f = g$ (ae.)이면 $\int_X f\,d\mu = \int_X g\,d\mu$이다.
iii. 만약 $\{x \in X : f(x) > 0\}$이 양의 측도를 가지면 $\int_X f\,d\mu > 0$이다.

한편, 적분가능한 함수 $f : X \to \overline{\mathbb{R}}$와 일반적인 가측함수 $g : X \to \overline{\mathbb{R}}$에 대해서도 iii을 제외한 나머지 성질들이 성립한다. 특히 ii의 경우 g의 적분가능성까지 보장된다.

ii° 만약 $f = g$ (ae.)이면 g도 적분가능하고 $\int_X f\,d\mu = \int_X g\,d\mu$이다.

Proof i. 이는 $f\mathbf{1}_N = 0$ (ae.)에서 정리 ? ? 의 i로부터 자명하다.

ii. 가정으로부터 적당한 영집합 $N \in \mathscr{A}$가 존재하여 $X \setminus N$에서 $f = g$이고, i로부터 $\int_X f\,d\mu = \int_{X\setminus N} f\,d\mu + \int_N f\,d\mu = \int_{X\setminus N} g\,d\mu = \int_{X\setminus N} g\,d\mu + \int_N g\,d\mu = \int_X g\,d\mu$이다.

iii. 만약 $\int_X f\,d\mu = 0$이면 정리 ? ? 의 ii로부터 $f = 0$ (ae.)의 모순이 발생하므로 명제가 자명하다.

한편, 적분가능한 함수 $f : X \to \overline{\mathbb{R}}$와 일반적인 가측함수 $g : X \to \overline{\mathbb{R}}$의 경우에는 이를 각각 양과 음의 부분으로 나누어 앞선 결과를 적용하면 된다. □

이제 이번 절의 후반부로서 Lebesgue 적분과 극한이 서로 어떻게 상호작용하는지에 대해 알아보도록 하자. 이러한 상호작용을 규정하는 정리들을 흔히 '수렴정리'라 부른다. 아쉽게도, 어느 경우에나 항상 사용할 수 있는 일반적인 수렴정리는 존재하지 않는다. 곧 각각의 수렴정리는 나름의 조건을 요구하는데, 수렴정리를 적용하기 전에 이러한 조건을 반드시 확인해야 한다.[12]첫 번째로 살펴볼 수렴정리는 함수열이 음이 아니며 증가할 것을 요구한다.

Theorem 1.102 (Monotone convergence theorem) 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f_i : X \to \overline{\mathbb{R}}_0^+$의 열 $\{f_i\}$가 증가하며 적당한 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 $f_i \uparrow f$ (ae.)라 하면 $\int_X f_i\,d\mu \uparrow \int_X f\,d\mu$이다.

Proof 먼저 $f_i \uparrow f$인 경우를 생각하면 수열 $\{\int_X f_i \, d\mu\}_i$가 증가하고 $\int_X f \, d\mu$에 의해 위로 유계이므로 $\lim_{i\to\infty} \int_X f_i \, d\mu \le \int_X f \, d\mu$이다. 따라서 역의 대소관계만 보인다면 이 경우에 대해서는 증명이 끝난다. 이를 위해 각 $i \in \mathbb{N}$에 대해 음이 아닌 단순함수 $g_{ij} : X \to \mathbb{R}_0^+$의 열 $\{g_{ij}\}_j$를 정리 ? ? 의 증명에서의 함수열과 같이 잡고 함수열 $\{h_i\}$를 $h_i := g_{ii}$로 두자. 이제 $h_i \uparrow f$임을 보이기 위해 $\{h_i\}$가 증가한다는 것과 f로 수렴한다는 것을 주장한다.

함수열 $\{h_i\}$가 증가한다는 점은 경우를 나누어 하나하나 확인해봄으로써 어렵지 않게 알 수 있다. 우선 임의의 $x \in X$와 임의의 $i \in \mathbb{N}$를 택하면 우리가 생각해야 할 경우는 다음의 네 가지이다.

i. $f_i(x) \ge i$ 이고 $f_{i+1}(x) \ge i+1$인 경우.
ii. $i \le f_i(x) \le f_{i+1}(x) < i+1$인 경우.
iii. $f_i(x) < i$ 이고 $f_{i+1}(x) \ge i+1$인 경우.
iv. $f_i(x) \le f_{i+1}(x) < i$인 경우.

연습삼아 처음 두 개의 경우에 대해서만 확인해보면, i의 경우에는 $h_{i+1}(x) = i+1 \ge i = h_i(x)$에서 자명하다. 한편, ii의 경우에는 적당한 자연수 $j \le (i+1)2^{i+1}$가 존재하여 $(j-1)2^{-i-1} \le f(x) < j2^{-i-1}$인데, 이는 $i2^{i+1} < j \le (i+1)2^{i+1}$임을 뜻하므로 곧 $h_{i+1}(x) = (j-1)2^{-i-1} \ge i = h_i(x)$에서 이 또한 자명하다. 남은 두 경우에 대해서도 이와 같이 쉽게 따져볼 수 있으므로 $\{h_i\}$는 증가함을 안다.

다음으로 $h_i \uparrow f$임을 확인해보자. 일단, 임의의 $x \in X$를 택하면 임의의 $i \in \mathbb{N}$에 대해 $|f_i(x) - h_i(x)| < 2^{-i}$이거나 $h_i(x) = i$이다. 여기서 만약 무한히 많은 $i \in \mathbb{N}$에 대해 $h_i(x) = i$라면 $f_i(x) \to \infty$에서 $f(x) = \infty$인 한편, $h_i(x) \to \infty$이므로 $h_i \uparrow f$이다. 만약 그렇지 않다면 충분히 큰 $i \in \mathbb{N}$에 대해 $|f_i(x) - h_i(x)| < 2^{-i}$이고, 곧 $h_i(x) \uparrow f(x)$임이 분명하다.

이제 $h_i \uparrow f$임을 알았으므로 정리 ? ? 로부터 $\int_X h_i \, d\mu \uparrow \int_X f \, d\mu$이고 $\{h_i\}$의 구성으로부터 모든 $i \in \mathbb{N}$에 대해 $h_i \le f_i$이므로 $\int_X f \, d\mu \le \lim_{i\to\infty} f_i \, d\mu$에서 증명이 끝난다.

한편, $f_i \uparrow f$ (ae.)인 경우에 대해서는 $N = \{x \in X : f_i(x) \not\to f(x)\}$을 생각하면 이가 영집합이고 정리 ? ? 의 iv로부터 가측이다. 이제 함수 $g_i : X \to \overline{\mathbb{R}}_0^+$의 열 $\{g_i\}$와 함수 $g : X \to \overline{\mathbb{R}}_0^+$를 각각 $g_i = f_i \mathbf{1}_{X\setminus N}, g = f\mathbf{1}_{X\setminus N}$과 같이 정의하면 $\{g_i\}$는 음이 아닌 가측함수의 열로, 증가하고 $g_i \uparrow g$이다. 그렇다면 앞서 얻은 결과로부터 $\int_X g_i \, d\mu \uparrow \int_X g \, d\mu$임을 바로 알 수 있고, 곧 모든 $i \in \mathbb{N}$에 대해 $f_i = g_i$ (ae.)이며 $f = g$ (ae.)이므로 따름정리 ? ? 의 ii로부터 $\int_X f_i \, d\mu \uparrow \int_X f \, d\mu$를 얻는다. □

두 번째 수렴정리는 함수열이 음이 아닐 것만을 요구하는 대신, 등호가 아닌 부등호의 관계만 보장한다.

Theorem 1.103 (Fatou's lemma) 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f_i : X \to \overline{\mathbb{R}}_0^+$의 열 $\{f_i\}$에 대해 $\liminf_{i\to\infty} \int_X f_i \, d\mu \ge \int_X \liminf_{i\to\infty} f_i \, d\mu$이다.

PROOF 함수열 $\{g_i\}$를 $g_i := \inf_{j\geq i} f_j$로 두면 이는 증가하는 함수열로 정리 ? ? 의 i로부터 각 g_i가 가측이므로 MCT에서 $\lim_{i\to\infty}\int_X g_i\,d\mu = \int_X \lim_{i\to\infty} g_i\,d\mu = \int_X \sup_{i\in\mathbb{N}}\inf_{j\geq i} f_j\,d\mu = \int_X \liminf_{i\to\infty} f_i\,d\mu$이다. 그렇다면 모든 $i\in\mathbb{N}$에 대해 $g_i \leq f_i$이므로 $\liminf_{i\to\infty}\int_X f_i\,d\mu \geq \liminf_{i\to\infty}\int_X g_i\,d\mu = \lim_{i\to\infty}\int_X g_i\,d\mu = \int_X \liminf_{i\to\infty} f_i\,d\mu$이다. □

마지막 수렴정리는 그 이름에서 알 수 있듯이 '지배'라는 독특한 조건을 요구한다. 즉, 함수열이 어떤 적분가능한 함수에 의해 지배되어야 한다.

Theorem 1.104 (Lebesgue's dominated convergence theorem) 측도공간 $(X,\mathscr{A},\mu)$에서 정의된 가측함수 $f_i : X\to\overline{\mathbb{R}}$의 열 $\{f_i\}$와 가측함수 $f : X\to\overline{\mathbb{R}}$에 대해 $f_i \to f$ (ae.)라 하자. 또한, 어떤 적분가능한 함수 $g : X\to\overline{\mathbb{R}}_0^+$가 존재하여 모든 $i\in\mathbb{N}$에 대해 $|f_i|\leq g$이면 각 f_i와 f는 적분가능하고 $\int_X f_i\,d\mu \to \int_X f\,d\mu$이다.

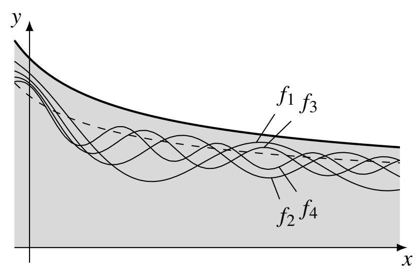


Figure 1.7 함수열 $\{f_i\}$가 g (굵은선)에 의해 지배당하고 있다. 그림에서 $\{f_i\}$는 결국 f (파선)로 점별수렴한다.

PROOF 가정으로부터 각 f_i에 대해 $\int_X |f_i|\,d\mu \leq \int_X g\,d\mu < \infty$이므로 이는 적분가능하고, 비슷한 이유로 f도 적분가능하다. 이제 $f_i\to f$인 경우를 생각해보자. 이 경우에 함수열 $\{g-f_i\},\{g+f_i\}$를 생각하면 이는 음이 아닌 가측함수열로서 Fatou의 보조정리로부터 $\liminf_{i\to\infty}\int_X g-f_i\,d\mu \geq \int_X g-f\,d\mu = \int_X g\,d\mu - \int_X f\,d\mu$이고, $\liminf_{i\to\infty}\int_X g-f_i\,d\mu = \int_X g\,d\mu - \limsup_{i\to\infty}\int_X f_i\,d\mu$에서 $\limsup_{i\to\infty}\int_X f_i\,\mu \leq \int_X f\,d\mu$를 얻는다. 한편, $\{g+f_i\}$에 대해서도 이와 비슷하게 하면 $\liminf_{i\to\infty}\int_X f_i\,d\mu \geq \int_X f\,d\mu$를 얻어 곧 $\int_X f_i\,d\mu \to \int_X f\,d\mu$가 된다.

한편, $f_i\to f$ (ae.)인 경우에 대해서는 $N=\{x\in X : f_i(x)\not\to f(x)\}$을 생각하면 이가 영집합이고 정리 ? ? 의 iv로부터 가측이다. 이제 함수 $h_i : X\to\overline{\mathbb{R}}$의 열 $\{h_i\}$와 함수 $h : X\to\overline{\mathbb{R}}$를 각각 $h_i = f_i\mathbf{1}_{X\setminus N}, h = f\mathbf{1}_{X\setminus N}$과 같이 정의하면 $\{h_i\}$는 음이 아닌 가측함수의 열로, 증가하고 모든 $i\in\mathbb{N}$에 대해 $|h_i|\leq g$이며 $h_i\uparrow h$이다. 그렇다면 앞서 얻은 결과로부터 $\int_X h_i\,d\mu \to \int_X h\,d\mu$임을 바로 알 수 있고, 곧 모든 $i\in\mathbb{N}$에 대해 $f_i = h_i$ (ae.)이며 $f=h$ (ae.)이므로 따름정리 ? ? 의 ii로부터 $\int_X f_i\,d\mu \to \int_X f\,d\mu$를 얻는다. □

이러한 수렴정리들로부터 얻어지는 다음 결과들도 유용하다.

Corollary 1.105 (Bounded convergence theorem) 유한 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f_i : X \to \overline{\mathbb{R}}$의 열 $\{f_i\}$에 대해 적당한 가측함수 $f : X \to \overline{\mathbb{R}}$가 존재하여 $f_i \uparrow f$ (ae.)라 하자. 또한, $\{f_i\}$가 균등하게 유계라 하자. 즉, 어떤 $M > 0$이 존재하여 모든 $i \in \mathbb{N}$에 대해 $|f_i| \le M$이라 하자. 그렇다면 각 f_i와 f는 적분가능하고 $\int_X f_i \, d\mu \to \int_X f \, d\mu$이다.

Proof 상수 함수 $g = M$가 명백히 적분가능하므로 DCT로부터 이는 자명하다. □

Corollary 1.106 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f_i : X \to \overline{\mathbb{R}}$의 열 $\{f_i\}$에 대해 다음이 성립한다.

i. 만약 각 f_i가 음이 아니면 $\int_X \sum_{i=1}^{\infty} f_i \, d\mu = \sum_{i=1}^{\infty} \int_X f_i \, d\mu$이다.
ii. 만약 $\sum_{i=1}^{\infty} f_i$가 거의 어디서나 수렴하거나 $\pm\infty$로 발산하며,[13]적당한 적분가능한 $g : X \to \overline{\mathbb{R}}_0^+$가 존재하여 모든 $k \in \mathbb{N}$에 대해 $|\sum_{i=1}^{k} f_i| \le g$라 하면 각 f_i와 $\sum_{i=1}^{\infty} f_i$는 적분가능하고 $\int_X \sum_{i=1}^{\infty} f_i \, d\mu = \sum_{i=1}^{\infty} \int_X f_i \, d\mu$이다.

Proof 각각 MCT와 DCT로부터 자명하다. 특히 ii의 경우 모든 $k \in \mathbb{N}$에 대해 $\sum_{i=1}^{k} f_i$가 적분가능하므로 곧 임의의 $i \in \mathbb{N}$에 대해 $f_i = \sum_{j=1}^{i} f_j - \sum_{j=1}^{i-1} f_j$가 적분가능함을 안다. □

Corollary 1.107 측도공간 $(X, \mathscr{A}, \mu)$와 열린집합 $U \subseteq \mathbb{R}^n$ 대해 함수 $f : X \times U \to \mathbb{R}$가 다음을 조건을 만족한다고 하자.

i. 임의의 $y \in U$에 대해 함수 $x \mapsto f(x, y)$가 적분가능하다.
ii. 거의 대부분의 $x \in X$에 대해 함수 $y \mapsto f(x, y)$가 연속이다.
iii. 적당한 적분가능한 함수 $g : X \to \overline{\mathbb{R}}_0^+$가 존재하여 $X \times U$에서 $|f| \le g$이다.

그렇다면 함수 $y \mapsto \int_X f(x, y) \, d\mu(x)$는 연속이다.

Proof DCT로부터 거의 자명하다. 임의의 $y \in U$를 고정하고 U에 속하는 임의의 수열 $\{y_i\}$를 생각하여 $y_i \to y$라 하면 주어진 조건으로부터 거의 대부분의 $x \in X$에 대해 $f(x, y_i) \to f(x, y)$이고 임의의 $i \in \mathbb{N}$와 임의의 $x \in X$에 대해 $|f(x, y_i)| \le g(x)$이므로, 곧 DCT에서 $\lim_{i\to\infty} \int_X f(x, y_i) \, d\mu(x) = \int_X f(x, y) \, d\mu(x)$가 되어 증명이 끝난다. □

Corollary 1.108 (Leibniz's rule)

측도공간 $(X, \mathscr{A}, \mu)$와 열린집합 $U \subseteq \mathbb{R}$ 대해 함수 $f : X \times U \to \mathbb{R}$가 다음 조건을 만족한다고 하자.

i. 임의의 $y \in U$에 대해 함수 $x \mapsto f(x, y)$가 적분가능하다.
ii. 거의 대부분의 $x \in X$에 대해 함수 $y \mapsto f(x, y)$가 미분가능하다.
iii. 임의의 $y \in U$에 대해 함수 $x \mapsto (\partial/\partial y) f(x, y)$가 가측이다.[14]
iv. 적당한 적분가능한 함수 $g : X \to \overline{\mathbb{R}}_0^+$가 존재하여 $X \times U$에서 $|\partial f/\partial y| \le g$이다.

그렇다면 함수 $y \mapsto \int_X f(x, y)\, d\mu(x)$는 미분가능하고 임의의 $y \in U$에 대해 함수 $x \mapsto (\partial f/\partial y)(x, y)$는 적분가능하며 다음이 성립한다.

$$\frac{d}{dy}\int_X f(x, y)\, d\mu(x) = \int_X \frac{\partial}{\partial y} f(x, y)\, d\mu(x)$$

Proof 편의를 위해 함수 $f_x : U \to \mathbb{R}$를 $f_x : y \mapsto f(x, y)$로 정의하자. 이제 임의의 $y \in U$를 고정하고 적당한 열린 ball B에 대해 함수 $k : X \times B \to \mathbb{R}$를

$$(x, h) \mapsto \begin{cases} \dfrac{f(x, y+h) - f(x, y)}{h} & h \neq 0 \text{ 이고 } f_x\text{가 미분가능한 경우} \\ \dfrac{\partial}{\partial y} f(x, y) & h = 0 \text{ 이고 } f_x\text{가 미분가능한 경우} \\ 0 & \text{ow.} \end{cases}$$

로 정의하면 이는 well-define되고, 만약 이가 따름정리 ? ? 의 조건을 모두 만족함을 보일 수 있으면 곧바로

$$\begin{aligned} \lim_{h\to 0}\int_X \frac{f(x, y+h) - f(x, y)}{h}\, d\mu(x) &= \lim_{h\to 0}\int_X k(x, h)\, d\mu(x) \\ &= \int_X \frac{\partial}{\partial y} f(x, y)\, d\mu(x) \end{aligned}$$

가 되어 증명이 끝난다. 그런데 주어진 조건으로부터 따름정리 ? ? 의 조건 i, ii가 만족됨은 자명하므로 iii이 만족됨을 보이는 것만 보이자. 임의의 $(x, h) \in X \times V$에 대해 만약 f_x가 미분가능하지 않다면 $k(x, h) = 0$에서 $|k(x, h)| \leq g(x)$임이 분명하므로 f_x가 미분가능하다고 하자. 그렇다면 $h = 0$인 경우에는 가정으로부터 $|k(x, h)| = |(\partial f/\partial y)(x, y)| \leq g(x)$이고, 그렇지 않은 경우에는 MVT에서 y와 $y + h$ 사이의 적당한 $z \in U$가 존재하여 $|k(x, h)| = |(\partial f/\partial y)(x, z)| \leq g(x)$가 되어 iii이 만족된다. 증명은 이로써 충분하다. □

측도를 이용해 적분을 정의한 것과 반대로 적분을 이용해 새로운 측도를 유도할 수도 있다. 이러한 기법은 확률론에서 널리 쓰이므로 여기서 잠시 살펴보는 것이 좋을 것 같다.

Theorem 1.109 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 함수 $\mu_f : \mathscr{A} \to \overline{\mathbb{R}}_0^+$를 $\mu_f : A \mapsto \int_A f\, d\mu$로 두면 μ_f는 $\mathscr{A}$ 위의 측도이다.

Proof 당장 $\mu_f(\emptyset) = 0$임은 분명하므로 μ_f가 σ–가법성을 가짐을 보이는 것으로 충분하다. 이를 위해 $\mathscr{A}$에 속하는 임의의 집합열 $\{A_i\}$를 택하면 MCT에서 $\mu_f(\bigsqcup_{i=1}^\infty A_i) = \int_{\bigsqcup_{i=1}^\infty A_i} f\, d\mu = \int_X \sum_{i=1}^\infty f \mathbf{1}_{A_i}\, d\mu = \sum_{i=1}^\infty \int_{A_i} f\, d\mu = \sum_{i=1}^\infty \mu_f(A_i)$가 되어 μ_f가 σ–가법성을 가짐을 알고, 곧 증명이 끝난다. □

Definition 1.110 가측공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 f **로부터 유도되는 측도(measure induced by** f)를 μ_f로 쓰고 $\mathscr{A}$ 위의 측도 $\mu_f : A \mapsto \int_A f\,d\mu$로 정의한다.

그렇다면 과연 이의 역도 성립할까에 대한 궁금증이 자연스럽게 피어난다. 어떤 고정된 측도 μ에 대해 임의의 측도 ν를 적당한 음이 아닌 가측함수 f를 사용하여 $\mu_f = \nu$와 같이 쓸 수 있을까? 만약 그렇다면 이러한 f는 유일할까? 만약 그 존재성이 일반적으로 보장되는 것이 아니라면, 어떤 조건에서 존재성을 말할 수 있을까? 이런 질문들에 대한 답은 자연스럽게 도함수의 일반화에 대한 아이디어로 이어지고, 이는 확률론이 꽃필 수 있는 견고한 기반이 된다. 다만, 그 자체로 꽤나 많은 논의가 필요하므로 일단 이와 관련된 논의는 잠시 미뤄두고, 여기에서는 적분에 관한 이야기를 마저 하도록 하자.

Theorem 1.111 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$와 $\mathscr{A}$에 속하는 집합열 $\{A_i\}$에 대해 다음이 성립한다.

i. 만약 $A_i \uparrow A$이면 $\int_{A_i} f\,d\mu \uparrow \int_A f\,d\mu$이다.
ii. 만약 $A_i \downarrow A$이고 f가 A_1 위에서 적분가능하면 $\int_{A_i} f\,d\mu \downarrow \int_A f\,d\mu$이다.

한편, 적분가능한 함수 $f : X \to \overline{\mathbb{R}}$에 대해서도 위의 성질들이 성립한다. 특히 ii의 경우 f의 A_1 위에서의 적분가능성이 자명해지므로 이 조건을 생략할 수 있다.

ii° 만약 $A_i \downarrow A$이면 $\int_{A_i} f\,d\mu \downarrow \int_A f\,d\mu$이다.

Proof 이는 μ_f가 $\mathscr{A}$ 위에서의 측도를 이루므로 자명하다. 한편, 적분가능한 함수 $f : X \to \overline{\mathbb{R}}$에 대해서는 이를 각각 양과 음의 부분으로 나누어 각각이 $\mathscr{A}$ 위에서의 측도를 유도한다는 사실을 이용하면 된다. □

Theorem 1.112 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f, g : X \to \overline{\mathbb{R}}_0^+$에 대해 $\int_X g\,d\mu_f = \int_X fg\,d\mu$이다. 한편, μ_f–적분가능한(혹은 fg가 μ–적분가능한 가측함수) $g : X \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며 이때 fg는 μ–적분가능하다(혹은 g는 μ_f–적분가능하다).

Proof 먼저 g가 음이 아닌 단순함수인 경우를 생각하여 이의 표준형을 $g = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하자. 그렇다면 $\int_X g\,d\mu_f = \sum_{i=1}^k a_i \mu_f(A_i) = \sum_{i=1}^k a_i \int_{A_i} f\,d\mu = \int_X \sum_{i=1}^n a_i f \mathbf{1}_{A_i}\,d\mu = \int_X fg\,d\mu$임이 분명하다. 이제 g를 음이 아닌 가측함수라 하면 정리 ? ? 로부터 음이 아닌 단순함수 $g_i : X \to \mathbb{R}_0^+$의 열 $\{g_i\}$가 존재하여 $g_i \uparrow g$이므로 MCT에서 $\int_X g d\mu_f = \lim_{i\to\infty} \int_X g_i d\mu_f = \lim_{i\to\infty} \int_X f g_i d\mu = \int_X fg d\mu$이다.

한편, μ_f–적분가능한 $g : X \to \overline{\mathbb{R}}$에 대해서, 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 fg가 μ–적분가능하며 $\int_X g\,d\mu_f = \int_X fg\,d\mu$임을 쉽게 알 수 있고, fg가 μ–적분가능한 가측함수 $g : X \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

Theorem 1.113 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f, g : X \to \overline{\mathbb{R}}_0^+$에 대해 만약 μ가 σ−유한하며 임의의 $A \in \mathscr{A}$에 대해 $\int_A f\,d\mu = \int_A g\,d\mu$이면 $f = g$ (ae.)이다. 한편, 적분가능한 $f, g : X \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며, 이 경우에는 μ가 σ−유한하다는 조건을 생략할 수 있다.

Proof 가정으로부터 $\mathscr{A}$에 속하는 적당한 집합열 $\{A_i\}$가 존재하여 $\mu(A_i) < \infty$이고 $\bigcup_{i=1}^{\infty} A_i = X$이고, WLOG, 필요하다면 각 항을 $C_i := \bigcup_{j=1}^{i} A_j$로 바꾸어 $\{A_i\}$가 처음부터 증가하는 집합열이라 해도 된다. 이제 집합열 $\{B_i\}$를 $B_i := \{x \in X : f(x) < g(x), f(x) < i\}$로 두면 이는 $\mathscr{A}$에 속하는 집합열이며 모든 $i \in \mathbb{N}$에 대해 $\int_{A_i \cap B_i} f\,d\mu = \int_{A_i \cap B_i} g\,d\mu \le \int_{A_i \cap B_i} i\,d\mu = i\mu(A_i \cap B_i) < \infty$이다. 이로부터 각 $i \in \mathbb{N}$에 대해 $\int_{A_i \cap B_i} g - f\,d\mu = 0$이고, 이는 보조정리 ? ? 의 iii으로부터 $A_i \cap B_i$가 영집합임을 의미하므로 $\{x \in X : f(x) < g(x)\} = \bigcup_{i=1}^{\infty}(A_i \cup B_i)$도 영집합이 되어 $f \ge g$ (ae.)이다. 한편, 비슷하게 하면 $f \le g$ (ae.)임도 보일 수 있으므로 곧 $f = g$ (ae.)임을 안다.

한편, 적분가능한 $f, g : X \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. 다만, 이 경우에는 f, g가 적분가능하여 $\int_{B_i} f\,d\mu = \int_{B_i} g\,d\mu < \infty$이므로 μ가 σ−유한하다는 조건이 필요하지 않다. □

측도공간의 완비화와 관련된 정리를 끝으로 길었던 Lebesgue 적분론의 전개를 마무리하도록 하겠다.

Theorem 1.114 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 $\mathscr{A}/\mathscr{B}_1$−가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 $\int_X f\,d\mu = \int_X f\,d\overline{\mu}$이다. 한편, μ−적분가능한(혹은 $\overline{\mu}$−적분가능한 $\mathscr{A}/\mathscr{B}_1$−가측함수) $f : X \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며 이때 f는 $\overline{\mu}$−적분가능하다(혹은 μ−적분가능하다).

Proof 먼저 f가 음이 아닌 단순함수인 경우를 생각하여 이의 표준형을 $f = \sum_{i=1}^{k} a_i \mathbf{1}_{A_i}$라 하자. 그렇다면 $\int_X f\,d\mu = \sum_{i=1}^{k} a_i \mu(A_i) = \sum_{i=1}^{k} a_i \overline{\mu}(A_i) = \int_X f\,d\overline{\mu}$임이 분명하다. 이제 f를 음이 아닌 가측함수라 하면 정리 ? ? 로부터 음이 아닌 단순함수 $f_i : X \to \mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i \uparrow f$이므로 MCT에서 $\int_X f\,d\mu = \lim_{i\to\infty} \int_X f_i\,d\mu = \lim_{i\to\infty} \int_X f_i\,d\overline{\mu} = \int_X f\,d\overline{\mu}$이다.

한편, μ−적분가능한 $f : X \to \overline{\mathbb{R}}$에 대해서, 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 f가 $\overline{\mu}$−적분가능하며 $\int_X f\,d\mu = \int_X f\,d\overline{\mu}$임을 쉽게 알 수 있고, $\overline{\mu}$−적분가능한 $\mathscr{A}/\mathscr{B}_1$−가측함수 $f : X \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

1.8 Product Measures and Contruction of The Lebesgue Measure II

이전 절에서 Lebesgue 적분론의 전개가 마무리 되었으므로 이제 우리의 원래 목적으로 되돌아와 Lebesgue 측도의 두 번째 구성을 본격적으로 시작해보려고 한다. 앞서 언급했듯이 이를 위해서는 우선 적당한 곱연산을 정의해야 한다.

Definition 1.115 가측공간 $(X,\mathscr{A}),(Y,\mathscr{B})$에 대해 적당한 $A\in\mathscr{A},B\in\mathscr{B}$가 존재하여 $C=A\times B$로 쓸 수 있는 집합 $C\subseteq X\times Y$를 **measurable rectangle**이라 하고, 모든 measurable rectangle의 모임을 $\mathscr{A}\times\mathscr{B}$로 쓴다. 이때 $\sigma(\mathscr{A}\times\mathscr{B})$를 $\mathscr{A},\mathscr{B}$의 **product $\sigma-$algebra**라 하며 $\mathscr{A}\otimes\mathscr{B}$로 쓴다.

위의 정의로부터 $\sigma-$대수의 $\times$곱은 결합법칙을 만족함이 분명하다. 따라서 가측공간 $(X_1,\mathscr{A}_1),\cdots,(X_i,\mathscr{A}_k)$에 대해 $\prod_{i=1}^k\mathscr{A}_i:=\mathscr{A}_1\times\cdots\times\mathscr{A}_k$가 well-define된다. 한편, $\sigma-$대수의 $\otimes$곱이 결합법칙을 만족하는가는 조금 덜 자명하다.

Theorem 1.116 가측공간 $(X,\mathscr{A}),(Y,\mathscr{B}),(Z,\mathscr{C})$에 대해 $(\mathscr{A}\otimes\mathscr{B})\otimes\mathscr{C}=\mathscr{A}\otimes(\mathscr{B}\otimes\mathscr{C})=\sigma(\mathscr{A}\times\mathscr{B}\times\mathscr{C})$이다.

Proof 표기의 편의를 위해 $\mathscr{D}=\sigma(\mathscr{A}\times\mathscr{B}\times\mathscr{C})$라 하자. 만약 $(\mathscr{A}\otimes\mathscr{B})\otimes\mathscr{C}=\mathscr{A}\otimes(\mathscr{B}\otimes\mathscr{C})=\mathscr{D}$임을 보일 수 있으면 증명이 끝난다. 이를 위해 임의의 $C\in\mathscr{C}$를 고정하고 $\mathscr{E}=\{A\in\mathscr{A}\otimes\mathscr{B}:A\times C\in\mathscr{D}\}$를 생각하면 $X\times Y\in\mathscr{E}$임이 분명하다. 또한, 임의의 $A\in\mathscr{E}$에 대해 $A^c\times C=[(X\times Y)\times C]\setminus(A\times C)\in\mathscr{D}$에서 $\mathscr{E}$는 여집합에 대해 닫혀있다. 한편, $\mathscr{E}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 $(\bigsqcup_{i=1}^{\infty}A_i)\times C=\bigsqcup_{i=1}^{\infty}(A_i\times C)\in\mathscr{D}$이므로 이상으로부터 $\mathscr{E}$이 $\lambda-$system임을 알 수 있다. 한편, 임의의 $A,B\in\mathscr{E}$에 대해 $(A\cap B)\times C=(A\times C)\cap(B\times C)\in\mathscr{E}$에서 $\mathscr{E}$는 교집합에 대해서도 닫혀있어 곧 $\pi-$system이기도 하다. 그렇다면 보조정리 ? ? 로부터 $\mathscr{E}$는 $\sigma-$대수이고, 정의로부터 $\mathscr{A}\times\mathscr{B}\subseteq\mathscr{E}$임이 분명하므로 $\mathscr{A}\otimes\mathscr{B}\subseteq\mathscr{E}$이다. 이제 이상의 결과가 임의의 $C\in\mathscr{C}$에 대해 성립함을 상기한다면 $(\mathscr{A}\otimes\mathscr{B})\times\mathscr{C}\subseteq\mathscr{D}$가 되어 곧 $(\mathscr{A}\otimes\mathscr{B})\otimes\mathscr{C}\subseteq\mathscr{D}$이다. 한편, $\mathscr{A}\times\mathscr{B}\times\mathscr{C}\subseteq(\mathscr{A}\otimes\mathscr{B})\otimes\mathscr{C}$가 분명하므로 $\mathscr{D}\subseteq(\mathscr{A}\otimes\mathscr{B})\otimes\mathscr{C}$에서 역의 포함관계도 성립하고, 곧 $(\mathscr{A}\otimes\mathscr{B})\otimes\mathscr{C}=\mathscr{D}$이다. 비슷한 방법으로 $\mathscr{A}\otimes(\mathscr{B}\otimes\mathscr{C})=\mathscr{D}$도 보일 수 있다. □

이로부터 $\sigma-$대수의 $\otimes$곱도 결합법칙을 만족하므로 $\bigotimes_{i=1}^k\mathscr{A}_i=\mathscr{A}_1\otimes\cdots\otimes\mathscr{A}_k$가 well-define된다. 만약 $\mathscr{A}_1=\cdots=\mathscr{A}_k=:\mathscr{A}$이면 이들의 $\otimes$곱을 $\mathscr{A}^{\otimes k}$로 간단히 쓰기도 한다.

Definition 1.117 집합 X,Y와 $A\subseteq X\times Y$, 점 $x_0\in X,y_0\in Y$에 대해 A의 **section** A_{x_0},A^{y_0}를 각각 $A_{x_0}=\{y\in Y:(x_0,y)\in A\},A^{y_0}=\{x\in X:(x,y_0)\in A\}$로 정의한다.

Definition 1.118 집합 X,Y,Z와 함수 $f:X\times Y\to Z$, 점 $x_0\in X,y_0\in Y$에 대해 f의 **section** $f_{x_0}:Y\to Z,f^{y_0}:X\to Z$를 각각 $f_{x_0}:y\mapsto f(x_0,y),f^{y_0}:x\mapsto f(x,y_0)$로 정의한다.

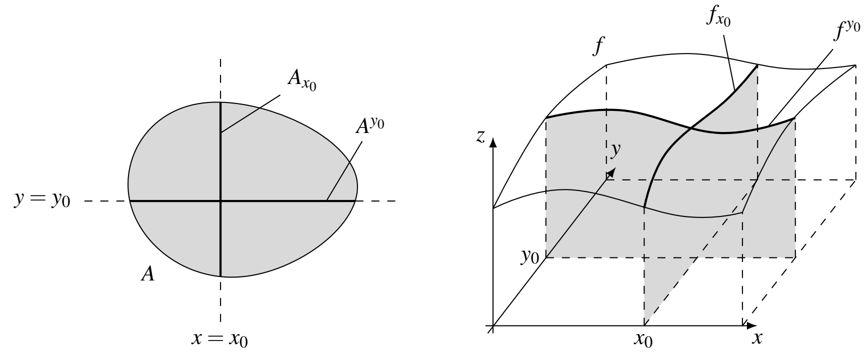


Figure 1.8 집합 $A \subseteq \mathbb{R}^2$(왼쪽)와 함수 $f : \mathbb{R}^2 \to \mathbb{R}$(오른쪽)의 section들.

다음 정리에 따르면 product σ−algebra에 속하는 measurable rectangle을 포함한 모든 집합은 항상 가측인 section을 가진다. 이로써 우리는 product σ−algebra의 구조를 조금 더 잘 이해할 수 있다.

Theorem 1.119 가측공간 $(X, \mathscr{A}), (Y, \mathscr{B}), (Z, \mathscr{C})$에 대해 다음이 성립한다.

i. 임의의 $A \in \mathscr{A} \otimes \mathscr{B}$와 임의의 $x \in X, y \in Y$에 대해 $A_x \in \mathscr{B}, A^y \in \mathscr{A}$이다.
ii. $\mathscr{A} \otimes \mathscr{B} / \mathscr{C}$−가측함수 $f : X \times Y \to Z$와 임의의 $x \in X, y \in Y$에 대해 f_x, f^y는 각각 $\mathscr{B}/\mathscr{C}$−가측이고 $\mathscr{A}/\mathscr{C}$−가측이다.

Proof i. 임의의 $x \in X$를 고정하고 $\mathscr{D} = \{A \in \mathscr{A} \otimes \mathscr{B} : A_x \in \mathscr{B}\}$를 생각하면 임의의 $A \times B \in \mathscr{A} \times \mathscr{B}$에 대해

$$(A \times B)_x = \begin{cases} B & x \in A\text{인 경우} \\ \emptyset & \text{ow.} \end{cases}$$

이므로 $\mathscr{A} \times \mathscr{B} \subseteq \mathscr{D}$이고, 따라서 $\mathscr{D}$가 σ−대수라는 것만 보이면 $\mathscr{A} \otimes \mathscr{B} \subseteq \mathscr{D}$가 되어 증명이 끝난다. 이를 위해 임의의 $A \in \mathscr{D}$를 생각하면 $(A^c)_x = (A_x)^c \in \mathscr{B}$이므로 $A^c \in \mathscr{D}$에서 $\mathscr{D}$는 여집합에 대해 닫혀있다. 또한 $\mathscr{D}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $(\bigcup_{i=1}^{\infty} A_i)_x = \bigcup_{i=1}^{\infty} (A_i)_x \in \mathscr{B}$이므로 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{D}$에서 $\mathscr{D}$는 가산 합집합에 대해서도 닫혀있다. 이제 이가 공집합이 아님은 분명하므로 곧 σ−대수이고, 증명이 끝난다. 한편, A^y에 대해서도 이와 비슷하게 하면 된다.

ii. 임의의 $A \in \mathscr{C}$에 대해 $(f_x)^{-1}(A) = [f^{-1}(A)]_x$이고 $(f^y)^{-1}(A) = [f^{-1}(A)]^y$이므로 이는 i로부터 자명하다. □

이제 product σ−algebra에 측도를 부여하는 방법에 대해 생각해보자. 만약 곱해지는 두 가측공간에 적당한 측도가 정의되어 있었다면 이를 사용하면서도 곱셈이라는 구조를 잘 반

영할 수 있도록 product $\sigma-$algebra에 측도를 부여하는 것이 자연스러울 것이다. 이런 관점에서, 다음 정리는 product $\sigma-$algebra에 자연스러운 측도를 정의할 구체적인 방법을 제공하는 동시에, 이러한 측도의 유일성까지 보장한다.

Lemma 1.120 $\sigma-$유한 측도공간 $(X,\mathscr{A},\mu),(Y,\mathscr{B},\nu)$와 임의의 $A\in\mathscr{A}\otimes\mathscr{B}$에 대해 함수 $\varphi_A : X\to\overline{\mathbb{R}}_0^+$를 $\varphi_A : x\mapsto\nu(A_x)$로 정의하면 이는 $\mathscr{A}/\mathscr{B}_1-$가측이다. 비슷하게, 함수 $\psi_A : Y\to\overline{\mathbb{R}}_0^+$를 $\psi_A : y\mapsto\mu(A^y)$로 정의하면 이는 $\mathscr{B}/\mathscr{B}_1-$가측이다.

Proof 먼저 ν가 유한한 경우를 생각하여 $\mathscr{L}=\{A\in\mathscr{A}\otimes\mathscr{B} : \varphi_A$가 $\mathscr{A}/\mathscr{B}_1-$가측$\}$을 생각하면 임의의 $A\times B\in\mathscr{A}\times\mathscr{B}$에 대해 $\varphi_{A\times B}(x)=\nu((A\times B)_x)=\nu(B)\mathbf{1}_A(x)$이므로 이가 단순함수가 되어 $A\times B\in\mathscr{L}$이고, 곧 $\mathscr{A}\times\mathscr{B}\subseteq\mathscr{L}$이다. 이제 $\mathscr{A}\times\mathscr{B}$가 $\pi-$system임은 분명하므로 만약 우리가 $\mathscr{L}$이 $\lambda-$system임을 보일 수 있으면 Dynkin의 $\pi-\lambda$ 정리로부터 $\mathscr{L}=\mathscr{A}\otimes\mathscr{B}$가 되어 증명이 끝난다. 이를 위해 임의의 $A\in\mathscr{L}$를 택하면 $\varphi_{A^c}(x)=\nu((A^c)_x)=\nu(Y\setminus A_x)=\nu(Y)-\nu(A_x)=\nu(Y)-\varphi_A(x)$에서 이가 $\mathscr{A}/\mathscr{B}_1-$가측이므로 $A^c\in\mathscr{L}$이고, 곧 $\mathscr{L}$은 여집합에 대해 닫혀있다. 또한, $\mathscr{L}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 $\varphi_{\bigsqcup_{i=1}^\infty A_i}(x)=\nu((\bigsqcup_{i=1}^\infty A_i)_x)=\nu(\bigsqcup_{i=1}^\infty (A_i)_x)=\sum_{i=1}^\infty\nu((A_i)_x)=\sum_{i=1}^\infty\varphi_{A_i}(x)$에서 이가 $\mathscr{A}/\mathscr{B}_1-$가측이므로 $\bigsqcup_{i=1}^\infty A_i\in\mathscr{L}$이고, $X\times Y\in\mathscr{L}$임은 분명하므로 곧 $\mathscr{L}$이 $\lambda-$system이 되어 증명이 끝난다.

한편, ν가 $\sigma-$유한한 경우에는 정의로부터 $\mathscr{B}$에 속하는 적당한 집합열 $\{B_i\}$가 존재하여 각 $i\in\mathbb{N}$에 대해 $\nu(B_i)<\infty$이고 $\bigcup_{i=1}^\infty B_i=Y$이다. WLOG, 필요하다면 각 항을 $C_i := B_i\setminus\bigcup_{j=1}^{i-1}A_j$로 바꾸어 $\{B_i\}$가 처음부터 서로소였다고 해도 된다. 이제 각 $i\in\mathbb{N}$에 대해 측도 ν_{B_i}가 유한하므로 곧 임의의 $A\in\mathscr{A}\otimes\mathscr{B}$에 대해 함수 $\varphi_A^i : X\to\overline{\mathbb{R}}_0^+$를 $\varphi_A^i : x\mapsto\nu_{B_i}(A_x)$로 두면 앞선 결과로부터 이가 $\mathscr{A}/\mathscr{B}_1-$가측이고, 곧 $\varphi_A=\sum_{i=1}^\infty\varphi_A^i$에서 φ_A가 $\mathscr{A}/\mathscr{B}_1-$가측이다. 남은 ψ_A에 대해서도 비슷하게 하면 된다. □

Theorem 1.121 $\sigma-$유한 측도공간 $(X,\mathscr{A},\mu),(Y,\mathscr{B},\nu)$에 대해 적당한 $\mathscr{A}\otimes\mathscr{B}$ 위의 측도 $\mu\otimes\nu : \mathscr{A}\otimes\mathscr{B}\to\overline{\mathbb{R}}_0^+$가 존재하여 임의의 $A\times B\in\mathscr{A}\times\mathscr{B}$에 대해 $(\mu\otimes\nu)(A\times B)=\mu(A)\nu(B)$이다. 나아가, 이러한 측도는 유일하고, $\sigma-$유한이며 $(\mu\otimes\nu)(A)=\int_X\nu(A_x)\,d\mu(x)=\int_Y\mu(A^y)\,d\nu(y)$이다.

Proof 가정으로부터 각각 $\mathscr{A},\mathscr{B}$에 속하는 적당한 집합열 $\{A_i\},\{B_j\}$가 존재하여 각 $i,j\in\mathbb{N}$에 대해 $\mu(A_i),\nu(B_j)<\infty$이고 $\bigcup_{i=1}^\infty A_i=X,\bigcup_{j=1}^\infty B_j=Y$이다. 이제, 집합열 $\{A_i\times B_j\}_{ij}$는 $\mathscr{A}\otimes\mathscr{B}$에 속하는 집합열로 $\bigcup_{i=1}^\infty\bigcup_{j=1}^\infty(A_i\times B_j)=X\times Y$이며 만약 정리의 조건을 만족하는 측도가 존재하기만 한다면 $(\mu\otimes\nu)(A_i\times B_j)=\mu(A_i)\nu(B_i)<\infty$가 되어 그 측도는 항상 $\sigma-$유한하다. 또한, 집합족 $\mathscr{A}\times\mathscr{B}$가 $\pi-$system임이 분명하므로 정리 ? ? 에서 우리가 원하는 측도는 존재하기만 한다면 유일하다. 따라서 증명은 정리의 조건을 만족하는 측도가 존재함을 보이는 것으로 충분하다.

이를 위해 함수 $\xi : \mathscr{A}\otimes\mathscr{B} \to \overline{\mathbb{R}}_0^+$를 $\xi : A \mapsto \int_X \nu(A_x)\,d\mu(x)$로 두고 이가 정리의 조건을 만족하는 측도임을 보이자. 우선 위의 보조정리로부터 이는 well-define되며 $\xi(\emptyset) = \int_X \nu(\emptyset)\,d\mu = 0$임은 분명하고, $\mathscr{A}\otimes\mathscr{B}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 MCT로부터

$$\begin{aligned}
\xi\left(\bigsqcup_{i=1}^{\infty} A_i\right) &= \int_X \nu\left(\left(\bigsqcup_{i=1}^{\infty} A_i\right)_x\right) d\mu \\
&= \int_X \nu\left(\bigsqcup_{i=1}^{\infty} (A_i)_x\right) d\mu \\
&= \int_X \sum_{i=1}^{\infty} \nu((A_i)_x) d\mu \\
&= \sum_{i=1}^{\infty} \int_X \nu((A_i)_x) d\mu \\
&= \sum_{i=1}^{\infty} \xi(A_i)
\end{aligned}$$

이므로 ξ가 $\mathscr{A}\otimes\mathscr{B}$ 위의 측도임을 안다. 한편, 임의의 $A\times B \in \mathscr{A}\times\mathscr{B}$에 대해 $\xi(A\times B) = \int_X \nu((A\times B)_x)\,d\mu(x) = \int_X \nu(B)\mathbf{1}_A(x)\,d\mu(x) = \mu(A)\nu(B)$에서 ξ는 정리의 조건도 만족하여 증명이 끝난다. 한편, 비슷한 방법으로 $(\mu\otimes\nu)(A) = \int_Y \mu(A^y)\,d\nu(y)$임도 보일 수 있다. □

Definition 1.122 $\sigma-$유한 측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$에 대해 μ와 ν의 **곱측도(product measure)**를 $\mu\otimes\nu$로 쓰고 모든 $A\times B \in \mathscr{A}\times\mathscr{B}$에 $\mu(A)\nu(B)$의 측도를 부여하는 $\mathscr{A}\otimes\mathscr{B}$ 위의 유일한 측도로 정의한다. 나아가, 이때의 측도공간 $(X\times Y, \mathscr{A}\otimes\mathscr{B}, \mu\otimes\nu)$를 $(X, \mathscr{A}, \mu)$와 $(Y, \mathscr{B}, \nu)$의 **product measure space**라 한다.

위의 정리에서 곱해지는 두 측도공간이 $\sigma-$유한해야 한다는 조건을 필수적이다. 만약 이들의 $\sigma-$유한성이 보장되지 않으면 Carathéodory의 확장정리로부터 존재성은 여전히 보장되지만 유일성은 더 이상 성립하지 않는다.[15]하지만, 우리가 주로 다룰 Lebesgue 측도공간, Borel 측도공간, 나중에 소개될 확률공간이 모두 $\sigma-$유한하므로 $\sigma-$유한의 가정이 깨어지는 경우에 대해 이 이상으로 논하지는 않겠다. 한편, 이렇게 정의된 측도 간의 곱이 과연 결합법칙을 만족시키는지에 대한 의문이 생겨난다.

Theorem 1.123 $\sigma-$유한 측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu), (Z, \mathscr{C}, \xi)$에 대해 $(\mu\otimes\nu)\otimes\xi = \mu\otimes(\nu\otimes\xi)$이다.

Proof 정리 ? ? 로부터 $\mathscr{A}\times\mathscr{B}\times\mathscr{C}$가 $\mathscr{A}\otimes\mathscr{B}\otimes\mathscr{C}$의 생성자이고, 이가 명백히 $\pi-$system이므로 정리 ? ? 로부터 $(\mu\otimes\nu)\otimes\xi$와 $\mu\otimes(\nu\otimes\xi)$가 $\mathscr{A}\times\mathscr{B}\times\mathscr{C}$에서 일치함을 보이는 것으로 충분한데, 이는 곱측도의 정의로부터 자명하다. □

따라서 측도의 곱이 결합법칙을 만족하므로 σ−유한 측도공간 $(X_1, \mathscr{A}_1, \mu_1), \cdots, (X_k, \mathscr{A}_k, \mu_k)$에 대해 $\bigotimes_{i=1}^k \mu_i = \mu_1 \otimes \cdots \otimes \mu_k$가 well-define된다. 만약 $\mu_1 = \cdots = \mu_k =: \mu$이면 이들의 곱을 $\mu^{\otimes k}$로 간단히 쓰기도 한다.

이로써 Lebesgue 측도의 두 번째 구성을 위해 준비가 끝났다.

Theorem 1.124 $l+m=n$인 $l, m, n \in \mathbb{N}$에 대해 다음이 성립한다.

i. Borel 측도공간 $(\mathbb{R}^n, \mathscr{B}_n, \mu_n)$은 $(\mathbb{R}^l, \mathscr{B}_l, \mu_l)$과 $(\mathbb{R}^m, \mathscr{B}_m, \mu_m)$의 product measure space 이다. 즉, $(\mathbb{R}^n, \mathscr{B}_n, \mu_n) = (\mathbb{R}^l \times \mathbb{R}^m, \mathscr{B}_l \otimes \mathscr{B}_m, \mu_l \otimes \mu_m)$.
ii. Lebesgue 측도공간 $(\mathbb{R}^n, \mathscr{M}_n, \lambda_n)$은 $(\mathbb{R}^l, \mathscr{M}_l, \lambda_l)$과 $(\mathbb{R}^m, \mathscr{M}_m, \lambda_m)$의 product measure space의 완비화이다. 즉, $(\mathbb{R}^n, \mathscr{M}_n, \lambda_n) = (\mathbb{R}^l \times \mathbb{R}^m, \overline{\mathscr{M}_l \otimes \mathscr{M}_m}, \overline{\lambda_l \otimes \lambda_m})$.

Proof i과 ii를 총 4단계에 걸쳐 한 번에 증명하도록 하자.

Step 1. $\mathscr{B}_n = \mathscr{B}_l \otimes \mathscr{B}_m$.

임의의 열린 $U \subseteq \mathbb{R}^n$에 대해 이는 $\mathbb{R}^l$과 $\mathbb{R}^m$의 열린집합의 Cartesian 곱의 가산 합집합으로 쓸 수 있으므로[16] $\mathscr{U}$를 $\mathbb{R}^n$의 모든 열린집합의 모임이라 하면 $\mathscr{U} \subseteq \mathscr{B}_l \otimes \mathscr{B}_m$에서 $\mathscr{B}_n \subseteq \mathscr{B}_l \otimes \mathscr{B}_m$이다. 이제 $\pi_1 : \mathbb{R}^n \to \mathbb{R}^l, \pi_2 : \mathbb{R}^n \to \mathbb{R}^m$를 각각 처음 l개의 좌표와 마지막 m개의 좌표로의 사영이라 한다면 이들이 연속함수이므로 Borel이고, 곧 임의의 $A \in \mathscr{B}_l, B \in \mathscr{B}_m$에 대해 $A \times \mathbb{R}^m = \pi_1^{-1}(A), \mathbb{R}^l \times B = \pi_2^{-1}(B)$가 모두 Borel이 되어 $A \times B = (A \times \mathbb{R}^m) \cap (\mathbb{R}^l \times B) \in \mathscr{B}_n$이므로 $\mathscr{B}_l \times \mathscr{B}_m \subseteq \mathscr{B}_n$에서 $\mathscr{B}_l \otimes \mathscr{B}_m \subseteq \mathscr{B}_n$이고, 이로써 $\mathscr{B}_n = \mathscr{B}_l \otimes \mathscr{B}_m$임을 안다.

Step 2. 임의의 $A \in \mathscr{B}_n$에 대해 이는 $\mathscr{M}_l \otimes \mathscr{M}_m$−가측이고 $\lambda_n(A) = (\lambda_l \otimes \lambda_m)(A)$이다.

1단계로부터 $\mathscr{B}_n = \mathscr{B}_l \otimes \mathscr{B}_m \subseteq \mathscr{M}_l \otimes \mathscr{M}_m$이고, 따라서 A가 $\mathscr{M}_l \otimes \mathscr{M}_m$−가측임은 분명하다. 이제 남은 부분을 보이기 위해 먼저 $A \in \mathscr{S}_n$인 경우를 생각해보면, 적당한 반열린구간 $I_1, \cdots, I_n \subseteq \mathbb{R}$이 존재하여 $A = \prod_{i=1}^n I_i$이다. 이로부터 $\lambda_n(A) = \prod_{i=1}^n \lambda_1(I_i) = [\prod_{i=1}^l \lambda_1(I_i)][\prod_{i=l+1}^n \lambda_1(I_i)] = \lambda_l(\prod_{i=1}^l I_i)\lambda_m(\prod_{i=l+1}^n I_i) = (\lambda_l \otimes \lambda_m)(A)$가 성립한다. 다음으로, A가 열린집합인 경우를 생각해보면 $\mathscr{S}_n$에 속하는 적당한 서로소인 집합열 $\{B_j\}$에 대해 $A = \bigsqcup_{j=1}^\infty B_j$로 쓸 수 있으므로[17] 이 경우에도 $\lambda_n(A) = \lambda_n(\bigsqcup_{j=1}^\infty B_j) = \sum_{j=1}^\infty \lambda_n(B_j) = \sum_{j=1}^\infty (\lambda_l \otimes \lambda_m)(B_j) = (\lambda_l \otimes \lambda_m)(\bigsqcup_{j=1}^\infty B_j) = (\lambda_l \otimes \lambda_m)(A)$가 성립한다. 이번에는 A가 compact한 경우를 생각해보면 이는 유계이므로 적당한 유계인 열린집합 $U \subseteq \mathbb{R}^n$가 존재하여 $A \subseteq U$이고 $\lambda_n(U) < \infty$이다. 따라서 $V = U \setminus A$라 하면 이 또한 열려있고 유계이므로 $\lambda_n(V) < \infty$이어서 이전의 결과로부터 $\lambda_n(A) = \lambda_n(U \setminus V) = \lambda_n(U) - \lambda_n(V) = (\lambda_l \otimes \lambda_m)(U) - (\lambda_l \otimes \lambda_m)(V) = (\lambda_l \otimes \lambda_m)(U \setminus V) = (\lambda_l \otimes \lambda_m)(A)$를 얻는다. 마지막으로, 일반적인 $A \in \mathscr{B}_n$에 대해서는 앞선 결과들과 정리 ? ? 로부터

$$\begin{aligned}\lambda_n(A) &= \inf\{\lambda_n(U) \in \overline{\mathbb{R}}_0^+ : A \subseteq U \subseteq \mathbb{R}^n, U\text{는 열린집합}\}\\ &= \inf\{(\lambda_l \otimes \lambda_m)(U) \in \overline{\mathbb{R}}_0^+ : A \subseteq U \subseteq \mathbb{R}^n, U\text{는 열린집합}\}\\ &\geq (\lambda_l \otimes \lambda_m)(A)\end{aligned}$$

이고

$$\begin{aligned}\lambda_n(A) &= \sup\{\lambda_n(K) \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K\text{는 compact 집합}\}\\ &= \sup\{(\lambda_l \otimes \lambda_m)(K) \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K\text{는 compact 집합}\}\\ &\leq (\lambda_l \otimes \lambda_m)(A)\end{aligned}$$

이므로 $\lambda_n(A) = (\lambda_l \otimes \lambda_m)(A)$임을 안다.

이상의 결론으로부터 임의의 $A \times B \in \mathscr{B}_l \times \mathscr{B}_m \subseteq \mathscr{B}_n$에 대해 $\mu_n(A \times B) = \mu_l(A)\mu_m(B)$이므로 정리 ? ? 로부터 $\mu_n = \mu_l \otimes \mu_m$이고, 곧 i의 증명이 끝난다.

Step 3. $\mathscr{M}_l \otimes \mathscr{M}_m \subseteq \mathscr{M}_n$.

임의의 $A \in \mathscr{M}_l = \overline{\mathscr{B}_l}$를 택하면 정리 ? ? 로부터 적당한 $B, C \in \mathscr{B}_l$가 존재하여 $B \subseteq A \subseteq C$이고 $\mu_l(C \setminus B) = 0$이다. 이제 $\pi : \mathbb{R}^n \to \mathbb{R}^l$를 처음 l개의 좌표로의 사영이라 하면 이가 연속이므로 Borel이고 곧 $B \times \mathbb{R}^m = \pi^{-1}(B), C \times \mathbb{R}^m = \pi^{-1}(C)$가 Borel이다. 여기에 앞선 단계의 결과들을 적용하면 $B \times \mathbb{R}^m \subseteq A \times \mathbb{R}^m \subseteq C \times \mathbb{R}^m$이고 $\lambda_n((C \times \mathbb{R}^m) \setminus (B \times \mathbb{R}^m)) = \lambda_n((C \setminus B) \times \mathbb{R}^m) = \lambda_l(C \setminus B)\lambda_m(\mathbb{R}^m) = 0$이 되어 정리 ? ? 에서 $A \times \mathbb{R}^m \in \mathscr{M}_n$이다. 비슷하게 임의의 $B \in \mathscr{M}_m$에 대해 $\mathbb{R}^l \times B \in \mathscr{M}_n$임을 보일 수 있으므로 임의의 $A \times B \in \mathscr{M}_l \times \mathscr{M}_m$에 대해 $A \times B = (A \times \mathbb{R}^m) \cap (\mathbb{R}^l \times B) \in \mathscr{M}_n$이 되어 $\mathscr{M}_l \times \mathscr{M}_m \subseteq \mathscr{M}_n$에서 $\mathscr{M}_l \otimes \mathscr{M}_m \subseteq \mathscr{M}_n$이다.

Step 4. $\mathscr{M}_n = \overline{\mathscr{M}_l \otimes \mathscr{M}_m}$이고 $\lambda_n = \overline{\lambda_l \otimes \lambda_m}$이다.

먼저 $\mathscr{M}_n \subseteq \overline{\mathscr{M}_l \otimes \mathscr{M}_m}$이고 $\lambda_n = \overline{\lambda_l \otimes \lambda_m}|_{\mathscr{M}_n}$임을 보이기 위해 임의의 $A \in \mathscr{M}_n = \overline{\mathscr{B}_n}$를 택하면 정리 ? ? 로부터 적당한 $B, C \in \mathscr{B}_n \subseteq \mathscr{M}_l \otimes \mathscr{M}_m$가 존재하여 $B \subseteq A \subseteq C$이고 $(\lambda_l \otimes \lambda_m)(C \setminus B) = \lambda_n(C \setminus B) = 0$이다. 그런데 이는 다시 정리 ? ? 로부터 $A \in \overline{\mathscr{M}_l \otimes \mathscr{M}_m}$임을 뜻하여 곧 $\mathscr{M}_n \subseteq \overline{\mathscr{M}_l \otimes \mathscr{M}_m}$이다. 나아가 $A = B \sqcup (A \setminus B)$이고 $A \setminus B \subseteq C \setminus B$가 $\lambda_l \otimes \lambda_m$–영집합이므로 $\overline{\lambda_l \otimes \lambda_m}(A) = (\lambda_l \otimes \lambda_m)(B) = \lambda_n(B) = \lambda_n(B) + \lambda_n(A \setminus B) = \lambda_n(B \sqcup (A \setminus B)) = \lambda_n(A)$에서 $\overline{\lambda_l \otimes \lambda_m}|_{\mathscr{M}_n} = \lambda_n$임을 안다. 이제 $\mathscr{M}_n \supseteq \overline{\mathscr{M}_l \otimes \mathscr{M}_m}$만 보이면 증명이 끝난다. 이를 위해 임의의 $A \in \overline{\mathscr{M}_l \otimes \mathscr{M}_m}$를 택하면 정리 ? ? 와 앞선 결과로부터 적당한 $B, C \in \mathscr{M}_l \otimes \mathscr{M}_m \subseteq \mathscr{M}_n$가 존재하여 $B \subseteq A \subseteq C$이고 $\lambda_n(C \setminus B) = (\lambda_l \otimes \lambda_m)(C \setminus B) = 0$이다. 그렇다면 $A \setminus B \subseteq C \setminus B$는 λ_n–영집합이고 정리 ? ? 로부터 $\mathscr{M}_n$이 완비성을 가지므로 $A \setminus B \in \mathscr{M}_n$이 되어 $A = B \sqcup (A \setminus B) \in \mathscr{M}_n$이다. 그렇다면 $\mathscr{M}_n \supseteq \overline{\mathscr{M}_l \otimes \mathscr{M}_m}$에서 ii의 증명도 끝난다. □

이로부터 우리는 n차원 Lebesgue 측도공간을 재귀적으로 구성할 수 있다. 먼저 첫 번째 Lebesgue 측도의 구성을 따라 $(\mathbb{R}, \mathscr{M}_1, \lambda_1)$을 구성하고, 이로써 2차원 Lebesgue 측도공간을 $(\mathbb{R}^2, \mathscr{M}_2, \lambda_2) := (\mathbb{R} \times \mathbb{R}, \overline{\mathscr{M}_1 \otimes \mathscr{M}_1}, \overline{\lambda_1 \otimes \lambda_1})$와 같이 정의한다. 다음으로 3차원 Lebesgue 측도공간을 $(\mathbb{R}^3, \mathscr{M}_3, \lambda_3) := (\mathbb{R}^2 \times \mathbb{R}, \overline{\mathscr{M}_2 \otimes \mathscr{M}_1}, \overline{\lambda_2 \otimes \lambda_1})$로 정의하고, 이를 충분히 반복하면 원하는 모든 $n \in \mathbb{N}$에 대해 n차원 Lebesgue 측도공간을 구성할 수 있다. 한편, n차원 Borel

측도공간은 Lebesgue 측도공간의 구성과 달리 완비화가 필요하지 않으므로 $(\mathbb{R}^n, \mathscr{B}_1^{\otimes n}, \mu_1^{\otimes n})$으로 더 깔끔하게 정의할 수 있다.

본 장에서의 우리의 이야기는 여기서 끝을 맺는다. 이어지는 절에서는 이후의 논의에서 필요하다고 생각되는 흥미로운 주제들에 대해 알아보도록 하겠다.

1.9 Riemann vs. Lebesgue

앞서 살펴본 바와 같이 Lebesgue 적분은 Riemann 적분이 극한과 매끄럽게 상호작용하지 못한다는 단점을 잘 보완해준다. 그런데, 우리는 여기서 과연 Lebesgue 적분이 정말 Riemann 적분의 일반화인지에 대한 질문을 조용히 숨겨두었다. 만약 우리가 그토록 공을 들여 전개한 Lebesgue 적분론이 Riemann 적분론과 다른 별개의 이론이라면 Riemann 적분론에서 성립하는 FTC와 같은 수많은 이론들을 Lebesgue 적분론에서는 사용할 수가 없을 것이다. 이건 빈대 잡자고 초가삼간 다 태우는 격이 아닌가! 당장 FTC가 없으면 아주 단순한 경우가 아니고서야 적분값을 계산할 수조차 없다. 이번 절에서는 우리가 회피했던 이 질문을 다루어보고자 한다. 한편, 다양한 문헌에서 서로 동등하지만 조금씩 다른 접근법을 통해 Riemann 적분을 정의하므로 일관된 논의를 위해 이번 절에서 사용할 Riemann 적분의 정의를 중간중간 옮겨 두었다. 이하의 정의는 [1]에서의 Riemann 적분의 정의를 측도론의 맥락에 맞도록 조금 수정한 것으로 Darboux의 접근방식이라 불린다.

Definition 1.125 유계인 함수 $f : \mathbb{R}^n \to \mathbb{R}$와 유계인 집합 $A \in \mathscr{M}_n$에 대해 $A \subseteq B$인 임의의 유계인 닫힌 box $B \subseteq \mathbb{R}^n$를 택하여 $\mathscr{P}$를 B의 분할이라 하자. 이때, $\mathscr{P}$에 대한 f의 **하합(lower sum)**을 $L(f, \mathscr{P})$로 쓰고 $L(f, \mathscr{P}) = \sum_{P\in\mathscr{P}} \inf_{x\in P} f\mathbf{1}_A \mathrm{vol}(P)$로 정의한다. (여기서 $\mathrm{vol}(\cdot)$은 $\mathbb{R}^n$의 유계인 box의 '부피'로서 각 모서리의 길이의 곱으로 정의된다.) 비슷하게, $\mathscr{P}$에 대한 f의 **상합(upper sum)**을 $U(f, \mathscr{P})$로 쓰고 $U(f, \mathscr{P}) = \sum_{P\in\mathscr{P}} \sup_{x\in P} f\mathbf{1}_A \mathrm{vol}(P)$로 정의한다. 이제 f의 **A에서의 하적분(lower integral over A)**을 $\underline{\int_A} f(x)\,dx$ 혹은 간단히 $\underline{\int_A} f$로 쓰고, $\underline{\int_A} f = \sup_{\mathscr{P}\text{ partition}} L(f, \mathscr{P})$로 정의한다. 비슷하게, f의 **A에서의 상적분(upper integral over A)**을 $\overline{\int_A} f(x)\,dx$ 혹은 간단히 $\overline{\int_A} f$로 쓰고 $\overline{\int_A} f = \inf_{\mathscr{P}\text{ partition}} U(f, \mathscr{P})$로 정의하여 만약 $\underline{\int_A} f = \overline{\int_A} f$이면 이때 f가 A **위에서 (Riemann) 적분가능(− integrable over A)**하다고 하고, 그 공통의 값을 $\int_A f(x)\,dx$ 혹은 간단히 $\int_A f$로 쓴다.

Lemma 1.126 임의의 $i \le n$와 임의의 $x_0 \in \mathbb{R}$에 대해 $\mathbb{R}^n$에 속하는 초평면 $\pi : x_i = x_0$는 영집합이다.

Proof 우선 π가 닫힌집합이므로 Borel이고, 따라서 이는 가측이다. 한편, Lebesgue 측도가 이동 불변성을 가지므로 $x_i = 0$인 경우에 대해서만 증명하는 것으로 충분하고, 간결한 논

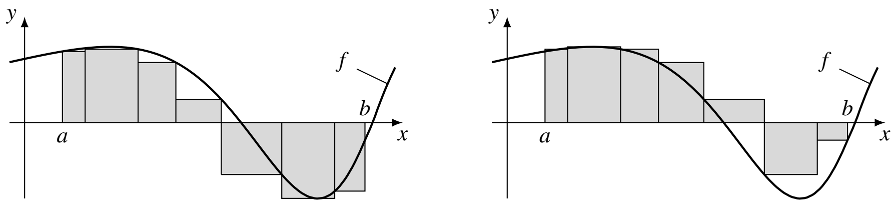


Figure 1.9 구간 $[a,b]$에서의 분할 $\mathscr{P}$에 대한 함수 f의 하합 $L(f,\mathscr{P})$ (오른쪽)와 상합 $U(f,\mathscr{P})$ (왼쪽).

의를 위해 $i=1$인 경우에 대해서만 보이도록 하자. (다른 경우에도 이와 비슷하게 보일 수 있다.) 이제 임의의 $\varepsilon>0$을 택하고 집합열 $\{A_i\}$를 $A_i=(-\varepsilon/2^{i+n}i^{n-1},\varepsilon/2^{i+n}i^{n-1}]\times(-i,i]^{n-1}$로 두면 이는 $\mathscr{S}_n$에 속하는 집합열로서 각 $i\in\mathbb{N}$에 대해 $\lambda_n(A_i)=\varepsilon/2^i$이고 $\pi\subseteq\bigcup_{i=1}^{\infty}A_i$이다. 이로부터 $\lambda_n(\pi)\le\lambda_n(\bigcup_{i=1}^{\infty}A_i)\le\sum_{i=1}^{\infty}\lambda_n(A_i)=\varepsilon$이 되어 보조정리가 성립한다. □

Theorem 1.127 유계인 함수 $f:\mathbb{R}^n\to\mathbb{R}$와 유계인 집합 $A\in\mathscr{M}_n$에 대해 만약 f가 A 위에서 Riemann 적분가능하다면 $f\mathbf{1}_A$는 Lebesgue 적분가능하며, 이때의 두 적분은 서로 일치한다. 즉, $\int_A f=\int_{\mathbb{R}^n}f\mathbf{1}_A\,d\lambda_n$이다.

Proof 먼저 A가 닫힌 box이고 f가 음이 아닌 특별한 경우를 생각하여 A의 임의의 분할 $\mathscr{P}$에 대해 단순함수 $\varphi_{\mathscr{P}},\psi_{\mathscr{P}}:\mathbb{R}^n\to\mathbb{R}_0^+$를 각각 $\varphi_{\mathscr{P}}=\sum_{P\in\mathscr{P}}\inf_{x\in P}f(x)\mathbf{1}_{\widetilde{P}}$, $\psi_{\mathscr{P}}=\sum_{P\in\mathscr{P}}\sup_{x\in P}f(x)\mathbf{1}_{\widetilde{P}}$로 두자. (여기서 $\widetilde{P}$는 box P의 남서쪽 면을 빼서 얻는 semi-open box로서 적당한 $x,y\in\mathbb{R}^n$에 대해 $P=[x,y]$라면 $\widetilde{P}=(x,y]$이다.) 그렇다면

$$\begin{aligned}\int_{\mathbb{R}^n}\varphi_{\mathscr{P}}\,d\lambda_n&=\sum_{P\in\mathscr{P}}\inf_{x\in P}f(x)\int_{\mathbb{R}^n}\mathbf{1}_{\widetilde{P}}\,d\lambda_n\\&=\sum_{P\in\mathscr{P}}\inf_{x\in P}f(x)\lambda_n(\widetilde{P})\\&=\sum_{P\in\mathscr{P}}\inf_{x\in P}f(x)\mathrm{vol}(P)\\&=L(f,\mathscr{P})\end{aligned}$$

이고 비슷하게 $\int_{\mathbb{R}^n}\psi_{\mathscr{P}}\,d\lambda_n=U(f,\mathscr{P})$이다. 이제 f가 A위에서 Riemann 적분가능하므로 임의의 $i\in\mathbb{N}$에 대해 적당한 A의 분할 $\mathscr{P}_i$가 존재하여 $U(f,\mathscr{P}_i)-L(f,\mathscr{P}_i)<1/i$이고, 이들을 모아 분할의 열 $\{\mathscr{P}_i\}$를 구성할 수 있고, WLOG, 필요하다면 각 항을 $\mathscr{Q}_i:=\bigcup_{j=1}^{i}\mathscr{P}_j$로 바꾸어 $\{\mathscr{P}_i\}$가 처음부터 증가했다고 해도 된다. 그렇다면 $\{\varphi_{\mathscr{P}_i}\},\{\psi_{\mathscr{P}_i}\}$는 각각 증가하고 감소하는 함수열로 $\{\psi_{\mathscr{P}_i}-\varphi_{\mathscr{P}_i}\}$가 감소하는 음이 아닌 함수열이 되어 이가 모든 점에서 수렴하고, 곧 Fatou의 보조정리로부터

$$\int_{\mathbb{R}^n}\lim_{i\to\infty}(\psi_{\mathscr{P}_i}-\varphi_{\mathscr{P}_i})\,d\lambda_n\le\liminf_{i\to\infty}\int_{\mathbb{R}^n}\psi_{\mathscr{P}_i}-\varphi_{\mathscr{P}_i}\,d\lambda_n$$

$$
\begin{aligned}
&= \lim_{i\to\infty} \left(\int_{\mathbb{R}^n} \psi_{\mathscr{P}_i}\, d\lambda_n - \int_{\mathbb{R}^n} \varphi_{\mathscr{P}_i}\, d\lambda_n \right) \\
&= \lim_{i\to\infty} [U(f, \mathscr{P}_i) - L(f, \mathscr{P}_i)] \\
&= 0
\end{aligned}
$$

이다. 이는 정리 ? ? 의 ii로부터 $\psi_{\mathscr{P}_i} - \varphi_{\mathscr{P}_i} \downarrow 0$ (ae.)임을 뜻하고, 특히 위의 보조정리로부터 모든 $i \in \mathbb{N}$에 대해 $\varphi_{\mathscr{P}_i} \le f\mathbf{1}_A \le \psi_{\mathscr{P}_i}$ (ae.)가 성립하므로[18] $\varphi_{\mathscr{P}_i} \uparrow f\mathbf{1}_A$ (ae.)가 되어 정리 ? ? 로부터 $f\mathbf{1}_A$가 가측임을 안다. 이제 MCT로부터 $\int_{\mathbb{R}^n} f\mathbf{1}_A\, d\lambda_n = \lim_{i\to\infty} \int_{\mathbb{R}^n} \varphi_{\mathscr{P}_i}\, d\lambda_n = \lim_{i\to\infty} U(f, \mathscr{P}_i) = \int_A f < \infty$이므로 $f\mathbf{1}_A$는 Lebesgue 적분가능하고 두 적분이 일치한다.

한편, 일반적인 경우에 대해서는 $A \subseteq B \in \mathbb{R}^n$인 임의의 닫힌 box B를 잡으면 이전의 결과로부터

$$
\begin{aligned}
\int_{\mathbb{R}^n} f\mathbf{1}_A\, d\lambda_n &= \int_{\mathbb{R}^n} (f\mathbf{1}_A)_+\, d\lambda_n - \int_{\mathbb{R}^n} (f\mathbf{1}_A)_-\, d\lambda_n \\
&= \int_{\mathbb{R}^n} f_+\mathbf{1}_A\mathbf{1}_B\, d\lambda_n - \int_{\mathbb{R}^n} f_-\mathbf{1}_A\mathbf{1}_B\, d\lambda_n \\
&= \int_B f_+\mathbf{1}_A - \int_B f_-\mathbf{1}_A \\
&= \int_A f_+ - \int_A f_- \\
&= \int_A f
\end{aligned}
$$

이고, 곧 일반적인 경우에도 $f\mathbf{1}_A$는 Lebesgue 적분가능하며 두 적분이 일치한다. □

Definition 1.128 유계인 음이 아닌 함수 $f : \mathbb{R}^n \to \mathbb{R}_0^+$와 집합 $A \in \mathscr{M}_n$에 대해 $f\mathbf{1}_A$가 적당한 $x > 0$에 대해 $[-x, x]^n$으로 표현되는 모든 닫힌 box 위에서 Riemann 적분가능하다고 하자. 만약 $\lim_{x\to\infty} \int_{[-x,x]^n} f\mathbf{1}_A < \infty$이면 이때 f가 A **위에서 (Riemann) 적분가능(− integrable over** A**)**하다고 하고, 그 극한을 $\int_A f(x)\, dx$ 혹은 간단히 $\int_A f$로 쓴다. 나아가, 함수 $f : \mathbb{R}^n \to \mathbb{R}$에 대해서는 $f_\pm$이 모두 A 위에서 Riemann 적분가능하다면 이때 f가 A **위에서 (Riemann) 적분가능(− integrable over** A**)**하다고 하고, $\int_A f_+ - \int_A f_-$의 값을 $\int_A f(x)\, dx$ 혹은 간단히 $\int_A f$로 쓴다.

Theorem 1.129 유계인 함수 $f : \mathbb{R}^n \to \mathbb{R}$와 집합 $A \in \mathscr{M}_n$에 대해 만약 f가 A 위에서 Riemann 적분가능하다면 $f\mathbf{1}_A$는 Lebesgue 적분가능하며, 이때의 두 적분은 서로 일치한다. 즉, $\int_A f = \int_{\mathbb{R}^n} f\mathbf{1}_A\, d\mu$이다.

Proof 함수 $f_i : \mathbb{R}^n \to \mathbb{R}$의 열 $\{f_i\}$를 $f_i = f\mathbf{1}_{A\cap[-i,i]^n}$으로 정의하면 이는 유계함수열이고 $f_i \to f\mathbf{1}_A$인데 각 $i \in \mathbb{N}$에 대해 f_i가 $[-i, i]^n$ 위에서 Riemann 적분가능하므로 정리 ? ? 로부터 가측이고, 곧 정리 ? ? 로부터 $f\mathbf{1}_A$도 가측이다. 이제 f가 음이 아닌 특별한 경우

를 생각하면 WLOG, $\{f_i\}$가 증가한다고 할 수 있다. 그렇다면 MCT와 정리 ? ? 로부터 $\int_{\mathbb{R}^n} f\mathbf{1}_A\,d\lambda_n = \lim_{i\to\infty}\int_{\mathbb{R}^n} f_i\,d\lambda_n = \lim_{i\to\infty}\int_{[-i,i]^n} f\mathbf{1}_A = \int_A f < \infty$가 되어 $f\mathbf{1}_A$는 Lebesgue 적분가능하고 두 적분이 일치한다. 한편, 일반적인 f에 대해서는

$$\begin{aligned}\int_{\mathbb{R}^n} f\mathbf{1}_A\,d\lambda_n &= \int_{\mathbb{R}^n}(f\mathbf{1}_A)_+\,d\lambda_n - \int_{\mathbb{R}^n}(f\mathbf{1}_A)_-\,d\lambda_n\\ &= \int_{\mathbb{R}^n} f_+\mathbf{1}_A\,d\lambda_n - \int_{\mathbb{R}^n} f_-\mathbf{1}_A\,d\lambda_n\\ &= \int_A f_+ - \int_A f_-\\ &= \int_A f\end{aligned}$$

이므로 이때에도 같은 결론이 성립한다. □

Definition 1.130 음이 아닌 함수 $f:\mathbb{R}^n\to\mathbb{R}_0^+$와 집합 $A\in\mathscr{M}_n$, 충분히 큰 $M\in\mathbb{R}$에 대해 $f_{(M)} := \min\{f, M\}$이 A 위에서 Riemann 적분가능하다고 하자. 만약 $\lim_{M\to\infty}\int_A f_{(M)} < \infty$이면 이때 f가 A **위에서 (Riemann) 적분가능(− integrable over** A)하다고 하고, 그 극한을 $\int_A f(x)\,dx$ 혹은 간단히 $\int_A f$로 쓴다. 나아가, 유계인 함수 $f:\mathbb{R}^n\to\mathbb{R}$에 대해서는 $f_\pm$이 모두 A 위에서 Riemann 적분가능하다면 이때 f가 A **위에서 (Riemann) 적분가능(− integrable over** A)하다고 하고, $\int_A f_+ - \int_A f_-$의 값을 $\int_A f(x)\,dx$ 혹은 간단히 $\int_A f$로 쓴다.

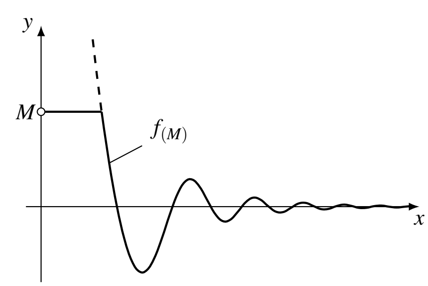


Figure 1.10 함수 f (파선)에 대한 $f_{(M)}$의 그래프. 여기서 f는 $x\downarrow 0$이면 ∞로 발산한다.

Theorem 1.131 함수 $f:\mathbb{R}^n\to\mathbb{R}$와 집합 $A\in\mathscr{M}_n$에 대해 만약 f가 A 위에서 Riemann 적분가능하다면 $f\mathbf{1}_A$는 Lebesgue 적분가능하며, 이때의 두 적분은 서로 일치한다. 즉, $\int_A f = \int_{\mathbb{R}^n} f\mathbf{1}_A\,d\mu$이다.

Proof 함수 $f_i:\mathbb{R}^n\to\mathbb{R}$의 열 $\{f_i\}$를 $f_i = f_{(i)}\mathbf{1}_A$로 정의하면 이는 증가하는 유계함수열이고 $f_i\uparrow f\mathbf{1}_A$인데 충분히 큰 $i\in\mathbb{N}$에 대해 f_i가 A 위에서 Riemann 적분가능하므로 정리 ? ? 로부터 가측이고, 곧 정리 ? ? 로부터 $f\mathbf{1}_A$도 가측이다. 여기서 WLOG, f_i가 모든 $i\in\mathbb{N}$에 대해 Riemann 적분가능하다고 하자. 이제 f가 음이 아닌 특별한 경우를 생각하면 MCT와

정리 ? ? 로부터 $\int_{\mathbb{R}^n} f\mathbf{1}_A\,d\lambda_n = \lim_{i\to\infty}\int_{\mathbb{R}^n} f_i\,d\lambda_n = \lim_{i\to\infty}\int_A f_{(i)} = \int_A f < \infty$가 되어 $f\mathbf{1}_A$는 Lebesgue 적분가능하고 두 적분이 일치한다. 한편, 일반적인 f에 대해서는

$$\begin{aligned}\int_{\mathbb{R}^n} f\mathbf{1}_A\,d\lambda_n &= \int_{\mathbb{R}^n}(f\mathbf{1}_A)_+\,d\lambda_n - \int_{\mathbb{R}^n}(f\mathbf{1}_A)_-\,d\lambda_n \\ &= \int_{\mathbb{R}^n} f_+\mathbf{1}_A\,d\lambda_n - \int_{\mathbb{R}^n} f_-\mathbf{1}_A\,d\lambda_n \\ &= \int_A f_+ - \int_A f_- \\ &= \int_A f\end{aligned}$$

이므로 이때에도 같은 결론이 성립한다. □

따라서 다행히도 Lebesgue 적분은 Riemann 적분을 포괄하는 개념으로, 이론 전개를 할 때는 Lebesgue 적분으로 생각하여 이론을 전개하고, 실제로 계산을 해야 할 때에는 이를 Riemann 적분으로 생각하여 FTC 등을 사용해 계산해도 대부분의 경우 문제가 없다. 다만, 여기서도 조심해야 할 부분이 있는데, 바로 improper한 Riemann 적분은 이가 절대수렴하는 경우에만 Lebesgue 적분과 그 값이 일치한다는 것이다. Riemann 적분이 조건수렴에 그치는 $\int_0^\infty(\sin x)/x\,dx$와 같은 경우 이의 Lebesgue 적분가능성은 일반적으로 보장되지 않는다. 물론, 이 책에서는 조건수렴에 그치는 Riemann 적분을 다룰 일이 (아마) 없을 것이므로 안심해도 좋다.

1.10 L^p Space and Radon-Nikodým Theorem

지금까지 우리는 측도라는 개념으로 넓이나 부피 따위를 일반화하였고, 이를 이용하여 Lebesgue 적분의 개념을 정립함으로써 $\mathbb{R}^n$에서 뿐만 아니라 측도가 주어진 임의의 공간에서 적분을 할 수 있는 이론적 토대를 마련하였다. 이번 절에서는 이런 추상화의 흐름에 발맞추어 적분의 역연산인 미분의 개념을 임의의 공간으로 확장하려고 한다. 한편, 이를 위해서는 L^p 공간이라는 경유지가 필요한데, L^p 공간은 그 자체로도 꽤나 흥미로운 탐구의 대상이어서 파고들자면 끝도 없으므로 여기서는 우리에게 필요한 정도만 선택적으로 그 내용을 소개하도록 하겠다. 이에 대해 조금 더 관심이 있는 독자들은 [2]와 같은 실해석학 교재를 고조하기 바란다.

Definition 1.132 측도공간 $(X,\mathscr{A},\mu)$에서 정의된 가측함수 $f:X\to\overline{\mathbb{R}}$와 $p\geq 1$에 대해 $(\int_X |f|^p\,d\mu)^{1/p}$를 f의 **p−노름(p−norm)**이라 하고 $||f||_{p,\mu}$ 혹은 간단히 $||f||_p$로 쓴다.

Proposition 1.133 측도공간 $(X,\mathscr{A},\mu)$에서 정의된 가측함수 $f:X\to\overline{\mathbb{R}}$와 $p\geq 1$에 대해 다음이 성립한다.

i. 임의의 $\lambda \in \mathbb{R}$에 대해 $||\lambda f||_p = |\lambda|||f||_p$이다.
ii. $||f||_p = 0$일 필요충분조건은 $f = 0$ (ae.)이다.

Proof i. 이는 $||\lambda f||_p = (\int_X |\lambda f|^p d\mu)^{1/p} = |\lambda|(\int_X |f|^p d\mu)^{1/p} = |\lambda|||f||_p$에서 자명하다.

ii. $\int_X |f|^p d\mu = ||f||_p^p = 0$이므로 정리 ? ? 의 ii로부터 $|f| = 0$ (ae.)이고, 곧 $f = 0$ (ae.)이다. 한편, 역은 자명하다. □

Theorem 1.134 (Hölder's inequality)

측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f, g : X \to \overline{\mathbb{R}}$와 $1/p + 1/q = 1$인 $p, q > 1$에 대해 $\int_X |fg| d\mu \leq ||f||_p||g||_q$이다. 이때, $||f||_p, ||g||_q < \infty$라면 등호가 성립할 필요충분조건은 모두 0은 아닌 적당한 $a, b \in \mathbb{R}$가 존재하여 $a|f|^p + b|g|^q = 0$ (ae.)인 것이다.

Proof WLOG, 필요하다면 f, g를 각각 $|f|, |g|$로 대체하여 f, g가 처음부터 음이 아니라 해도 된다. 한편, 만약 $||f||_p = 0$이면 명제 ? ? 의 ii로부터 $f = 0$ (ae.)이고, 곧 $\int_X |fg| d\mu = 0$이 되어 부등식이 자명하므로 $||f||_p > 0$이라 하고, 비슷한 이유로 $||g||_p > 0$이라 하자. 또한 $||f||_p = \infty$이거나 $||g||_p = \infty$인 경우에도 부등식이 자명하므로 $||f||_p, ||g||_p < \infty$라 하자. 이제 $||f||_p = ||g||_q = 1$인 특별한 경우를 생각하면 Young의 부등식으로부터 $\int_X fg d\mu \leq \int_X f^p/p + g^q/q d\mu = ||f||_p^p/p + ||g||_q^q/q = 1/p + 1/q = 1$이다. 이어서 일반적인 경우에 $\widetilde{f} = f/||f||_p, \widetilde{g} = g/||g||_q$라 하면 명제 ? ? 의 i로부터 $||\widetilde{f}||_p = ||\widetilde{g}||_q = 1$이므로 앞선 결과로부터 $\int_X |\widetilde{f}\widetilde{g}| d\mu \leq 1$이 되어 곧 $\int_X |fg| d\mu \leq ||f||_p||g||_q$임을 안다.

다음으로, 등호조건을 보이기 위해 $||f||_p, ||g||_q < \infty$이고 $\int_X |fg| d\mu = ||f||_p||g||_q$라 하자. 만약 $||f||_p = 0$이면 $f = 0$ (ae.)에서 $|f|^p + 0 \cdot |g|^q = 0$ (ae.)가 되어 더 이상 보일 것이 없으므로 $||f||_p \neq 0$이라 하고, 비슷한 이유에서 $||g||_q \neq 0$이라 하자. 이제 $\widetilde{f} = |f|/||f||_p, \widetilde{g} = |g|/||g||_q$라 하면 $\int_X \widetilde{f}\widetilde{g} d\mu = (\int_X |fg| d\mu)/||f||_p||g||_q = 1 = ||\widetilde{f}||_p^p/p + ||\widetilde{g}||_q^q/q = \int_X \widetilde{f}^p/p + \widetilde{g}^q/q d\mu$이므로 Young의 부등식과 정리 ? ? 의 ii로부터 $\widetilde{f}\widetilde{g} = \widetilde{f}^p/p + \widetilde{g}^q/q$ (ae.)이다. 이는 다시 Young의 부등식의 등호조건으로부터 $\widetilde{f}^p = \widetilde{g}^q$ (ae.)를 뜻하고, 따라서 $||g||_q^q|f|^p - ||f||_p^p|g|^q = 0$ (ae.)이다. 역으로 모두 0은 아닌 적당한 $a, b \in \mathbb{R}$가 존재하여 $a|f|^p + b|g|^q = 0$ (ae.) 경우에 등호가 성립함은 자명하므로 증명은 이로써 충분한다. □

Theorem 1.135 (Minkowski's inequality) 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f, g : X \to \overline{\mathbb{R}}$와 $p \geq 1$에 대해 $||f + g||_p \leq ||f||_p + ||g||_p$이다. 이때, $||f||_p, ||g||_p < \infty$이면 등호가 성립한 필요충분조건은 모두 0은 아닌 적당한 $a, b \in \mathbb{R}$가 존재하여 $a|f|^p + b|g|^p = 0$ (ae.) 인 것이다.

Proof 만약 $p = 1$이면 $||f + g||_1 = \int_X |f + g| d\mu \leq \int_X |f| d\mu + \int_X |g| d\mu = ||f||_1 + ||g||_1$에서 부등식이 자명하므로 $p > 1$이라 하자. 또한, 만약 $||f||_p = \infty$이거나 $||g||_p = \infty$이거나 $||f + g||_p = 0$인 경우에도 부등식이 자명하므로 $||f||_p, ||g||_p < \infty$이고 $||f + g||_p > 0$이라 하

자. 그렇다면 $|f|,|g| \le (|f|^p+|g|^p)^{1/p}$이므로 $|f+g|^p \le (|f|+|g|)^p \le 2^p(|f|^p+|g|^p)$에서 $\int_X |f+g|^p\,d\mu \le 2^p \int_X |f|^p+|g|^p\,d\mu = 2^p(||f||_p^p+||g||_p^p) < \infty$이고, 곧 $||f+g||_p < \infty$이다. 따라서 Hölder의 부등식으로부터 $1/p+1/q=1$인 $q>1$에 대해

$$\begin{aligned}
||f+g||_p^p &= \int_X |f+g|^p\,d\mu \\
&\le \int_X |f||f+g|^{p-1}\,d\mu + \int_X |g||f+g|^{p-1}\,d\mu \\
&\le ||f||_p|||f+g|^{p-1}||_q + ||g||_p|||f+g|^{p-1}||_q \\
&= (||f||_p+||g||_p)\left(\int_X |f+g|^{(p-1)q}\,d\mu\right)^{1/q} \\
&= (||f||_p+||g||_p)\left(\int_X |f+g|^p\,d\mu\right)^{(p-1)/p} \\
&= (||f||_p+||g||_p)||f+g||_p^{p-1}
\end{aligned}$$

이 성립하여 $||f+g||_p \le ||f||_p+||g||_p$임을 안다.

다음으로, 등호조건을 보이기 위해 $||f||_p, ||g||_p < \infty$이고 $||f+g||_p = ||f||_p+||g||_p$라 하자. 만약 $f=0$ (ae.)이면 $|f|^p+0\cdot|g|^p=0$ (ae.)에서 더 이상 보일 것이 없으므로 $f=0$ (ae.)이 아니라 하고, 비슷한 이유에서 $g=0$ (ae.)도 아니라 하자. 그렇다면 위의 증명과정과 Hölder의 부등식의 등호조건으로부터 적당한 $a_1,a_2,b_1,b_2 \in \mathbb{R}$가 존재하여 $a_1|f|^p+b_2|f+g|^{(p-1)q} = a_1|f|^p+b_2|f+g|^p = 0$ (ae.)이고 $a_2|g|^p+b_2|f+g|^{(p-1)q} = a_2|g|^p+b_2|f+g|^p = 0$ (ae.)이다. 여기서 만약 $a_1=0$이면 $f+g=0$ (ae.)가 되어 $||f||_p+||g||_p = ||f+g||_p = 0$에서 $f=g=0$ (ae.)의 모순이 발생하므로 $a_1 \ne 0$이고 $b_1=0$이면 $f=0$ (ae.)의 모순이 발생하므로 $b_1 \ne 0$이다. 비슷하게 $a_2,b_2 \ne 0$이므로 이상을 종합하면 $a_1b_2|f|^p - a_2b_1|g|^p$ (ae.)이고 $a_1b_2, a_2b_1 \ne 0$이다. 역으로 모두 0은 아닌 적당한 $a,b \in \mathbb{R}$가 존재하여 $a|f|^p+b|g|^p=0$ (ae.)인 경우에 등호가 성립함은 자명하므로 증명은 이로써 충분하다. □

당장은 $p-$노름은 노름이 아니라는 점에 주의하자. 명제 ? ? 의 i과 Minkowski의 부등식으로부터 $p-$노름이 양의 동차성과 삼각부등식은 만족하지만, 아직 이가 양의 정부호성을 만족하는지는 알지 못한다. 그리고 조금만 생각해보면, $p-$노름이 양의 정부호성을 갖길 기대하는 건 부질없는 것임을 깨달을 수 있다. 적분값이 0이라고 해서 피적분함수가 반드시 0일 필요는 없기 때문이다. 그럼에도 이에 $p-$노름이라는 이름을 괜히 붙인 것은 아닐테니, $p-$노름을 정말 노름으로 만들기 위해서는 '도약' 이 필요하다.

Definition 1.136 측도공간 $(X,\mathscr{A},\mu)$에서 정의된 가측함수 $f: X \to \overline{\mathbb{R}}$와 $p \ge 1$에 대해 만약 $||f||_p < \infty$이면 이를 $L^p(\mu)$**−함수**($L^p(\mu)$**−function**) 혹은 간단히 L^p**−함수**(L^p**−function**)라 한다. 나아가, X에서 정의된 모든 L^p-함수의 모임을 $\mathscr{L}^p(\mu)$ 혹은 간단히 $\mathscr{L}^p$로 쓴다.

Definition 1.137 측도공간 $(X, \mathscr{A}, \mu)$와 $p \geq 1$에 대해 $\mathscr{L}^p(\mu)$ 위의 동치관계 $\overset{\mu}{\sim}$를 $f \overset{\mu}{\sim} g :\Leftrightarrow f = g$ (ae.)로 정의하고 이의 몫공간 $\mathscr{L}^p(\mu)/\overset{\mu}{\sim}$를 **Lebesgue 공간(− space)**, $L^p(\mu)$**−공간($L^p(\mu)$−space)** 혹은 간단히 L^p**−공간(L^p−space)**이라 하고 $L^p(\mu)$ 혹은 간단히 L^p로 쓴다.

언듯 보아서는 뭔가 이상한 짓을 한 것 같지만 이러한 L^p 공간의 구성은 굉장히 직관적인 구성이다. 앞서 보았듯이 p−노름이 가지는 문제점은 양의 정부호성이 성립하지 않는다는 것이었는데, 명제 ? ? 의 ii에 의하면 $||f||_p = 0$이면 f가 거의 어디서나 0인 것까지는 보장이 되었다. 우리는 이에 착안하여 거의 어디서나 같은 것을 정말 같은 것으로 볼 수 있는 공간을 구성한 것이다. 이제 우리의 도약은 L^p 공간에 적당한 노름공간의 구조를 부여함으로써 마무리된다.

Proposition 1.138 측도공간 $(X, \mathscr{A}, \mu)$와 $p \geq 1$에 대해 함수 $+ : (L^p)^2 \to L^p$, $\cdot : \mathbb{R} \times L^p \to L^p$를 각각 $+ : ([f], [g]) \mapsto [f+g]$, $\cdot : (\lambda, [f]) \mapsto [\lambda f]$로 정의하면 $(L^p, +, \cdot, [0], 1)$은 벡터공간이다.

Proof 정리에서 정의된 덧셈과 스칼라 곱이 well-define된다는 것만 보이면 $(L^p, +, \cdot, [0], 1)$이 벡터공간이 되기 위한 조건은 쉽게 확인할 수 있으므로 여기서는 well-definedness만 보이도록 하자. 이를 위해 임의의 $f, g \in \mathscr{L}^p$와 임의의 $\lambda \in \mathbb{R}$를 생각하면 Minkowski의 부등식으로부터 $||f+g||_p \leq ||f||_p + ||g||_p < \infty$이고 명제 ? ? 의 i로부터 $||\lambda f||_p = |\lambda| ||f||_p < \infty$이므로 $f+g, \lambda f \in \mathscr{L}^p$이다. 나아가, 주어진 덧셈과 스칼라 곱의 결과가 representative의 선택에 의존하지 않음이 자명하므로 이는 well-define된다. □

Proposition 1.139 측도공간 $(X, \mathscr{A}, \mu)$와 $p \geq 1$에 대해 함수 $||\cdot||_p : L^p \to \mathbb{R}_0^+$를 $||\cdot||_p : [f] \mapsto ||f||_p$로 정의하면 $(L^p, ||\cdot||_p)$는 노름공간이다.

Proof 우선 따름정리 ? ? 의 ii로부터 $||\cdot||_p$는 representative의 선택에 의존하지 않으므로 well-define된다. 나아가 명제 ? ? 의 i과 Minkowski의 부등식으로부터 양의 동차성과 삼각부등식이 성립하고, 명제 ? ? 의 ii와 L^p 공간의 구성으로부터 이가 양의 정부호성도 가지므로 $(L^p, ||\cdot||_p)$은 노름공간이다. □

보통 특별히 정하지 않는 이상 L^p 공간에는 p−노름이 주어진 것으로 생각하고, 이는 이 책에서도 마찬가지이다. 또한, 일반적으로 노름공간 $(L^p, ||\cdot||_p)$의 원소 $[f]$를 표기를 남용하여 그냥 f로 쓴다. 이렇게 집합인 동치류를 마치 함수처럼 표기하는 것에 처음에는 거부감을 느낄 수 있다. 하지만 L^p에서 진행되는 논의는 대부분 적분에 관한 논의이고, 곧 이러한 맥락에서는 함수의 특정 점에서의 값보다는 그 함수의 적분값이 중요한데, 거의 어디서나 같은 함수는 적어도 적분에서는 구별할 필요가 없으므로 이러한 표기법은 굉장히 편리하다. 또한, L^p 공간의 기저에 깔려있는 거의 어디서나 같은 것은 정말 같다는 그 철학을 생각한다면 오히려 이러한 표기가 자연스러운 표기법이라고 생각할 수도 있다.

다만, 이러한 표기의 남용으로 인하여 L^p에 속하는 동치류로서 $[f]$를 뜻하는 f와 $[f]$의 임의의 representative로서 $\mathscr{L}^p$에 속하는 구체적인 함수를 뜻하는 f가 적어도 표기로써는 구별되지 않으므로 이를 논의의 맥락으로써 잘 구별해야 한다. 대부분의 경우, 이러한 구별은 그다지 어렵지 않다. 한 예시로 아래의 따름정리 ? ? 의 'L^p에 속하는 함수열 $\{f_i\}$가 적당한 $f \in L^p$에 대해 $f_i \to f$ in sense of L^p이면 $\{f_i\}$의 적당한 부분함수열이 존재하여 거의 어디서나 f로 수렴한다.' 는 표현을 보자. 먼저 'L^p에 속하는 함수열 $\{f_i\}$' 에서의 f_i는 L^p에 속하는 동치류로 $[f_i]$를 이르는 것이고, 이를 표기를 남용하여 f_i라 쓰고 그에 걸맞게 '함수열' 이라 표현하였다. 이어지는 '적당한 $f \in L^p$에 대해 $f_i \to f$ in sense of L^p' 또한 동치류의 열 $\{[f_i]\}$가 $p-$노름으로써 동치류 $[f] \in L^p$로 수렴한다는 뜻으로, 'in sense of L^p' 라는 표현에서 그 의미가 분명하다. 한편, 끝부분의 '$\{f_i\}$의 적당한 부분함수열이 존재하여 거의 어디서나 f로 수렴한다' 는 $[f_i]$의 임의의 representative를 택하여 구성한 $\mathscr{L}^p$에 속하는 (진짜) 함수열 $\{f_i\}$에 대해 적당한 부분함수열 $\{f_{i_j}\}$가 존재하여 $[f]$의 임의의 representative f로 거의 어디서나 점별수렴함을 뜻하는 것으로, 이러한 거의 어디서나 점별수렴한다는 결론이 $[f_i]$나 $[f]$의 representative의 선택에 의존하지 아니하므로 그 의미의 전달이 분명하다. 이와 같이 맥락에 따라 그 의미를 파악하며 읽는 것은, 처음에는 다소 시간이 걸리겠지만, 머지않아 익숙해 질 것이다.

우리는 $p-$노름이 정말 노름이 되도록 하기 위해 L^p 공간을 도입했지만, L^p 공간은 우리가 생각하는 것보다 훨씬 더 아름다운 공간이다. 지금부터는 L^p 공간이 함수공간으로서 가지는 성질을 알아보자.

Definition 1.140 노름공간 $(\mathsf{V}, ||\cdot||)$에 대해 $||\cdot||$이 완비성을 가지면즉 임의의 Cauchy 수열이 수렴하면, 이때의 $(\mathsf{V}, ||\cdot||)$를 **Banach 공간(− space)**이라 한다.

Theorem 1.141 측도공간 $(X, \mathscr{A}, \mu)$와 $p \geq 1$에 대해 $(L^p, ||\cdot||_p)$는 Banach 공간이다.

Proof L^p 공간에 속하는 임의의 Cauchy 수열 $\{f_i\}$를 택하면 임의의 $j \in \mathbb{N}$에 대해 적당한 $i_j \in \mathbb{N}$가 존재하여 임의의 $i, i' \geq i_j$에 대해 $||f_i - f_{i'}||_p < 2^{-j}$이다. WLOG, 필요하다면 i_j를 더 크게 잡음으로써 $\{i_j\}_j$를 증가수열이라 해도 되고, 이로부터 부분열 $\{f_{i_j}\}_j$를 생각하면 이는 임의의 $j \in \mathbb{N}$에 대해 $||f_{i_j} - f_{i_{j+1}}||_p < 2^{-j}$인 함수열이다. 이제 함수 $g_j : X \to \overline{\mathbb{R}}$의 열 $\{g_j\}$를 $g_j = \sum_{k=1}^{j} |f_{i_{k+1}} - f_{i_k}|$로 정의하고 함수 $g : X \to \overline{\mathbb{R}}$를 $g = \sum_{k=1}^{\infty} |f_{i_{k+1}} - f_{i_k}| = \lim_{j\to\infty} g_j$로 두자. 그렇다면 임의의 $j \in \mathbb{N}$에 대해 Minkowski의 부등식으로부터 $||g_j||_p = ||\sum_{k=1}^{j} |f_{i_{k+1}} - f_{i_k}|||_p \leq \sum_{k=1}^{j} ||f_{i_{k+1}} - f_{i_k}||_p < \sum_{k=1}^{\infty} 2^{-k} = 1$이다. 나아가 임의의 $j \in \mathbb{N}$에 대해 $g_j^p \leq g_{j+1}^p$이고 $g_j^p \to g^p$이므로 MCT로부터 $||g||_p = (\int_X |g|^p \, d\mu)^{1/p} = \lim_{j\to\infty} (\int_X |g_j|^p \, d\mu)^{1/p} = \lim_{j\to\infty} ||g_j||_p \leq 1$에서 $g \in L^p$이다. 이는 $|g|^p$가 적분가능함을 뜻하므로 정리 ? ? 의 ii로부터 g는 거의 어디서나 유한하고, 곧 적당한 영집합 $N \subseteq X$에 대해 $X \setminus N$에서 $\sum_{j=1}^{\infty} (f_{i_{j+1}} - f_{i_j})$는 절대수렴한다. 이제 함수 $f : X \to \overline{\mathbb{R}}$를 $f = [f_{i_1} + \sum_{j=1}^{\infty} (f_{i_{j+1}} -$

$f_{i_j})]\mathbf{1}_{X\setminus N}$이라 하면 $f_{i_j}\mathbf{1}_{X\setminus N} \to f$이므로 정리 ? ? 의 ii로부터 f는 가측이다. 한편, 임의의 $\varepsilon > 0$을 택하면 적당한 $i_0 \in \mathbb{N}$가 존재하여 임의의 $i, i' \geq i_0$에 대해 $||f_i - f_{i'}||_p < \varepsilon/2$이므로 Fatou의 보조정리로부터 $i \geq i_0$이면

$$\begin{aligned}
\int_X |f_i - f|^p \, d\mu &= \int_X \lim_{j\to\infty} |f_i - f_{i_j}\mathbf{1}_{X\setminus N}|^p \, d\mu \\
&\leq \liminf_{j\to\infty} \int_X |f_i - f_{i_j}\mathbf{1}_{X\setminus N}|^p \, d\mu \\
&= \liminf_{j\to\infty} \int_X |f_i - f_{i_j}|^p \, d\mu \\
&= \liminf_{j\to\infty} ||f_i - f_{i_j}||_p^p \\
&\leq \left(\frac{\varepsilon}{2}\right)^p
\end{aligned}$$

가 되어 곧 임의의 $i \geq i_0$에 대해 $||f_i - f||_p < \varepsilon$이다. 이는 $f \in L^p$임을 뜻하기도 하므로 곧 $f_i \to f$ in sense of L^p가 되어 $(L^p, ||\cdot||_p)$가 Banach 공간임을 안다. □

Corollary 1.142 측도공간 $(X, \mathscr{A}, \mu)$와 $p \geq 1$에 대해 L^p에 속하는 함수열 $\{f_i\}$가 적당한 $f \in L^p$에 대해 $f_i \to f$ in sense of L^p이면 $\{f_i\}$의 적당한 부분함수열이 존재하여 거의 어디서나 f로 수렴한다.

PROOF 함수열 $\{f_i\}$가 수렴하므로 이는 Cauchy이고, 따라서 위의 정리의 증명에서와 같이 $\{f_{i_j}\}$를 잡으면 이는 거의 어디서나 f로 수렴한다. □

요컨대, L^p 공간은 적어도 그 구조에서는 실수체에 버금갈 정도로 좋은 공간이고, 이러한 이유로 전술하였다시피 L^p 공간 자체에 대해 많은 연구가 이루어져 있다. 한편, $p = 2$인 경우에는 여기서 한 발 더 나아갈 수 있다.

Definition 1.143 내적공간 $(\mathsf{V}, \langle\cdot,\cdot\rangle)$에 대해 $\langle\cdot,\cdot\rangle$로부터 유도되는 노름이 완비성을 가지면 이때의 $(\mathsf{V}, \langle\cdot,\cdot\rangle)$를 **Hilbert 공간(− space)**이라 한다.

Theorem 1.144 측도공간 $(X, \mathscr{A}, \mu)$에 대해 함수 $\langle\cdot,\cdot\rangle_\mu : (L^2)^2 \to \mathbb{R}$를 $\langle\cdot,\cdot\rangle_\mu : (f, g) \mapsto \int_X fg\,d\mu$로 정의하면 $(L^2, \langle\cdot,\cdot\rangle_\mu)$는 Hilbert 공간이다.

PROOF 우선 임의의 $f, g \in L^2$에 대해 Hölder의 부등식으로부터 $\int_X |fg|\,d\mu \leq ||f||_2||g||_2 < \infty$이므로 fg는 적분가능하고, 곧 $\langle\cdot,\cdot\rangle_\mu$는 well-define된다. 이제 $\langle\cdot,\cdot\rangle_\mu$가 내적임은 거의 자명하고, 이가 유도하는 노름이 2−노름이므로 정리 ? ? 로부터 $(L^2, \langle\cdot,\cdot\rangle_\mu)$는 Hilbert 공간이다. □

보통 특별히 정하지 않는 이상 L^2 공간에는 위의 정리에서와 같은 내적이 주어진 것으로 생각하고, 혼동의 우려가 없다면 위 내적의 아래첨자 μ를 생략하여 그냥 $\langle\cdot,\cdot\rangle$로 쓴다.

한편, 위의 정리로부터 $(L^2, \langle\cdot,\cdot\rangle)$이 Hilbert 공간을 이루므로, 우리는 잠시 Hilbert 공간이라는 추상의 세계로 건너가 이가 가지는 성질을 알아보고자 한다. 이번 절을 시작하며 L^p 공간은 도함수를 일반화하기 위한 경유지라고 했는데 어쩐지 점점 더 우리의 목적으로부터 멀어지는 듯하다. 그러나 대부분의 추상화가 그러하듯, 추상의 세계에서 우리는 본질을 더 잘 바라볼 수 있는 법이다. 우리가 Hilbert 공간에서 특히 관심을 가질 부분은 이에서 정의된 선형사상의 characterization이다.

Theorem 1.145 Hilbert 공간 $(\mathsf{V}, ||\cdot||)$의 공집합이 아닌 부분집합 $K \subseteq \mathsf{V}$가 닫혀있고 볼록하다면 임의의 $v \in K$에 대해 $||v^*|| \leq ||v||$인 $v^* \in K$가 유일하게 존재한다.

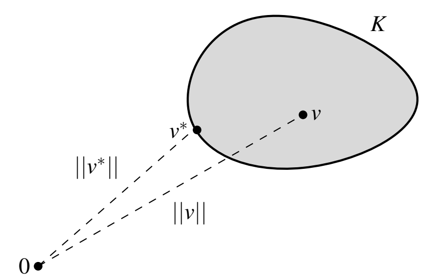


Figure 1.11 정리 ? ? 에서의 집합 K와 점 v^*.

Proof 먼저 존재성을 보이기 위해 $\delta = \inf_{v\in K} ||v||$라 하고 $||v_i|| \to \delta$인 K에 속하는 수열 $\{v_i\}$를 생각하자. 이제 임의의 $\varepsilon > 0$을 택하면 적당한 $i_0 \in \mathbb{N}$가 존재하여 임의의 $i \geq i_0$에 대해 $|||v_i||^2 - \delta^2| < \varepsilon/4$ 즉, $||v_i||^2 < \delta^2 + \varepsilon/4$이다. 한편, 임의의 $i, j \geq i_0$에 대해 K가 볼록집합이므로 $(v_i + v_j)/2 \in K$이고, 따라서 $||v_i + v_j|| \geq 2\delta$가 되어 이상으로부터 $||v_i - v_j||^2 = 2||v_i||^2 + 2||v_j||^2 - ||v_i + v_j||^2 < 4\delta^2 + \varepsilon - 4\delta^2 = \varepsilon$이므로 $\{v_i\}$는 Cauchy이다. 그런데 가정으로부터 $\{v_i\}$는 K에서 수렴하고 그 수렴값을 $v^* \in K$라 하면 $||v_i|| \to ||v^*|| = \delta$이므로 임의의 $v \in K$에 대해 $||v|| \geq ||v^*||$가 되어 존재성이 보여진다.

유일성을 보이는 것은 보다 간단하다. 만약 $||v^*|| = ||\widetilde{v}|| = \delta$인 v^*와 다른 $\widetilde{v} \in K$가 존재한다면 K가 볼록집합이므로 $(v^* + \widetilde{v})/2 \in K$이고, 곧 $||v^* + \widetilde{v}|| \geq 2\delta$이다. 그런데 이는 $||v^* - \widetilde{v}||^2 = 2||v^*||^2 + 2||\widetilde{v}||^2 - ||v^* + \widetilde{v}||^2 \leq 4\delta^2 - 4\delta^2 = 0$에서 $v^* = \widetilde{v}$의 모순을 함의하므로 유일성이 보여진다. □

마침내, 짧았던 L^p 공간과 추상세계에서의 여정에 마침표를 찍을 정리를 소개한다. Riesz의 표현정리라 불리는 다음 정리의 결론은 Hilbert 공간에서의 모든 유계인 선형사상은 그 공간의 벡터 하나로 완벽히 특정된다는 것이다. 즉, 선형사상이 유계이면 그 선형사상에 대한 정보는 점 하나에 모두 담겨있다!

Theorem 1.146 (Riesz representation theorem) Hilbert 공간 $(\mathsf{V}, ||\cdot||)$에서 정의된 선형사상 $T:\mathsf{V}\to\mathbb{R}$가 유계라 하면 임의의 $v\in\mathsf{V}$에 대해 $T(v)=\langle v, v^*\rangle$인 $v^*\in\mathsf{V}$가 유일하게 존재한다. 나아가, 이러한 v^*에 대해 $||v^*||=||T||_{\text{op}}$가 성립한다.

Proof 먼저 존재성을 보이자. 만약 $T=0$이면 $v^*=0$이 정리의 조건을 만족하고 곧 정리가 자명하므로 $T\neq 0$이라 하자. 이제 집합 $K=T^{-1}(1)$을 생각하면 가정으로부터 $T(v)\neq 0$인 $v\in\mathsf{V}$를 적어도 하나 택할 수 있고, 이러한 v에 대해 $T(v/||T(v)||)=1$에서 $v/||T(v)||\in K$이므로 K는 공집합이 아니다. 또한, T가 유계이므로 이는 연속이고, 따라서 K는 닫혀있으며, 임의의 $v, w\in K$와 임의의 $t\in[0,1]$에 대해 $T(tv+(1-t)w)=tT(v)+(1-t)T(w)=1$이므로 $tv+(1-t)w\in K$가 되어 K는 볼록집합이다. 그렇다면 정리 ? ? 로부터 임의의 $v\in K$에 대해 $||\widetilde{v}||\leq||v||$인 $\widetilde{v}\in K$가 유일하게 존재한다.

존재성에 대한 증명을 마무리하기 위해서는 임의의 $v\in\ker T$에 대해 $\langle v,\widetilde{v}\rangle=0$임을 보이는 것으로 충분하다. 만약 이를 보일 수 있으면 임의의 $v\in\mathsf{V}$에 대해 $T(v-T(v)\widetilde{v})=T(v)-T(v)=0$에서 $v-T(v)\widetilde{v}\in\ker T$이므로 곧 $0=\langle v-T(v)\widetilde{v},\widetilde{v}\rangle=\langle v,\widetilde{v}\rangle-T(v)||\widetilde{v}||^2$에서 $v^*=\widetilde{v}/||\widetilde{v}||^2$라 하면 $T(v)=\langle v,\widetilde{v}\rangle/||\widetilde{v}||^2=\langle v, v^*\rangle$가 되어 존재성이 명백하기 때문이다. 이제 임의의 $v\in\ker T$를 택하면 임의의 $t\in\mathbb{R}$에 대해 $T(\widetilde{v}+tv)=T(\widetilde{v})+tT(v)=1$이므로 $\widetilde{v}+tv\in K$에서 $||\widetilde{v}||^2\leq||\widetilde{v}+tv||^2=||\widetilde{v}||^2+2t\langle v,\widetilde{v}\rangle+t^2||v||^2$이다. 이는 곧 함수 $t\mapsto ||\widetilde{v}||^2+2t\langle v,\widetilde{v}\rangle+t^2||v||^2$가 $t=0$에서 최댓값을 가짐을 뜻하므로 그 도함수가 0을 근으로 가진다는 사실로부터 $\langle v,\widetilde{v}\rangle=0$이다.

유일성을 보이는 것은 보다 간단하다. 만약 임의의 $v\in\mathsf{V}$에 대해 $T(v)=\langle v, v^*\rangle=\langle v,\widetilde{v}\rangle$인 v^*가 아닌 $\widetilde{v}\in\mathsf{V}$가 존재한다면 $\langle v, v^*-\widetilde{v}\rangle=0$인데, 여기서 $v=v^*-\widetilde{v}$를 택하면 $||v^*-\widetilde{v}||^2=\langle v^*-\widetilde{v}, v^*-\widetilde{v}\rangle=0$이 되어 $v^*=\widetilde{v}$의 모순이 발생하므로 유일성이 보여진다.

마지막으로, $||v^*||=||T||_{\text{op}}$임을 보이자. 만약 $v^*=0$이면 명백히 $T=0$이므로 $||v^*||=||T||_{\text{op}}=0$에서 정리가 자명하므로 $v^*\neq 0$이라 하자. 그렇다면 $||v^*||=\langle v^*, v^*\rangle/||v^*||=T(v^*)/||v^*||\leq|T||_{\text{op}}$이다. 한편, Cauchy-Schwarz 부등식으로부터 임의의 $v\in\mathsf{V}$에 대해 $|T(v)|=|\langle v, v^*\rangle|\leq||v||||v^*||$이므로 $||v^*||\geq|T(v)|/||v||$에서 $||T||_{\text{op}}\leq||v^*||$이 되어 이상으로부터 $||v^*||=||T||_{\text{op}}$임을 안다. □

Corollary 1.147 측도공간 $(X,\mathscr{A},\mu)$와 선형사상 $T:L^2\to\mathbb{R}$에 대해 T가 유계라면 임의의 $g\in L^2$에 대해 $T(g)=\int_X fg\,d\mu$인 $f\in L^2$가 존재하고, 거의 어디서나 같은 함수를 하나로 본다면 유일하다. 나아가 $||T||_{\text{op}}=||f||_2$이다.

Proof 정리 ? ? 로부터 $(L^2,\langle\cdot,\cdot\rangle)$이 Hilbert 공간으므로 Riesz의 표현정리로부터 이는 거의 자명하다. □

이제 다시 측도공간으로 돌아오자. 다음의 정리나 개념들은 훨씬 전에 소개할 수도 있었지만 논의의 전개가 산만해지는 것을 피하기 위해 일부러 조금 미루어 두었던 것들이다. 먼저 σ−대수의 제한을 소개한다.

Definition 1.148 가측공간 $(X,\mathscr{A})$와 공집합이 아닌 $Y\in\mathscr{A}$에 대해 집합족 $\{A\in\mathscr{A}:A\subseteq Y\}$를 $\mathscr{A}$의 Y로의 **제한(restriction)**이라 하고 $\mathscr{A}|_Y$로 쓴다.

Proposition 1.149 가측공간 $(X,\mathscr{A})$와 집합 $Y\in\mathscr{A}$에 대해 $\mathscr{A}|_Y$는 Y 위의 $\sigma-$대수이다.

Proof 우선 $Y\in\mathscr{A}|_Y$이므로 이는 비어있지 않고 임의의 $A\in\mathscr{A}|_Y$에 대해 $Y\setminus A\in\mathscr{A}|_Y$에서 $\mathscr{A}|_Y$는 (Y를 전체 공간으로 보았을 때의) 여집합에 대해서 닫혀있다. 한편, 이가 가산 합집합에 대해 닫혀있음은 분명하므로 이상으로부터 $\mathscr{A}|_Y$는 Y 위의 $\sigma-$대수이다. $\square$

Theorem 1.150 측도공간 $(X,\mathscr{A},\mu)$와 집합 $Y\in\mathscr{A}$에 대해 $\mathscr{B}=\mathscr{A}|_Y$라 하면 다음이 성립한다.

i. 제한 $\mu|_{\mathscr{B}}$는 $\mathscr{B}$ 위의 측도이다. 따라서 $(X,\mathscr{B},\mu|_{\mathscr{B}})$는 측도공간을 이룬다.
ii. 가측공간 $(Z,\mathscr{C})$와 $\mathscr{A}/\mathscr{C}-$가측함수 $f:X\to Z$에 대해 $f|_Y$는 $\mathscr{B}/\mathscr{C}-$가측이다.
iii. 음이 아닌 $\mathscr{A}/\mathscr{B}_1-$가측함수 $f:X\to\overline{\mathbb{R}}_0^+$에 대해 $\int_Y f\,d\mu=\int_Y f|_Y\,d\mu|_{\mathscr{B}}$이다. 한편, Y에서 $\mu-$적분가능한(혹은 $f|_Y$가 $\mu|_{\mathscr{B}}-$적분가능한 $\mathscr{A}/\mathscr{B}_1-$가측함수) $f:X\to\overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며, 이때 $f|_Y$는 $\mu|_{\mathscr{B}}-$적분가능하다(혹은 f는 Y에서 $\mu-$적분가능하다).

Proof i. 이는 명제 ? ? 로부터 $\mathscr{B}$가 Y 위의 $\sigma-$대수라는 점에서 자명하다.

ii. 이 또한 거의 자명하다. 임의의 $A\in\mathscr{C}$에 대해 $(f|_Y)^{-1}(A)=f^{-1}(A)\cap Y\in\mathscr{B}$이므로 $f|_Y$가 $\mathscr{B}/\mathscr{C}-$가측이다.

iii. 먼저 f가 음이 아닌 단순함수인 경우를 생각하여 이의 표준형을 $f=\sum_{i=1}^k a_i\mathbf{1}_{A_i}$라 하면 $f|_Y=\sum_{i=1}^k a_i\mathbf{1}_{A_i\cap Y}$이고, 이로부터 $\int_Y f|_Y\,d\mu|_{\mathscr{B}}=\sum_{i=1}^k a_i\mu|_{\mathscr{B}}(A_i\cap Y)=\sum_{i=1}^k a_i\mu(A_i\cap Y)=\int_X f\mathbf{1}_Y\,d\mu=\int_Y f\,d\mu$이다. 이제 f를 음이 아닌 가측함수라 하면 정리 ? ? 로부터 음이 아닌 단순함수 $f_i:X\to\mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i\uparrow f$이므로 MCT에서 $\int_Y f|_Y\,d\mu|_{\mathscr{B}}=\lim_{i\to\infty}\int_Y f_i|_Y\,d\mu|_{\mathscr{B}}=\lim_{i\to\infty}\int_Y f_i\,d\mu=\int_Y f\,d\mu$이다.

한편, Y에서 $\mu-$적분가능한 $f:X\to\overline{\mathbb{R}}$에 대해서, 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 $f|_Y$가 $\mu|_{\mathscr{B}}-$적분가능하며 $\int_Y f|_Y\,d\mu|_{\mathscr{B}}=\int_Y f\,d\mu$임을 쉽게 알 수 있고, $f|_Y$가 $\mu|_{\mathscr{B}}-$적분가능한 $\mathscr{A}/\mathscr{B}_1-$가측함수 $f:X\to\overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. $\square$

위의 정리는, 비록 그 표현은 복잡해 보이지만, 본질적으로는 피적분함수와 적분에 사용되는 측도의 적분영역 밖의 정보는 그 적분에 있어 아무런 필요가 없다는 지극히 상식적인 아이디어를 확인한 것에 불과하다.

다음으로, 측도의 선형결합에 대한 내용으로 넘어가자. 이 또한 매우 직관적인 결과들이다.

Theorem 1.151 가측공간 $(X,\mathscr{A})$ 위의 측도 μ,ν에 대해 다음이 성립한다.

i. 임의의 $c \geq 0$에 대해 함수 $c\mu$도 측도이다.
ii. 함수 $\mu + \nu$도 측도이다.

Proof 자명하다. i의 $c\mu$와 ii의 $\mu + \nu$가 측도의 조건들을 만족함을 쉽게 확인할 수 있다. □

Theorem 1.152 가측공간 $(X, \mathscr{A})$ 위의 측도 μ, ν와 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 다음의 성립한다.

i. 임의의 $c \geq 0$에 대해 $\int_X f\, d(c\mu) = c \int_X f\, d\mu$이다.
ii. $\int_X f\, d(\mu + \nu) = \int_X f\, d\mu + \int_X f\, d\nu$이다.

한편, μ−적분가능한 $f : X \to \overline{\mathbb{R}}$에 대해서도 i의 성질이 성립하고, 이때 f는 $c\mu$−적분가능하다. 비슷하게, μ−적분가능한 동시에 ν−적분가능한(혹은 $(\mu + \nu)$−적분가능한) $f : X \to \overline{\mathbb{R}}$에 대해서도 ii의 성질이 성립하고, 이때 f는 $(\mu + \nu)$−적분가능하다(혹은 μ−적분가능한 동시에 ν−적분가능하다).

Proof 편의를 위해 i만 증명하도록 한다. (ii는 i과 비슷하게 하면 된다.) 먼저 f가 음이 아닌 단순함수인 경우를 생각하여 이의 표준형을 $f = \sum_{i=1}^{k} a_i \mathbf{1}_{A_i}$라 하면 $\int_X f\, d(c\mu) = \sum_{i=1}^{k} a_i (c\mu)(A_i) = c \sum_{i=1}^{k} a_i \mu(A_i) = c \int_X f\, d\mu$가 성립한다. 이제 f를 음이 아닌 가측함수라 하면 정리 ? ? 로부터 음이 아닌 단순함수 $f_i : X \to \mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i \uparrow f$이므로 MCT에서 $\int_X f\, d(c\mu) = \lim_{i\to\infty} \int_X f_i\, d(c\mu) = c \lim_{i\to\infty} \int_X f_i\, d\mu = c \int_X f\, d\mu$임을 안다.

한편, μ−적분가능한 $f : X \to \overline{\mathbb{R}}$에 대해서, 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 f가 $c\mu$−적분가능하며 $\int_X f\, d(c\mu) = c \int_X f\, d\mu$임을 쉽게 알 수 있다. □

이로써 도함수를 일반화하기 위한 준비가 모두 끝났다. 아마 아직은 감이 오지 않겠지만, 우리가 L^p 공간과 Hilbert 공간을 오가며 얻은 결과들과 뜬금없이 소개한 σ−대수의 제한과 같은 내용들이 마법과 같이 하나로 합쳐져 우리가 원하는 도함수의 일반화로 이어질 것이다.

Definition 1.153 가측공간 $(X, \mathscr{A})$ 위의 측도 μ, ν를 생각하자. 만약 $\nu(N) = 0$인 모든 $N \in \mathscr{A}$에 대해 $\mu(N) = 0$이면 이때 μ는 ν에 대해 **절대연속(absolutely continuous)**이라 하고 $\mu \ll \nu$로 쓴다.

Lemma 1.154 유한 측도공간 $(X, \mathscr{A}, \mu)$에 대해 $L^2 \subseteq L^1$이다.

Proof 임의의 $f \in L^2$에 대해 Cauchy-Schwarz 부등식으로부터 $\int_X |f|\, d\mu = \langle |f|, 1 \rangle \leq ||f||_2 ||1||_2 = \sqrt{(\int_X f^2\, d\mu)(\int_X d\mu)} = \sqrt{\mu(X)}||f||_2 < \infty$이고, 곧 $f \in L^1$이므로 $L^2 \subseteq L^1$임을 안다. □

Theorem 1.155 (Radon-Nikodým) 가측공간 $(X, \mathscr{A})$ 위의 측도 μ와 σ−유한 측도 ν에 대해 TFAE.

i. 측도 μ는 σ−유한하고 $\mu \ll \nu$이다.
ii. 음이 아닌 가측함수 $f : X \to \mathbb{R}_0^+$가 존재하여 $\mu = \nu_f$이다.

특히, 만약 i이 성립한다면 ii에서의 가측함수 f는 ν−거의 어디서나 같은 함수를 하나로 본다면 유일하다. 나아가, 만약 μ가 유한하다면 f는 ν−적분가능하다.

Proof ii $\Rightarrow$ i. 우선 임의의 $N \in \mathscr{A}$에 대해 $\nu(N) = 0$이라 하면 보조정리 ? ? 의 i로부터 $\mu(N) = \nu_f(N) = 0$이 되어 $\mu \ll \nu$임은 분명하므로 μ가 σ−유한함만 보이면 된다. 가정으로부터 ν가 σ−유한하므로 $\mathscr{A}$에 속하는 적당한 $\{A_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $\nu(A_i) < \infty$이고 $\bigcup_{i=1}^{\infty} A_i = X$이며, WLOG, 필요하다면 각 항을 $C_i := \bigcup_{j=1}^{i} A_j$로 바꾸어 $\{A_i\}$가 처음부터 증가하는 집합열이라 해도 된다. 이제 집합열 $\{B_i\}$를 $B_i = f^{-1}((-\infty, i]) \cap A_i$라 하면 이는 명백히 증가하는 집합열로 각 $i \in \mathbb{N}$에 대해 $\mu(B_i) = \int_{B_i} f\,d\nu \leq \int_{A_i} i\,d\nu = i\nu(A_i) < \infty$이고 $\bigcup_{i=1}^{\infty} B_i = X$임이 분명하므로 μ도 σ−유한하다.

i $\Rightarrow$ ii. 우선 ν, μ가 모두 유한한 특별한 경우를 생각하자. 이제 함수 $T : L^2(\nu+\mu) \to \mathbb{R}$를 $T : g \mapsto \int_X g\,d\mu$로 두면 ν, μ가 유한하므로 $\nu+\mu$도 유한하여 위의 보조정리와 정리 ? ? 의 ii로부터 $L^2(\nu+\mu) \subseteq L^1(\nu+\mu) \subseteq L^1(\mu)$이고, 곧 T는 well-define된다. 또한, T가 선형사상임이 분명하고 임의의 $g \in L^2(\nu+\mu)$에 대해 Cauchy-Schwarz 부등식으로부터

$$
\begin{aligned}
|T(g)| &= \left| \int_X g\,d\mu \right| \\
&\leq \int_X |g|\,d\mu \\
&\leq \int_X |g|\,d\nu + \int_X |g|\,d\mu \\
&= \int_X |g|\,d(\nu+\mu) \\
&= \langle |g|, 1 \rangle_{\nu+\mu} \\
&\leq ||g||_{2,\nu+\mu} ||1||_{2,\nu+\mu} \\
&= \sqrt{\left[\int_X g^2\,d(\nu+\mu)\right]\left[\int_X d(\nu+\mu)\right]} \\
&= \sqrt{(\nu+\mu)(X)}\,||g||_{2,\nu+\mu}
\end{aligned}
$$

이므로 T는 유계인 선형사상이 되어 Riesz의 표현정리로부터 임의의 $g \in L^2(\nu+\mu)$에 대해 $\int_X g\,d\mu = T(g) = \int_X gh\,d(\nu+\mu)$인 $h \in L^2(\nu+\mu)$가 존재한다. 나아가 임의의 $g \in L^2(\nu+\mu)$에 대해 $\int_X g(1-h)\,d(\nu+\mu) = \int_X g\,d(\nu+\mu) - \int_X gh\,d(\nu+\mu) = \int_X g\,d(\nu+\mu) - \int_X g\,d\mu = \int_X g\,d\nu$도 성립한다.

더 진행하기 전에 $0 \leq h < 1$ $((\nu+\mu)$−ae.$)$임을 확인하고 가자. 이를 위해 집합 $A = h^{-1}((-\infty, 0)), B = h^{-1}([1, \infty))$를 생각하면 $\mu(A) = \int_X \mathbf{1}_A\,d\mu = \int_X h\mathbf{1}_A\,d(\nu+\mu) \leq 0$이므

로 $\int_X h\mathbf{1}_A\,d(\nu+\mu)=0$이고 곧 정리 ? ? 의 ii로부터 $h\mathbf{1}_A=0$ ($(\nu+\mu)$–ae.)가 되어 $(\nu+\mu)(A)=(\nu+\mu)((h\mathbf{1}_A)^{-1}((-\infty,0)))=0$이다. 비슷하게, $\nu(B)=\int_X \mathbf{1}_B\,d\nu=\int_X(1-h)\mathbf{1}_B\,d(\nu+\mu)\le 0$이므로 $\nu(B)=0$이고, $\mu\ll\nu$라는 가정으로부터 $\mu(B)=0$이 되어 $(\nu+\mu)(B)=0$이다. 이상으로부터 A,B가 모두 $(\nu+\mu)$–영집합이 되어 $0\le h<1$ ($(\nu+\mu)$–ae.)이다.

WLOG, 필요하다면 $(\nu+\mu)$–영집합에서의 h의 값을 0으로 바꾸어 그냥 $0\le h<1$이라 해도 된다. 한편, 임의의 $A\in\mathscr{A}$에 대해 $\nu(A)=\int_X \mathbf{1}_A\,d\nu=\int_X(1-h)\mathbf{1}_A\,d(\nu+\mu)=\int_A 1-h\,d(\nu+\mu)=(\nu+\mu)_{1-h}(A)$이므로 임의의 음이 아닌 가측함수 $g:X\to\overline{\mathbb{R}}_0^+$에 대해 $\int_X g\,d\nu=\int_X g\,d(\nu+\mu)_{1-h}=\int_X g(1-h)\,d(\nu+\mu)$이다. 따라서 함수 $f:X\to\overline{\mathbb{R}}_0^+$를 $f=h/(1-h)$로 두면 이는 가측이고 임의의 $A\in\mathscr{A}$에 대해 $\mu(A)=\int_X \mathbf{1}_A\,d\mu=\int_X h\mathbf{1}_A\,d(\nu+\mu)=\int_X f(1-h)\mathbf{1}_A\,d(\nu+\mu)=\int_X f\mathbf{1}_A\,d\nu=\int_A f\,d\nu$가 되어 $\mu=\nu_f$에서 정리가 성립한다.

이제 σ–유한한 ν,μ를 생각하면 정의로부터 $\mathscr{A}$에 속하는 적당한 $\{A_i\},\{B_i\}$가 존재하여 $\bigcup_{i=1}^\infty A_i=\bigcup_{i=1}^\infty B_i=X$이고 각 $i\in\mathbb{N}$에 대해 $\nu(A_i),\mu(B_i)<\infty$이다. 또한, 앞서 ii가 i을 함의함을 증명할 때와 비슷하게 하여 WLOG, $\{A_i\},\{B_i\}$가 증가한다고 해도 된다. 이제 집합열 $\{Y_i\}$를 $Y_i=A_i\cap B_i$로 두면 이는 증가하는 집합열로서 $\bigcup_{i=1}^\infty Y_i=\bigcup_{i=1}^\infty A_i\cap\bigcup_{i=1}^\infty B_i=X$이고 각 $i\in\mathbb{N}$에 대해 $\nu(Y_i)\le\nu(A_i)<\infty$이며 $\mu(Y_i)\le\mu(B_i)<\infty$이다. 또한, 표기의 편의를 위해 각 $i\in\mathbb{N}$에 대해 $\mathscr{B}_i=\mathscr{A}|_{Y_i},\nu_i=\nu|_{\mathscr{B}_i},\mu_i=\mu|_{\mathscr{B}_i}$로 두면 ν_i,μ_i는 유한하며 $\mathscr{B}_{i+1}|_{Y_i}=\mathscr{B}_i,\nu_{i+1}|_{\mathscr{B}_i}=\nu_i,\mu_{i+1}|_{\mathscr{B}_i}=\mu_i$이다. 그렇다면 각 $i\in\mathbb{N}$에 대해 $\mu_i\ll\nu_i$임이 분명하므로 앞선 결과로부터 적당한 음이 아닌 가측함수 $f_i:Y_i\to\mathbb{R}_0^+$가 존재하여 $\mu_i=(\nu_i)_{f_i}$이고, 각 $i\in\mathbb{N}$와 임의의 $A\in\mathscr{B}_i$에 대해 정리 ? ? 로부터 $\int_A f_i\,d\nu_i=(\nu_i)_{f_i}(A)=\mu_i(A)=\mu_{i+1}(A)=(\nu_{i+1})_{f_{i+1}}(A)=\int_A f_{i+1}\,d\nu_{i+1}=\int_A f_{i+1}|_{Y_i}\,d\nu_i$이다. 이는 곧 정리 ? ? 로부터 각 $i\in\mathbb{N}$에 대해 $f_i=f_{i+1}|_{Y_i}$ (ν_i–ae.)임을 함의하므로 WLOG, 필요하다면 ν_i–영집합에서의 f_{i+1}의 값을 조금씩 바꾸어 그냥 $f_i=f_{i+1}|_{Y_i}$라 해도 되고, 이로부터 함수 $f:X\to\mathbb{R}_0^+$를 임의의 $i\in\mathbb{N}$에 대해 $f|_{Y_i}=f_i$가 되도록 정의할 수 있다. 그렇다면 임의의 $x\in\mathbb{R}$에 대해 $f^{-1}((-\infty,x])=\bigcup_{i=1}^\infty[f^{-1}((-\infty,x])\cap Y_i]=\bigcup_{i=1}^\infty(f_i)^{-1}((-\infty,x])$이므로 f는 가측이다. 이제 임의의 $C\in\mathscr{A}$에 대해 집합열 $\{C_i\}$를 $C_i=C\cap Y_i$로 두면 $C_i\uparrow C$이므로 정리 ? ? 의 i과 정리 ? ? 의 i로부터

$$\begin{aligned}
\mu(C)&=\lim_{i\to\infty}\mu(C_i)\\
&=\lim_{i\to\infty}\mu_i(C_i)\\
&=\lim_{i\to\infty}(\nu_i)_{f_i}(C_i)\\
&=\lim_{i\to\infty}\int_{C_i} f_i\,d\nu_i\\
&=\lim_{i\to\infty}\int_{C_i} f\,d\nu
\end{aligned}$$

$$= \int_C f \, d\nu$$
$$= \nu_f(C)$$

가 되어 $\mu = \nu_f$이다.

한편, 정리 ? ? 로부터 이러한 f가 $\nu-$거의 어디서나 같은 함수를 하나로 본다면 유일함이 분명하고, 만약 μ가 유한하다면 $\int_X f \, d\nu = \nu_f(X) = \mu(X) < \infty$에서 f가 $\nu-$적분가능하므로 증명이 끝난다. □

Definition 1.156 가측공간 $(X, \mathscr{A})$ 위의 $\sigma-$유한 측도 μ, ν에 대해 $\mu \ll \nu$이면 이때 $\mu = \nu_f$를 만족하는 함수 $f : X \to \mathbb{R}_0^+$를 μ의 ν에 대한 **Radon-Nikodým 도함수(− derivative)**라 하고 $d\mu/d\nu$로 쓴다.

치환적분법을 떠오르게 하는 다음 따름정리를 보면 방금 우리가 얻은 정의를 왜 도함수의 일반화로 볼 수 있는지가 더욱 분명해 질 것이다.

Corollary 1.157 가측공간 $(X, \mathscr{A})$ 위의 $\sigma-$유한 측도 ν, μ에 대해 $\mu \ll \nu$라 하자. 그렇다면 임의의 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 $\int_X f \, d\mu = \int_X f(d\mu/d\nu) \, d\nu$이다. 한편, $\mu-$적분가능한(혹은 $f(d\mu/d\nu)$가 $\nu-$적분가능한 가측함수) $f : X \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며 이때 $f(d\mu/d\nu)$는 $\nu-$적분가능하다(혹은 f는 $\mu-$적분가능하다).

Proof 이는 정리 ? ? 로부터 자명하다. □

1.11 Fubini's Theorem

이번 절에서는 그 제목에서와 같이 다중적분을 반복적분으로써 계산할 수 있도록 해 주는 Fubini의 정리를 증명한다. 사실 이전 절에서 보인 바와 같이 Lebesgue 적분은 Riemann 적분을 포괄하는 개념이므로 실제로 그 적분값을 구해야 할 필요가 있을 때에는 이를 Riemann 적분으로 생각하여 계산하면 되고, 미적분학이나 해석학에서 Riemann 적분에서의 Fubini의 정리를 이미 배웠으므로 이번 절에서의 논의가 없더라도 계산에 있어 Fubini의 정리를 사용할 수 있다. 그럼에도 불구하고, 하나의 절을 할애하여 이에 관한 논의를 진행하는 것에는 나름의 이유가 있다.

먼저, 종전에 우리가 배운 Riemann 적분에서의 Fubini의 정리는 피적분함수에 대해, 예컨대 이가 미분가능해야 한다거나, 혹은 조각적 연속이어야 한다거나 하는 다소 강한 조건을 요구한다. 또한, 어디까지나 Lebesgue 적분이 Riemann 적분을 포괄하는 것이지 이 두 개념이 서로 동등한 것이 아니므로 Lebesgue 적분을 언제나 Riemann 적분으로 생각하여 계

산할 수 있는 것은 아니다. 예컨대 Riemann 적분은 불가능하지만 Lebesgue 적분은 가능한 대표적인 함수인 $\mathbf{1}_{\mathbb{Q}\cap[0,1]}$의 적분을 계산해야 하는 경우, 이는 Lebesgue 적분으로만 계산할 수 있다.[19]비록 이러한 경우가 이 책에서는 (아마) 없겠지만, 그래도 이론적인 측면에서 조금 불안한 것이 사실이다. 이에 이번 절에서 Fubini의 정리를 Lebesgue 적분론의 틀 안에서 최대한 일반적으로 다시 증명함으로써 이런 이론적인 불안정함을 해소하고자 한다.

우선 가장 기본적인 형태로 product measure space에서 정의된 음이 아닌 가측함수에 대한 Fubini의 정리를 보자.

Lemma 1.158 $\sigma-$유한 측도공간 $(X,\mathscr{A},\mu),(Y,\mathscr{B},\nu)$와 음이 아닌 $(\mathscr{A}\otimes\mathscr{B})/\mathscr{B}_1-$가측함수 $f:X\times Y\to\overline{\mathbb{R}}_0^+$에 대해 함수 $\varphi:X\to\overline{\mathbb{R}}_0^+,\ \psi:Y\to\overline{\mathbb{R}}_0^+$를 각각 $\varphi:x\mapsto\int_Y f_x\,d\nu,\ \psi:y\mapsto\int_X f^y\,d\mu$라 하면 이는 well-define되며 각각 $\mathscr{A}/\mathscr{B}_1-$가측이고 $\mathscr{B}/\mathscr{B}_1-$가측이다.

Proof 우선 정리 ? ? 로부터 모든 $x\in X$에 대해 f_x가 $\mathscr{B}/\mathscr{B}_1-$가측이므로 함수 φ는 well-define된다. 이제 f가 음이 아닌 단순함수인 특별한 경우를 생각하여 이의 표준형을 $f=\sum_{i=1}^k a_i\mathbf{1}_{A_i}$라 하자. 그렇다면 $\varphi(x)=\int_Y(\sum_{i=1}^k a_i\mathbf{1}_{A_i})_x\,d\nu=\sum_{i=1}^k a_i\int_Y\mathbf{1}_{(A_i)_x}\,d\nu=\sum_{i=1}^k a_i\nu((A_i)_x)$이므로 보조정리 ? ? 로부터 φ가 $\mathscr{A}/\mathscr{B}_1-$가측이다. 한편, 음이 아닌 가측함수 f의 경우에는 정리 ? ? 로부터 음이 아닌 단순함수 $f_i:X\times Y\to\mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i\uparrow f$이고, 각 $i\in\mathbb{N}$에 대해 함수 $\varphi_i:X\to\overline{\mathbb{R}}_0^+$를 $\varphi_i:x\mapsto\int_Y(f_i)_x\,d\nu$라 하면 앞선 결과로부터 이는 $\mathscr{A}/\mathscr{B}_1-$가측이다. 그렇다면 MCT로부터 $\lim_{i\to\infty}\varphi_i(x)=\int_Y\lim_{i\to\infty}(f_i)_x\,d\nu=\int_Y f_x\,d\nu=\varphi(x)$이므로 정리 ? ? 의 ii로부터 φ는 $\mathscr{A}/\mathscr{B}_1-$가측이다. 이제 ψ에 대해서도 이와 비슷하게 하면 된다. □

Theorem 1.159 (Fubini) $\sigma-$유한 측도공간 $(X,\mathscr{A},\mu),(Y,\mathscr{B},\nu)$와 $(\mathscr{A}\otimes\mathscr{B})/\mathscr{B}_1-$가측함수 $f:X\times Y\to\overline{\mathbb{R}}_0^+$에 대해

$$\int_{X\times Y}f(x,y)\,d(\mu\otimes\nu)(x,y)=\int_X\left[\int_Y f(x,y)\,d\nu(y)\right]d\mu(x)=\int_Y\left[\int_X f(x,y)\,d\mu(x)\right]d\nu(y)$$

가 성립한다.

Proof 먼저 위의 보조정리로부터 이중적분들이 well-define됨을 안다. 이제 함수 $\varphi:X\to\overline{\mathbb{R}}_0^+$를 위의 보조정리에서와 같이 두고 f가 음이 아닌 단순함수인 특별한 경우를 생각하여 이의 표준형을 $f=\sum_{i=1}^k a_i\mathbf{1}_{A_i}$라 하면 위의 보조정리의 증명에서 $\varphi(x)=\sum_{i=1}^k a_i\nu((A_i)_x)$이므로 정리 ? ? 로부터 $\int_X\varphi\,d\mu=\sum_{i=1}^k a_i\int_X\nu((A_i)_x)\,d\mu(x)=\sum_{i=1}^k a_i(\mu\otimes\nu)(A_i)=\int_{X\times Y}f\,d(\mu\otimes\nu)$이다. 한편, 음이 아닌 가측함수 f의 경우에는 정리 ? ? 로부터 음이 아닌 단순함수 $f_i:X\times Y\to\mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i\uparrow f$이고 함수열 $\{\varphi_i\}$를 위의 보조정리의 증명에서와 같이 두면 $\varphi_i\uparrow\varphi$이므로 MCT와 앞선 결론으로부터 $\int_X\varphi\,d\mu=\lim_{i\to\infty}\int_X\varphi_i\,d\mu=\lim_{i\to\infty}\int_{X\times Y}f_i\,d(\mu\otimes\nu)=\int_{X\times Y}f\,d(\mu\otimes\nu)$이다. 이제, 두 번째 등식에 대해서도 이와 비슷하게 하면 된다. □

다음으로, product measure space에서 정의된 적분가능한 함수에 대한 Fubini의 정리를 볼텐데, 이 경우에는 피적분함수의 section이 항상 적분가능하지 않으므로 조금 더 복잡하다.

Lemma 1.160 σ−유한 측도공간 $(X, \mathscr{A}, \mu)$, $(Y, \mathscr{B}, \nu)$와 $(\mu \otimes \nu)$−적분가능한 함수 $f : X \times Y \to \overline{\mathbb{R}}$에 대해 다음이 성립한다.

i. Section $f_x : Y \to \overline{\mathbb{R}}$는 μ−거의 대부분의 $x \in X$에 대해 ν−적분가능하며 함수 $\varphi : X \to \mathbb{R}$를

$$\varphi : x \mapsto \begin{cases} \displaystyle\int_Y f_x \, d\nu & f_x\text{가 } \nu\text{−적분가능한 경우} \\ 0 & \text{ow.} \end{cases}$$

라 하면 이는 μ−적분가능하다.

ii. Section $f^y : X \to \overline{\mathbb{R}}$는 ν−거의 대부분의 $y \in Y$에 대해 μ−적분가능하며 함수 $\psi : Y \to \mathbb{R}$를

$$\psi : x \mapsto \begin{cases} \displaystyle\int_X f^y \, d\mu & f^y\text{가 } \mu\text{−적분가능한 경우} \\ 0 & \text{ow.} \end{cases}$$

라 하면 이는 ν−적분가능하다.

Proof 간결한 논의를 위해 i만 보이도록 하자. (ii는 i과 비슷하게 하면 된다.) 함수 $f_\pm$가 $(\mathscr{A} \otimes \mathscr{B})/\mathscr{B}_1$−가측이므로 함수 $\widetilde{\varphi}^\pm : X \to \overline{\mathbb{R}}_0^+$를 $\widetilde{\varphi}^\pm : x \mapsto \int_Y (f_\pm)_x \, d\nu$로 두면 (여기서 위첨자로 쓰인 $\pm$은 정의 ? ? 에서 아래첨자로 사용된 $\pm$와는 달리 함수의 양과 음의 부분을 의미하는 것이 아니다.) 이는 보조정리 ? ? 로부터 $\mathscr{A}/\mathscr{B}_1$−가측이고, 곧 Fubini의 정리로부터 $\int_X \widetilde{\varphi}^\pm \, d\mu = \int_{X \times Y} f_\pm \, d(\mu \otimes \nu) \le \int_{X \times Y} |f| \, d(\mu \otimes \nu) < \infty$가 되어 $\widetilde{\varphi}^\pm$이 μ−적분가능함을 안다. 그렇다면 $\widetilde{\varphi}^\pm$이 μ−거의 어디서나 유한하므로 명백히 $\mathscr{A}$−가측인 집합 $N = (\widetilde{\varphi}^\pm)^{-1}(\infty) = \{x \in X : \int_Y |f_x| \, d\nu = \infty\}$에 대해 $\mu(N) = 0$이고, 곧 f_x는 μ−거의 대부분의 $x \in X$에 대해 ν−적분가능하다. 이제 함수 $\varphi^\pm : X \to \mathbb{R}_0^+$를 $\varphi^\pm = \widetilde{\varphi}^\pm \mathbf{1}_{X \setminus N}$로 두면 $\varphi^\pm = \widetilde{\varphi}^\pm$ (μ−ae.)이므로 따름정리 ? ? 의 ii로부터 $\varphi^\pm$은 μ−적분가능하고, 임의의 $x \in X$에 대해 $(f_x)_\pm = (f_\pm)_x$가 성립하여 $\varphi = \varphi^+ - \varphi^-$이므로 증명이 끝난다. □

Theorem 1.161 (Fubini) σ−유한 측도공간 $(X, \mathscr{A}, \mu)$, $(Y, \mathscr{B}, \nu)$와 $(\mu \otimes \nu)$−적분가능한 함수 $f : X \times Y \to \overline{\mathbb{R}}$에 대해

$$\int_{X \times Y} f(x, y) \, d(\mu \otimes \nu)(x, y) = \int_X \left[\int_Y f(x, y) \, d\nu(y) \right] d\mu(x) = \int_Y \left[\int_X f(x, y) \, d\mu(x) \right] d\nu(y)$$

가 성립한다. 단, 만약 고정된 $x \in X$에 대해 $f(x,y)$가 ν–적분가능하지 않다면 이때의 $\int_Y f(x,y)\,d\nu(y)$의 값은 0으로 간주하고, 이는 고정된 $y \in Y$에 대해 $f(x,y)$가 μ–적분가능하지 않은 경우에도 마찬가지이다.

Proof 먼저 위의 보조정리로부터 이중적분들이 well-define됨을 안다. 이제 함수 $\varphi : X \to \mathbb{R}$와 $\varphi^{\pm} : X \to \mathbb{R}_0^+$를 각각 위의 보조정리와 그 증명에서와 같이 두면 $\varphi = \varphi^+ - \varphi^-$에서 $\int_X \varphi\,d\mu = \int_X \varphi^+\,d\mu - \int_X \varphi^-\,d\mu = \int_{X\times Y} f_+\,d(\mu\otimes\nu) - \int_{X\times Y} f_-\,d(\mu\otimes\nu) = \int_{X\times Y} f\,d(\mu\otimes\nu)$임을 쉽게 알 수 있다. 이제 두 번째 등식에 대해서도 이와 비슷하게 하면 된다. □

비록 위의 정리에서는 적분불가능한 section에 대해 그 적분값을 0이라는 dummy value로 두었지만, 보조정리 ? ? 로부터 피적분함수의 거의 대부분의 section이 적분가능하므로 전체 적분의 값은 dummy value의 선택과는 무관하다. 그리고 이로부터 위의 반복적분 표기는 나름 make sense한다. 이러한 사실은 뒤따르는 나머지 Fubini의 정리에 대해서도 마찬가지이다.

이제 중간 결과로서 Borel 측도공간에서의 Fubini의 정리를 얻는다.

Corollary 1.162 $l+m=n$인 $l,m,n \in \mathbb{N}$과 음이 아닌 Borel 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$에 대해

$$\int_{\mathbb{R}^n} f(x,y)\,d\mu_n(x,y) = \int_{\mathbb{R}^l}\left[\int_{\mathbb{R}^m} f(x,y)\,d\mu_m(y)\right]d\mu_l(x) = \int_{\mathbb{R}^m}\left[\int_{\mathbb{R}^l} f(x,y)\,d\mu_l(x)\right]d\mu_m(y)$$

가 성립한다. 한편, 적분가능한 Borel 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립한다. 단, 이 경우에는 만약 고정된 $x \in \mathbb{R}^l$에 대해 $f(x,y)$가 적분가능하지 않다면 이때의 $\int_{\mathbb{R}^m} f(x,y)\,d\mu_m(y)$의 값은 0으로 간주하고, 이는 고정된 $y \in \mathbb{R}^m$에 대해 $f(x,y)$가 적분가능하지 않은 경우에도 마찬가지이다.

Proof 이는 Fubini의 정리와 정리 ? ? 의 i로부터 자명하다. □

한편, Lebesgue 측도공간은 Borel 측도공간과는 달리 product measure space의 완비화로써 구성되므로 조금 더 복잡한 Fubini의 정리가 필요하다. 논의의 전체적인 전개 방향은 이전과 비슷하다.

Lemma 1.163 σ–유한 완비측도공간 $(X,\mathscr{A},\mu), (Y,\mathscr{B},\nu)$와 음이 아닌 $\overline{\mathscr{A}\otimes\mathscr{B}}/\mathscr{B}_1$–가측 함수 $f : X\times Y \to \overline{\mathbb{R}}_0^+$에 대해 다음이 성립한다.

i. Section $f_x : Y \to \overline{\mathbb{R}}$는 μ–거의 대부분의 $x \in X$에 대해 $\mathscr{B}/\mathscr{B}_1$–가측이며 함수 $\varphi : X \to \overline{\mathbb{R}}_0^+$를

$$\varphi : x \mapsto \begin{cases} \displaystyle\int_Y f_x\,d\nu & f_x\text{가 } \mathscr{B}/\mathscr{B}_1\text{–가측인 경우} \\ 0 & \text{ow.} \end{cases}$$

라 하면 이는 $\mathscr{A}/\mathscr{B}_1$−가측이다.

ii. Section $f^y : X \to \overline{\mathbb{R}}$는 ν−거의 대부분의 $y \in Y$에 대해 $\mathscr{A}/\mathscr{B}_1$−가측이며 함수 $\psi : Y \to \overline{\mathbb{R}}_0^+$를

$$\psi : x \mapsto \begin{cases} \int_X f^y \, d\mu & f^y\text{가 } \mathscr{A}/\mathscr{B}_1\text{−가측인 경우} \\ 0 & \text{ow.} \end{cases}$$

라 하면 이는 $\mathscr{B}/\mathscr{B}_1$−가측이다.

Proof 간결한 논의를 위해 i만 보이도록 하자. (ii는 i과 비슷하게 하면 된다.) 정리 ? ? 로부터 적당한 $(\mathscr{A} \otimes \mathscr{B})/\mathscr{B}_1$−가측함수 $g : X \times Y \to \overline{\mathbb{R}}$가 존재하여 $f = g$ $((\mu \otimes \nu)\text{−ae.})$이고, WLOG, 필요하다면 $(\mu \otimes \nu)$−영집합에서의 g의 값을 0으로 바꾸어 g를 음이 아니라 해도 된다. 그렇다면 집합 $M = \{(x, y) \in X \times Y : f(x, y) \neq g(x, y)\}$에 대해 $(\mu \otimes \nu)(M) = 0$이므로 정리 ? ? 로부터 $\int_X \nu(M_x) \, d\mu(x) = (\mu \otimes \nu)(M) = 0$이고, 정리 ? ? 로부터 명백히 $\mathscr{A}$−가측인 집합 $N = \{x \in X : \nu(M_x) \neq 0\}$에 대해 $\mu(N) = 0$이다. 이상을 종합하면 임의의 $x \in X \setminus N$에 대해 $[(X \times Y) \setminus M]_x = Y \setminus M_x$ 위에서 $f_x = g_x$이므로 곧 μ−거의 대부분의 $x \in X$에 대해 $f_x = g_x$ $(\nu\text{−ae.})$이고, 정리 ? ? 와 보조정리 ? ? 로부터 f_x는 μ−거의 대부분의 $x \in X$에 대해 $\mathscr{B}/\mathscr{B}_1$−가측이다. 나아가, 함수 $\widetilde{\varphi} : X \to \overline{\mathbb{R}}_0^+$를 $\widetilde{\varphi} : x \mapsto \int_Y g_x \, d\nu$로 두면 다시 보조정리 ? ? 로부터 이는 $\mathscr{A}/\mathscr{B}_1$−가측이 되고, 따름정리 ? ? 의 ii에서 $\varphi = \widetilde{\varphi}$ $(\mu\text{−ae.})$이므로 정리 ? ? 로부터 φ는 $\mathscr{A}/\mathscr{B}_1$−가측이다. □

Theorem 1.164 (Fubini) σ−유한 완비측도공간 $(X, \mathscr{A}, \mu)$, $(Y, \mathscr{B}, \nu)$와 음이 아닌 $\overline{\mathscr{A} \otimes \mathscr{B}}/\mathscr{B}_1$−가측함수 $f : X \times Y \to \overline{\mathbb{R}}_0^+$에 대해

$$\int_{X \times Y} f(x, y) \, d\overline{\mu \otimes \nu}(x, y) = \int_X \left[\int_Y f(x, y) \, d\nu(y) \right] d\mu(x) = \int_Y \left[\int_X f(x, y) \, d\mu(x) \right] d\nu(y)$$

가 성립한다. 단, 만약 고정된 $x \in X$에 대해 $f(x, y)$가 $\mathscr{B}/\mathscr{B}_1$−가측이 아니라면 이때의 $\int_Y f(x, y) \, d\nu(y)$의 값은 0으로 간주하고, 이는 고정된 $y \in Y$에 대해 $f(x, y)$가 $\mathscr{A}/\mathscr{B}_1$−가측이 아닌 경우에도 마찬가지이다.

f

Proof 먼저 위의 보조정리로부터 이중적분들이 well-define됨을 안다. 이제 함수 $\widetilde{\varphi} : X \to \overline{\mathbb{R}}_0^+$와 $g : X \times Y \to \overline{\mathbb{R}}_0^+$를 위의 보조정리의 증명에서와 같이 두면 $\varphi = \widetilde{\varphi}$ $(\mu\text{−ae.})$에서 Fubini의 정리와 따름정리 ? ? 의 ii, 정리 ? ? 로부터 $\int_X \varphi \, d\mu = \int_X \widetilde{\varphi} \, d\mu = \int_{X \times Y} g \, d(\mu \otimes \nu) = \int_{X \times Y} g \, d\overline{\mu \otimes \nu} = \int_{X \times Y} f \, d\overline{\mu \otimes \nu}$이다. 이제 두 번째 등식에 대해서도 이와 비슷하게 하면 된다. □

Lemma 1.165 $\sigma-$유한 완비측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$와 $\overline{\mu \otimes \nu}-$적분가능한 함수 $f: X \times Y \to \overline{\mathbb{R}}$에 대해 다음이 성립한다.

i. Section $f_x : Y \to \overline{\mathbb{R}}$는 $\mu-$거의 대부분의 $x \in X$에 대해 $\nu-$적분가능하며 함수 $\varphi : X \to \mathbb{R}$를

$$\varphi : x \mapsto \begin{cases} \displaystyle\int_Y f_x \, d\nu & f_x\text{가 } \nu\text{−적분가능한 경우} \\ 0 & \text{ow.} \end{cases}$$

라 하면 이는 $\mu-$적분가능하다.

ii. Section $f^y : X \to \overline{\mathbb{R}}$는 $\nu-$거의 대부분의 $y \in Y$에 대해 $\mu-$적분가능하며 함수 $\psi : Y \to \mathbb{R}$를

$$\psi : x \mapsto \begin{cases} \displaystyle\int_X f^y \, d\mu & f^y\text{가 } \mu\text{−적분가능한 경우} \\ 0 & \text{ow.} \end{cases}$$

라 하면 이는 $\nu-$적분가능하다.

Proof 증명은 보조정리 ? ? 와 거의 비슷하다. 이번에도 간결한 논의를 위해 i만 보이도록 하자. (ii는 i과 비슷하게 하면 된다.) 함수 $f_\pm$가 $\overline{\mathscr{A} \otimes \mathscr{B}}/\mathscr{B}_1-$가측이므로 함수 $\widetilde{\varphi}^\pm : X \to \overline{\mathbb{R}}_0^+$를

$$\widetilde{\varphi}^\pm : x \mapsto \begin{cases} \displaystyle\int_Y (f_\pm)_x \, d\nu & (f_\pm)_x\text{가 } \mathscr{B}/\mathscr{B}_1\text{−가측인 경우} \\ 0 & \text{ow.} \end{cases}$$

로 두면(여기서 위첨자로 쓰인 $\pm$은 정의 ? ? 에서 아래첨자로 사용된 $\pm$와는 달리 함수의 양과 음의 부분을 의미하는 것이 아니다.) 보조정리 ? ? 로부터 $(f_\pm)_x$ 중 하나라도 $\mathscr{B}/\mathscr{B}_1-$가측이 아닌 $x \in X$의 집합 M은 $\mu-$영집합이고, $\widetilde{\varphi}^\pm$는 $\mathscr{A}/\mathscr{B}_1-$가측이 되어 곧 Fubini의 정리로부터 $\int_X \widetilde{\varphi}^\pm \, d\mu = \int_{X \times Y} f_\pm \, d\overline{\mu \otimes \nu} \leq \int_{X \times Y} |f| \, d\overline{\mu \otimes \nu} < \infty$에서 $\widetilde{\varphi}^\pm$이 $\mu-$적분가능함을 안다. 그렇다면 $\widetilde{\varphi}^\pm$이 $\mu-$거의 어디서나 유한하므로 명백히 $\mathscr{A}-$가측인 집합 $N = (\widetilde{\varphi}^\pm)^{-1}(\infty) = \{x \in X \setminus M : \int_Y |f_x| \, d\nu = \infty\}$에 대해 $\mu(N) = 0$이고, 곧 f_x는 $\mu-$거의 대부분의 $x \in X$에 대해 $\nu-$적분가능하다. 이제 함수 $\varphi^\pm : X \to \mathbb{R}_0^+$를 $\varphi^\pm = \widetilde{\varphi}^\pm \mathbf{1}_{X \setminus (M \cup N)}$으로 두면 $\varphi^\pm = \widetilde{\varphi}^\pm$ ($\mu-$ae.)이므로 따름정리 ? ? 의 ii로부터 $\varphi^\pm$은 $\mu-$적분가능하고, 임의의 $x \in X$에 대해 $(f_x)_\pm = (f_\pm)_x$가 성립하여 $\varphi = \varphi^+ - \varphi^-$이므로 증명이 끝난다. □

Theorem 1.166 (Fubini) $\sigma-$유한 완비측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$와 $\overline{\mu \otimes \nu}-$적분가능한 함수 $f : X \times Y \to \overline{\mathbb{R}}$에 대해

$$\int_{X\times Y} f(x,y)\,d\overline{\mu\otimes\nu}(x,y)=\int_X\left[\int_Y f(x,y)\,d\nu(y)\right]d\mu(x)=\int_Y\left[\int_X f(x,y)\,d\mu(x)\right]d\nu(y)$$

가 성립한다. 단, 만약 고정된 $x\in X$에 대해 $f(x,y)$가 $\nu-$적분가능하지 않다면 이때의 $\int_Y f(x,y)\,d\nu(y)$의 값은 0으로 간주하고, 이는 고정된 $y\in Y$에 대해 $f(x,y)$가 $\mu-$적분가능하지 않은 경우에도 마찬가지이다.

Proof 증명은 앞선 보인 Fubini의 정리와 거의 비슷하다. 먼저 위의 보조정리로부터 이 중적분들이 well-define됨을 안다. 이제 함수 $\varphi:X\to\mathbb{R}$와 $\varphi^{\pm}:X\to\mathbb{R}_0^+$를 각각 위의 보조정리와 그 증명에서와 같이 두면 $\varphi=\varphi^+-\varphi^-$에서 $\int_X\varphi\,d\mu=\int_X\varphi^+\,d\mu-\int_X\varphi^-\,d\mu=\int_{X\times Y}f_+\,d\overline{\mu\otimes\nu}-\int_{X\times Y}f_-\,d\overline{\mu\otimes\nu}=\int_{X\times Y}f\,d\overline{\mu\otimes\nu}$임을 쉽게 알 수 있다. 이제 두 번째 등식에 대해서도 이와 비슷하게 하면 된다. □

마침내, Lebesgue 측도공간에서 적용될 수 있는 일반적인 Fubini의 정리를 얻을 수 있다.

Corollary 1.167 $l+m=n$인 $l,m,n\in\mathbb{N}$과 음이 아닌 가측함수 $f:\mathbb{R}^n\to\overline{\mathbb{R}}_0^+$에 대해

$$\int_{\mathbb{R}^n} f(x,y)\,d\lambda_n(x,y)=\int_{\mathbb{R}^l}\left[\int_{\mathbb{R}^m} f(x,y)\,d\lambda_m(y)\right]d\lambda_l(x)=\int_{\mathbb{R}^m}\left[\int_{\mathbb{R}^l} f(x,y)\,d\lambda_l(x)\right]d\lambda_m(y)$$

가 성립한다. 한편, 적분가능한 $f:\mathbb{R}^n\to\overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립한다. 단, 어느 경우에나 만약 고정된 $x\in\mathbb{R}^l$에 대해 $f(x,y)$가 가측이 아니거나 적분가능하지 않다면 이때의 $\int_{\mathbb{R}^m} f(x,y)\,d\lambda_m(y)$의 값은 0으로 간주하고, 이는 고정된 $y\in\mathbb{R}^m$에 대해 $f(x,y)$가 가측이 아니거나 적분가능하지 않은 경우에도 마찬가지이다.

Proof 이는 Fubini의 정리와 정리 ? ? , ? ? 의 ii로부터 자명하다. □

이렇게 얻은 일반적인 측도에 대한 Fubini의 정리를 사용하여 앞서 증명한 Minkowski의 부등식을 일반화하는 것으로 이번 절을 마무리한다. 이 결과는 이후 합성곱에 대해 공부할 때 유용하게 사용될 것이다.

Lemma 1.168 측도공간 $(X,\mathscr{A},\mu)$와 $1/p+1/q=1$인 $p,q>1$, 음이 아닌 가측함수 $f:X\to\overline{\mathbb{R}}_0^+$를 생각하자. 만약 적당한 $C\geq 0$가 존재하여 음이 아닌 임의의 $g\in L^q$에 대해 $\int_X fg\,d\mu\leq C||g||_q$를 만족하면 다음이 성립한다.

i. 만약 $||f||_p<\infty$이면 $||f||_p\leq C$이다.
ii. 측도 μ가 $\sigma-$유한하면 $||f||_p\leq C$이다.

Proof i. 만약 $||f||_p=0$이면 정리가 자명하므로 $||f||_p>0$이라 하자. 가정과 정리 ? ? 의 ii로부터 $N=f^{-1}(\infty)$는 영집합이므로 함수 $h:X\to\mathbb{R}_0^+$를 $h=f\mathbf{1}_{X\setminus N}$으로 두면 $f=h$ (ae.)이고 $||h||_p=||f||_p<\infty$이며 임의의 음이 아닌 $g\in L^q$에 대해 $\int_X hg\,d\mu=\int_X fg\,d\mu\leq$

$C||g||_q$이다. 이제 함수 $g : X \to \mathbb{R}_0^+$를 $g = h^{p-1}$로 두면 $g^q = h^{(p-1)q} = h^p = hg$에서 $||g||_q = (\int_X h^p \, d\mu)^{1-1/p} = ||h||_p^{p-1} < \infty$가 되어 $g \in L^q$이고 곧 $||h||_p^p = \int_X h^p \, d\mu = \int_X hg \, d\mu \leq C||g||_q = C||h||_p^{p-1}$이다. 그런데 앞서 $||h||_p = ||f||_p > 0$이라 가정하였으므로 이는 $||f||_p = ||h||_p \leq C$임을 함의하여 증명이 끝난다.

ii. 모순을 유도하기 위해 $||f||_p > C$라 하면 정의로부터 적당한 음이 아닌 단순함수 $h : X \to \mathbb{R}_0^+$가 존재하여 $h \leq f$이고 $||h||_p > C$이다. 여기서 $||h||_p < \infty$인 경우에는 $g \geq 0$인 임의의 $g \in L^q$에 대해 $\int_X hg \, d\mu \leq \int_X fg \, d\mu \leq C||g||_q$이므로 i로부터 $||h||_p \leq C$의 모순이 발생하여 증명이 끝나므로 $||h||_p = \infty$인 경우에 대해서만 조금 더 생각해보자. 이 경우에는 h의 표준형을 $h = \sum_{i=1}^l a_i \mathbf{1}_{A_i}$라 하면 적당한 $i_0 \leq k$에 대해 $a_{i_0} > 0$이고 $\mu(A_{i_0}) = \infty$이다. 한편, 가정으로부터 $\mathscr{A}$에 속하는 적당한 집합열 $\{B_j\}$가 존재하여 각 $j \in \mathbb{N}$에 대해 $\mu(B_j) < \infty$이고 $\bigcup_{j=1}^\infty B_j = X$이므로 집합열 $\{C_k\}$를 $C_k := A_{i_0} \cap \bigcup_{j=1}^k B_j$로 두면 이는 증가하는 집합열로서 A_{i_0}으로 수렴함이 분명하므로 정리 ? ? 의 i로부터 충분히 큰 $k_0 \in \mathbb{N}$에 대해 $(C/a_{i_0})^p < \mu(C_{k_0}) \leq \mu(\bigcup_{j=1}^{k_0} B_j) \leq \sum_{j=1}^{k_0} \mu(B_j) < \infty$이다. 따라서 단순함수 $\widetilde{h} : X \to \mathbb{R}_0^+$를 $\widetilde{h} = a_{i_0} \mathbf{1}_{C_{k_0}}$로 두면 $||\widetilde{h}||_p = a_{i_0}[\mu(C_{k_0})]^{1/p}$에서 $C < ||\widetilde{h}||_p < \infty$이고 $\widetilde{h} \leq h \leq f$이므로 앞서 $||h||_p < \infty$인 경우에 모순을 유도한 것과 비슷하게 $||\widetilde{h}||_p \leq C$의 모순을 유도할 수 있고, 증명은 이로써 충분하다. □

Theorem 1.169 (Minkowski's integral inequality) $\sigma-$유한 측도공간 $(X, \mathscr{A}, \mu)$, $(Y, \mathscr{B}, \nu)$에 대해 함수 $f : X \times Y \to \overline{\mathbb{R}}$가 $(\mathscr{A} \otimes \mathscr{B})/\mathscr{B}_1-$가측이라 하자. 그렇다면 $p \geq 1$에 대해

$$\left\{ \int_X \left[\int_Y |f(x,y)| \, d\nu(y) \right]^p d\mu(x) \right\}^{1/p} \leq \int_Y \left[\int_X |f(x,y)|^p \, d\mu(x) \right]^{1/p} d\nu(y)$$

이다. 즉, $|| \int_Y |f_x| \, d\nu ||_p \leq \int_Y ||f^y||_p \, d\nu$이다.

Proof 우선 Fubini의 정리로부터 위 부등식의 적분은 모두 well-define된다. 한편, $p = 1$이면 부등식이 자명하여 더 이상 보일 것이 없으므로 $p > 1$이라 하자. 또한, $\int_Y ||f^y||_p \, d\nu = \infty$인 경우에도 부등식이 자명하므로 $C := \int_Y ||f^y||_p \, d\nu < \infty$라 하자. 이제 $1/p + 1/q = 1$인 $q > 1$에 대해 음이 아닌 임의의 $g \in L^q(\mu)$를 택하고 함수 $\varphi : X \to \overline{\mathbb{R}}_0^+$를 $\varphi : x \mapsto \int_Y |f_x| \, d\nu$로 두면 이는 Fubini의 정리에서 $\mathscr{A}/\mathscr{B}_1-$가측이고 φg는 $(\mathscr{A} \otimes \mathscr{B})/\mathscr{B}_1-$가측이므로 다시 Fubini의 정리와 Hölder의 부등식으로부터 $\int_X \varphi g \, d\mu = \int_X [\int_Y |f(x,y)| g(x) \, d\nu(y)] \, d\mu(x) = \int_Y [\int_X |f(x,y)| g(x) \, d\mu(x)] \, d\nu(y) \leq \int_Y ||f^y||_p ||g||_q \, d\nu(y) = C||g||_q$이다. 그렇다면 위의 보조정리로부터 $||\varphi||_p \leq C$가 되어 증명이 끝난다. □

1.12 Transformation formula

이번 절에서는 변수변환 공식이라는 단 하나의 정리를 증명하도록 한다. 이렇게 하나의 정리를 증명하는데 하나의 절을 오롯이 할애하는 것에는 이전 절이 Fubini의 정리의 증명에 할애된 것과 같은 맥락의 이유와 더불어, 그 증명이 굉장히 길기 때문이다. (아마 이 책을 통틀어 가장 긴 증명이 아닐까 싶다.) 대부분의 독자들이 변수변환 공식이 어떤 공식인지, 언제 사용하는지 등에 대해 잘 알고 있을 것이고, 긴 증명과정과는 달리 이에 별다른 준비가 필요하지는 않으므로 이만 각설하고, 바로 증명을 시작하자. 증명은 2개의 보조정리의 도움을 받아 총 10단계로 이루어져 있다.

Lemma 1.170 집합 $N \subseteq \mathbb{R}^n$이 영집합일 필요충분조건은 임의의 $\varepsilon > 0$에 대해 적당한 열린 ball의 열 $\{B(x_i, r_i)\}$가 존재하여 $N \subseteq \bigcup_{i=1}^{\infty} B(x_i, r_i)$이고 $\sum_{i=1}^{\infty} r_i^n < \varepsilon$인 것이다.

Proof 먼저 충분조건임을 보이기 위해 임의의 $\varepsilon > 0$을 택하면 Lebesgue 측도가 regular하므로 적당한 열린집합 $U \subseteq \mathbb{R}^n$가 존재하여 $N \subseteq U$이고 $\lambda_n(U) < (2/\sqrt{n})^n \varepsilon$이다. 또한, $\mathscr{S}_n$에 속하는 적당한 집합열 $\{B_i\}$가 존재하여 $U = \bigsqcup_{i=1}^{\infty} B_i$이고 각 $i \in \mathbb{N}$에 대해 B_i의 모든 모서리의 길이가 l_i로 같다.[20] 그렇다면 $\sum_{i=1}^{\infty} l_i^n = \sum_{i=1}^{\infty} \lambda_n(B_i) = \lambda_n(\bigsqcup_{i=1}^{\infty} B_i) = \lambda_n(U) < (2/\sqrt{n})^n \varepsilon$이 성립하므로 각 $i \in \mathbb{N}$에 대해 x_i를 B_i의 중심이라 하여 $\{B(x_i, \sqrt{n} l_i/2)\}$로 정의된 열린 ball의 열을 생각하면 $\sum_{i=1}^{\infty} (\sqrt{n} l_i/2)^n < \varepsilon$이다. 이제 각 $i \in \mathbb{N}$에 대해 $B_i \subseteq B(x_i, \sqrt{n} l_i/2)$임은 분명하므로 $r_i = \sqrt{n} l_i/2$라 하면 충분조건임이 보여진다.

한편, 필요조건임은 거의 자명하다. 가정으로부터 임의의 $\varepsilon > 0$을 택하면 적당한 열린 ball의 열 $\{B(x_i, r_i)\}$가 존재하여 $\overline{N} \subseteq \bigcup_{i=1}^{\infty} \overline{B(x_i, r_i)}$이고 $\sum_{i=1}^{\infty} r_i^n < \varepsilon/4^n$인데, 각 $i \in \mathbb{N}$에 대해 $\overline{B(x_i, r_i)} \subseteq (-2r_i, 2r_i]^n + x_i$이므로 $\lambda_n(\overline{B(x_i, r_i)}) \le (4r_i)^n$에서 $\lambda_n(\overline{N}) \le \lambda_n(\bigcup_{i=1}^{\infty} \overline{B(x_i, r_i)}) \le \sum_{i=1}^{\infty} \lambda_n(\overline{B(x_i, r_i)}) \le \sum_{i=1}^{\infty} (4r_i)^n < \varepsilon$이다. 이는 곧 $\lambda_n(\overline{N}) = 0$임을 뜻하므로 필요조건임도 보여진다. □

Lemma 1.171 영집합 $N \subseteq \mathbb{R}^n$과 Lipschitz 연속인 $f : \mathbb{R}^n \to \mathbb{R}^n$에 대해 $f(N)$도 영집합이다.

Proof 함수 f의 Lipschitz 상수를 $L > 0$이라 하자. 위의 보조정리로부터 임의의 $\varepsilon > 0$을 택하면 적당한 열린 ball의 열 $\{B(x_i, r_i)\}$가 존재하여 $N \subseteq \bigcup_{i=1}^{\infty} B(x_i, r_i)$이고 $\sum_{i=1}^{\infty} r_i^n < \varepsilon/L^n$이다. 한편, 각 $i \in \mathbb{N}$와 임의의 $x \in B(x_i, r_i)$에 대해 $||f(x) - f(x_i)|| < L||x - x_i|| < Lr_i$에서 $f(B(x_i, r_i)) \subseteq B(f(x_i), Lr_i)$이다. 이로부터 $f(N) \subseteq f(\bigcup_{i=1}^{\infty} B(x_i, r_i)) = \bigcup_{i=1}^{\infty} f(B(x_i, r_i)) \subseteq \bigcup_{i=1}^{\infty} B(f(x_i), Lr_i)$이고 $\sum_{i=1}^{\infty} (Lr_i)^n < \varepsilon$이므로 다시 위의 보조정리로부터 $f(N)$이 영집합임을 안다. □

Theorem 1.172 (Transformation formula) 열린집합 $U, V \subseteq \mathbb{R}^n$ 사이에서 정의된 $\mathscr{C}^1$급 미분동형사상 $\Phi : U \to V$와 음이 아닌 가측함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$에 대해 합성 $f \circ \Phi : U \to \overline{\mathbb{R}}_0^+$

는 가측이고 $\int_V f\,d\lambda_n = \int_U (f\circ\Phi)|\det\mathbf{D}\Phi|\,d\lambda_n$이다. 한편, V에서 적분가능한 (혹은 $(f\circ\Phi)|\det\mathbf{D}\Phi|$가 U에서 적분가능한) $f:\mathbb{R}^n\to\overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며, 이때 $(f\circ\Phi)|\det\mathbf{D}\Phi|$는 U에서 적분가능하다 (혹은 f는 V에서 적분가능하다). [21]

Proof 증명은 총 10단계로 구성되어 있다. 증명이 상당히 짜증날 것이므로 인내심을 갖길 바란다. (대략 5~6단계를 전후하여 고비가 온다.) 이제부터 임의의 음이 아닌 가측함수 f에 대해 위의 정리를 만족하는 미분동형사상 Φ를 good이라 하자.

Step 1. 미분동형사상 $\Phi:U\to V$가 good일 필요충분조건은 임의의 가측인 $A\subseteq V$에 대해 $\Phi^{-1}(A)$도 가측이고 $\lambda_n(A)=\int_{\Phi^{-1}(A)}|\det\mathbf{D}\Phi|\,d\lambda_n$인 것이다.

먼저 충분조건임을 보이기 위해 임의의 가측집합 $A\subseteq V$를 고정하면 $\mathbf{1}_A$가 가측이고 가정으로부터 Φ가 good이므로 $\mathbf{1}_{\Phi^{-1}(A)}=\mathbf{1}_A\circ\Phi$가 가측이 되어 곧 $\Phi^{-1}(A)$가 가측이고, $\lambda_n(A)=\int_V \mathbf{1}_A\,d\lambda_n=\int_U(\mathbf{1}_A\circ\Phi)|\det\mathbf{D}\Phi|\,d\lambda_n=\int_{\Phi^{-1}(A)}|\det\mathbf{D}\Phi|\,d\lambda_n$이다. 이제 필요조건임을 보이기 위해 먼저 f가 음이 아닌 단순함수인 경우를 생각하여 단순함수 $f\mathbf{1}_V$의 표준형을 $f\mathbf{1}_V=\sum_{i=1}^k a_i\mathbf{1}_{A_i}$라 하면 $f\circ\Phi=(f\mathbf{1}_V)\circ\Phi=\sum_{i=1}^k a_i\mathbf{1}_{\Phi^{-1}(A_i)}$가 가측이고

$$\begin{aligned}
\int_V f\,d\lambda_n &= \sum_{i=1}^k a_i\lambda_n(A_i)\\
&= \sum_{i=1}^k a_i\int_{\Phi^{-1}(A_i)}|\det\mathbf{D}\Phi|\,d\lambda_n\\
&= \int_U\sum_{i=1}^k a_i\mathbf{1}_{\Phi^{-1}(A_i)}|\det\mathbf{D}\Phi|\,d\lambda_n\\
&= \int_U\sum_{i=1}^k a_i(\mathbf{1}_{A_i}\circ\Phi)|\det\mathbf{D}\Phi|\,d\lambda_n\\
&= \int_U(f\circ\Phi)|\det\mathbf{D}\Phi|\,d\lambda_n
\end{aligned}$$

이 성립한다. 다음으로, 일반적인 음이 아닌 가측함수 f를 생각하면 정리 ? ? 로부터 음이 아닌 단순함수 $f_i:\mathbb{R}^n\to\overline{\mathbb{R}}_0^+$의 열 $\{f_i\}$에 대해 $f_i\uparrow f$인데, 각 $i\in\mathbb{N}$에 대해 앞선 결과로부터 $f_i\circ\Phi$가 가측이므로 정리 ? ? 의 ii로부터 $f\circ\Phi$도 가측이다. 또한 MCT로부터 $\int_V f\,d\lambda_n=\lim_{i\to\infty}\int_V f_i\,d\lambda_n=\lim_{i\to\infty}\int_U(f_i\circ\Phi)|\det\mathbf{D}\Phi|\,d\lambda_n=\int_U(f\circ\Phi)|\det\mathbf{D}\Phi|\,d\lambda_n$도 성립하므로 Φ가 good임을 안다.

Step 2. 열린집합 $U,V,W\subseteq\mathbb{R}^n$에 대해 미분동형사상 $\Phi:U\to V,\Psi:V\to W$가 good이라면 합성 $\Psi\circ\Phi:U\to W$도 good이다.

임의의 음이 아닌 가측함수 f에 대해 가정으로부터 $f\circ(\Psi\circ\Phi)$가 가측임은 자명하다. 또한 $\int_W f\,d\lambda_n=\int_V(f\circ\Psi)|\det\mathbf{D}\Psi|\,d\lambda_n$인데, 여기서 $g:=(f\circ\Psi)|\det\mathbf{D}\Psi|$도 음이 아닌 가측함수이므로 다시 가정으로부터 $\int_V g\,d\lambda_n=\int_U(g\circ\Phi)|\det\mathbf{D}\Phi|\,d\lambda_n$이 되어 이상을 종합하면

$$\begin{aligned}\int_W f\,d\lambda_n &= \int_U (f\circ\Psi\circ\Phi)|\det\mathbf{D}\Psi\circ\Phi||\det\mathbf{D}\Phi|\,d\lambda_n\\ &= \int_U [f\circ(\Psi\circ\Phi)]|\det(\mathbf{D}\Psi\circ\Phi)\mathbf{D}\Phi|\,d\lambda_n\\ &= \int_U [f\circ(\Psi\circ\Phi)]|\det\mathbf{D}(\Psi\circ\Phi)|\,d\lambda_n\end{aligned}$$

에서 $\Psi\circ\Phi$가 good임을 안다.

Step 3. 모든 평행이동 $T:\mathbb{R}^n\to\mathbb{R}^n$는 good이다.

평행이동인 T에 대해 $|\det\mathbf{D}T|=1$이고 Lebesgue 측도가 이동 불변성을 가지므로 임의의 가측인 $A\subseteq\mathbb{R}^n$에 대해 $T^{-1}(A)$가 가측이며 $\lambda_n(A)=\lambda_n(T^{-1}(A))=\int_{T^{-1}(A)}|\det\mathbf{D}T|\,d\lambda_n$이므로 1단계에서 T는 good임을 안다.

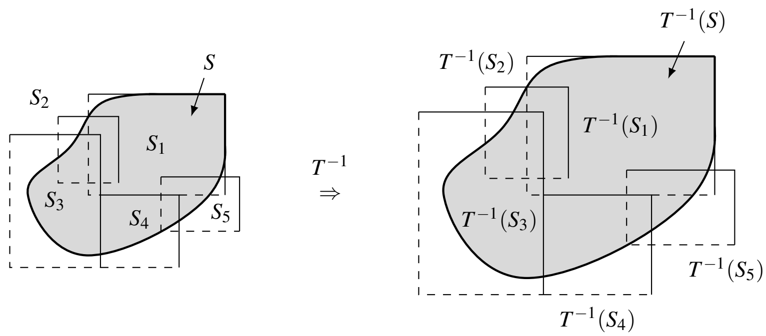
Step 4. 모든 homothety $T:\mathbb{R}^n\to\mathbb{R}^n$는 good이다.

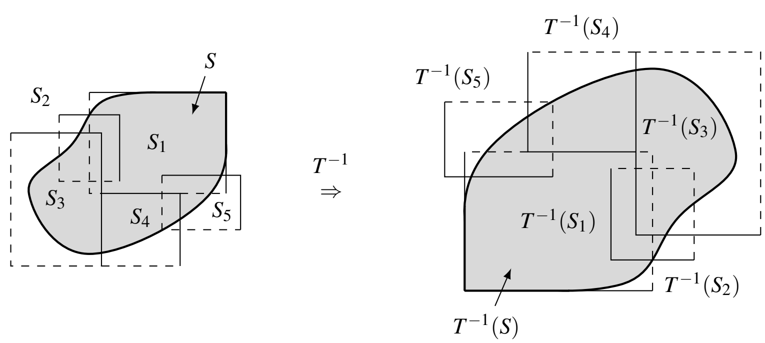
Homothety인 T는 적당한 0이 아닌 $c\in\mathbb{R}$에 대해 $T:x\mapsto cx$로 쓸 수 있다. 이번에도 단계 1로써 T가 good임을 보이도록 하자. 먼저 임의의 semi-open box $B\subseteq\mathbb{R}^n$를 생각하면 T가 연속이므로 곧 가측이고, 따라서 $T^{-1}(B)$도 가측이다. 이제 $\lambda_n(T^{-1}(B))=\rho_n(B)/|c|^n$임을 보이고자 하는데, 만약 B가 유계가 아니라면 $T^{-1}(B)$도 유계가 아니어서 이가 자명하므로 B가 유계라 하자. 그렇다면 적당한 $x,y\in\mathbb{R}^n$가 존재하여 $B=(x,y]$이고, 곧 보조정리 ? ? 로부터 $\lambda_n(T^{-1}(B))=\lambda_n(B/c)=\prod_{i=1}^n(y_i-x_i)/|c|=\rho_n(B)/|c|^n$이다.[22] 나아가, 정리 ? ? 로부터 임의의 $A\in A(\mathscr{S}_n)$에 대해서도 $T^{-1}(A)$가 가측이고 $\lambda_n(T^{-1}(A))=\rho_n(A)/|c|^n$임이 자명하다.

마지막으로, 임의의 $S\subseteq\mathbb{R}^n$를 택하고 이의 $A(\mathscr{S}_n)$에 속하는 임의의 가산 덮개 $\{S_i\}$를 생각하여, WLOG, 필요하다면 공집합을 추가하여 이가 집합열을 이룬다고 하자. 만약 $c>0$이라면 $\{T^{-1}(S_i)\}$가 $A(\mathscr{S}_n)$에 속하는 $T^{-1}(S)$의 가산 덮개임이 분명하므로 앞선 결과로부터 $\rho_n^*(T^{-1}(S))=\rho_n^*(S)/|c|^n$임이 분명하다. 한편, $c<0$인 경우에는 조금 복잡하다. 먼저 각 $i\in\mathbb{N}$에 대해 $T^{-1}(S_i)$를 구성하는 서로소인 box들에서 원래 속한 면은 빼고, 원래 빠진 면은 더하여 얻는 집합을 T_i라 하면 이는 다시 semi-open box들의 서로소 합집합이 되고, 앞선 결과와 보조정리 ? ? 로부터 $\rho_n(T_i)=\lambda_n(T^{-1}(S_i))=\rho_n(S_i)/|c|^n$이다. (각주 ? ? 참조.) 그러나 원래 속한 면을 빼는 과정에서 $\{T_i\}$가 더 이상 $T^{-1}(S)$의 가산 덮개가 되지 못할 수 있다. 이 문제를 해결하기 위해 임의의 $\varepsilon>0$을 택하여 각 $i\in\mathbb{N}$에 대해 T_i를 이루는 semi-open box들을 조금씩 늘려 얻은 T_i'가 $\rho_n(T_i')=\rho_n(T_i)+\varepsilon/2^i$가 되도록 하면 이때의 $\{T_i'\}$는 다시 $T^{-1}(S)$의 가산 덮개가 되면서 $\sum_{i=1}^\infty\rho_n(T_i')=\rho_n(S_i)/|c|^n+\varepsilon$을 만족한다. 이로부터 $\rho_n^*(T^{-1}(S))\le\rho_n(S_i)/|c|^n+\varepsilon$이고, 곧 $\rho_n^*(T^{-1}(S))\le\rho_n^*(S)/|c|^n$이다. 그런데 여기서 T는 임의의 homothety이고, T^{-1}도 $T^{-1}:x\mapsto x/c$인 homothety이므로 $\rho_n^*(T(S))/|c|^n\le\rho_n^*(S)$에서 $\rho_n^*(S)/|c|^n\le\rho_n^*(T^{-1}(S))$가 되어 이상으로부터 c의 부호에 무관하게 $\rho_n^*(T^{-1}(S))=\rho_n^*(S)/|c|^n$임을 안다. 이러한 결론이 임의의 $S\subseteq\mathbb{R}^n$에 대해 성립한다는 점을 상기하면, 임의의 가측인 $A\subseteq\mathbb{R}^n$와 임의의 $S\subseteq\mathbb{R}^n$에 대해

$$\begin{aligned}
\rho_n^*(S\cap T^{-1}(A))+\rho_n^*(S\cap[T^{-1}(A)]^c) &= \rho_n^*(S\cap T^{-1}(A))+\rho_n^*(S\cap T^{-1}(A^c)) \\
&= \frac{\rho_n^*(T(S)\cap A)+\rho_n^*(T(S)\cap A^c)}{|c|^n} \\
&= \frac{\rho_n^*(T(S))}{|c|^n} \\
&= \rho_n^*(S)
\end{aligned}$$

이므로 $T^{-1}(A)$도 가측이며, 앞선 결론으로부터 $\lambda_n(T^{-1}(A))=\lambda_n(A)/|c|^n$이다. 이제 $\lambda_n(A)=|c|^n\lambda_n(T^{-1}(A))=\int_{T^{-1}(A)}|\det\mathbf{D}T|\,d\lambda_n$이므로 단계 1로부터 T가 good임을 안다.





Step 5. 모든 가역인 선형사상 $T:\mathbb{R}^n\to\mathbb{R}^n$는 good이다.

모든 가역인 선형사상은 (표준기저에 대해) 기본행렬로 나타내어지는 유한개의 선형사상의 합성으로 쓸 수 있으므로 2단계로부터 T가 다음 3가지와 같은 경우에 대해서만 생각하면 된다.

i. $T:(\cdots,x_i,\cdots)\mapsto(\cdots,cx_i,\cdots)$ (단, $c\neq 0$이다.)

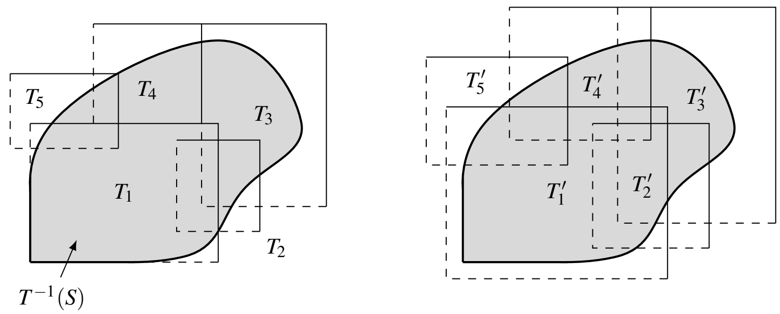


Figure 1.12 변수변환 공식의 증명의 4단계에서의 집합 S와 임의의 가산 덮개 $\{S_i\}$가 homothety T에 의해 변환되는 모습. 각각 $c > 0$인 경우(위)와 $c < 0$인 경우(중간). 한편, $c < 0$인 경우에는 $\{T_i\}$가 $T^{-1}(S)$의 가산 덮개가 되지 못할 수도 있으므로(왼쪽 아래) 각 T_i를 조금씩 늘려 $\{T_i'\}$를 구성한다(오른쪽 아래).

ii. $T : (\cdots, x_i, \cdots) \mapsto (\cdots, x_i + cx_j, \cdots)$
iii. $T : (\cdots, x_i, \cdots, x_j, \cdots) \mapsto (\cdots, x_j, \cdots, x_i, \cdots)$

간결한 논의를 위해 $i = 1$, $j = 2$인 경우에 대해서만 보이도록 하고, 이하 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$를 임의의 가측함수라 하자. (다른 i, j에 대해서도 이와 비슷하게 하면 된다.) 우선 $f \circ T$가 가측임을 보이려고 하는데, 만약 f가 Borel이면 정리 ? ? 로부터 $f \circ T$가 가측임이 분명하다. 한편, 정리 ? ? 와 ? ? 로부터 적당한 Borel 함수 $g : \mathbb{R}^n \to \overline{\mathbb{R}}$가 존재하여 $f = g$ (ae.)이다. 이제 $N = \{x \in \mathbb{R}^n : f(x) \neq g(x)\}$라 하면 이가 영집합이고 T^{-1}가 Lipschitz 연속이므로 위의 보조정리로부터 $T^{-1}(N) = \{x \in \mathbb{R}^n : (f \circ T)(x) \neq (g \circ T)(x)\}$가 영집합이 되어 $f \circ T = g \circ T$ (ae.)이다. 그렇다면 Lebesgue 측도공간이 완비성을 가진다는 사실과 정리 ? ? 로부터 $f \circ T$도 가측이다.

다음으로, 위의 세 가지 경우에 대해 $\int_{\mathbb{R}^n} f\, d\lambda_n = \int_{\mathbb{R}^n} (f \circ T)|\det \mathbf{D}T|\, d\lambda_n$임을 보이자. 먼저 첫 번째 경우를 보면 Fubini의 정리와 4단계의 결과로부터

$$\begin{aligned}
\int_{\mathbb{R}^n} f\, d\lambda_n &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \cdots, x_n)\, d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(cx_1, \cdots, x_n)|c|\, d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} (f \circ T)(x_1, \cdots, x_n)|c|\, d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}^n} (f \circ T)|\det \mathbf{D}T|\, d\lambda_n
\end{aligned}$$

이다. 두 번째 경우에도 Fubini의 정리와 3단계의 결과로부터

$$\int_{\mathbb{R}^n} f\, d\lambda_n = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \cdots, x_n)\, d\lambda_1(x_1) \cdots d\lambda_1(x_n)$$

$$
\begin{aligned}
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1 + cx_2, \cdots, x_n)\, d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} (f \circ T)(x_1, \cdots, x_n)\, d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}^n} (f \circ T) |\det \mathbf{D}T|\, d\lambda_n
\end{aligned}
$$

이며 마지막 경우에도 Fubini의 정리로부터

$$
\begin{aligned}
\int_{\mathbb{R}^n} f\, d\lambda_n &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, x_2, \cdots, x_n)\, d\lambda_1(x_2)\, d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_2, x_1, \cdots, x_n)\, d\lambda_1(x_1)\, d\lambda_1(x_2) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} (f \circ T)(x_1, \cdots, x_n)\, d\lambda_1(x_1)\, d\lambda_1(x_2) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}^n} (f \circ T) |\det \mathbf{D}T|\, d\lambda_n
\end{aligned}
$$

이다. 이상으로부터 어느 경우에나 T가 good이 되어 임의의 가역인 선형사상이 good임을 안다.

2단계, 3단계, 5단계의 결과를 종합하면 임의의 Affine 변환이 good임을 안다. 이제 증명의 후반부를 위해 임의의 $\mathscr{C}^1$급 미분동형사상 $\Phi : U \to V$와 임의의 compact한 $K \subseteq U$를 고정하고, $\mathbb{R}^n$에 최대값 노름 $||\cdot||_\infty$를 장착하자.

Step 6. 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 임의의 $x, y \in K$가 $||x - y||_\infty < \delta$와 $\prod_{i=1}^n [x_i, y_i] \subseteq K$를 만족하면 $||\Phi(x) - \Phi(y) - \mathbf{D}\Phi(x)(y - x)||_\infty < \varepsilon||x - y||_\infty$이다.

편의를 위해 $\prod_{i=1}^n [x_i, y_i] \subseteq K$를 만족하는 임의의 $x, y \in K$를 고정하자. 이제 함수 $\gamma : [0, 1] \to \mathbb{R}^n$를 $\gamma : t \mapsto \Phi((1 - t)x + ty)$로 두면 이는 $\mathscr{C}^1$급이고, FTC로부터 $\Phi(y) - \Phi(x) = \int_0^1 \nabla\gamma = \int_0^1 \mathbf{D}\Phi((1 - t)x + ty)(y - x)\, dt$이다. 이로부터 함수 $\Gamma : [0, 1] \to \mathsf{L}(\mathbb{R}^n, \mathbb{R}^n)$를 $\Gamma : t \mapsto \mathbf{D}\Phi((1 - t)x + ty) - \mathbf{D}\Phi(x)$로 두면 $\Phi(y) - \Phi(x) = \int_0^1 \mathbf{D}\Phi((1 - t)x + ty)(y - x)\, dt = \int_0^1 [\mathbf{D}\Phi(x) + \Gamma(t)](y - x)\, dt = \mathbf{D}\Phi(x)(y - x) + \int_0^1 \Gamma(t)(y - x)\, dt$이다. 한편, $\mathbf{D}\Phi : \mathbb{R}^n \to \mathsf{L}(\mathbb{R}^n, \mathbb{R}^n)$가 K 위에서 균등연속이므로 임의의 $\varepsilon > 0$을 택하면 적당한 $\delta > 0$가 존재하여 $||x - y||_\infty < \delta$이면 $||\mathbf{D}\Phi(x) - \mathbf{D}\Phi(y)||_{\text{op}} < \varepsilon$이다. 이제 $||x - y||_\infty < \delta$라 하고 임의의 $t \in [0, 1]$에 대해 $z = (1 - t)x + ty$라 하면 $||x - z||_\infty \le ||x - y||_\infty < \delta$에서 $||\Gamma(t)||_{\text{op}} = ||\mathbf{D}\Phi(x) - \mathbf{D}\Phi(z)||_{\text{op}} < \varepsilon$이 되어 이상으로부터 각 $i \le n$에 대해 $\int_0^1 \Gamma_i(t)(y - x)\, dt \le \int_0^1 ||\Gamma(t)(y - x)||_\infty\, dt \le \int_0^1 ||\Gamma(t)||_{\text{op}} ||x - y||_\infty\, dt < \varepsilon||x - y||_\infty$이므로 곧 $||\Phi(x) - \Phi(y) - \mathbf{D}\Phi(x)(y - x)||_\infty = ||\int_0^1 \Gamma(t)(y - x)\, dt||_\infty < \varepsilon||x - y||_\infty$가 성립한다.

이제 임의의 $x \in U$에 대해 함수 $\Psi_x : \mathbb{R}^n \to \mathbb{R}^n$를 $\Psi_x : y \mapsto \Phi(x) + \mathbf{D}\Phi(x)(x - y)$로 두면 이는 Affine 변환이며 곧 good이다. 나아가 6단계의 결과로부터 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 임의의 $x, y \in K$가 $||x - y||_\infty < \delta$와 $\prod_{i=1}^n [x_i, y_i] \subseteq K$를 만족하면 $||\Phi(y) - \Psi_x(y)||_\infty < \varepsilon||x - y||_\infty$이다.

Step 7. 임의의 $0 < \varepsilon < 1$에 대해 적당한 $\delta > 0$가 존재하여 임의의 열린 ball $B(x, r) \subseteq K$이 $r < \delta$를 만족하면 $\Psi_x(B(x, (1-\varepsilon)r)) \subseteq \Phi(B(x, r)) \subseteq \Psi_x(B(x, (1+\varepsilon)r))$이다.

미분동형사상 Φ가 $\mathscr{C}^1$급이고 집합 K가 compact하여 $(\mathbf{D}\Phi(\cdot))^{-1} : K \to \mathsf{L}(\mathbb{R}^n, \mathbb{R}^n)$가 유계이므로 $M := \sup_{x \in K} ||(\mathbf{D}\Phi(x))^{-1}||_{\text{op}} < \infty$이다. 그렇다면 임의의 $x \in K$와 임의의 $y, z \in \mathbb{R}^n$에 대해 $\Psi_x^{-1}(y) = (\mathbf{D}\Phi(x))^{-1}(\Phi(x) - y) + x$이고, 이는 z에 대해서도 마찬가지이므로 $||\Psi_x^{-1}(y) - \Psi_x^{-1}(z)||_\infty = ||(\mathbf{D}\Phi(x))^{-1}(y-z)||_\infty \leq ||(\mathbf{D}\Phi(x))^{-1}||_{\text{op}}||y-z||_\infty \leq M||y-z||_\infty$이다. 이제 임의의 $x \in K^\circ$를 고정하고 임의의 $0 < \varepsilon < 1$을 택하면 적당한 $\delta > 0$가 존재하여 임의의 $y \in K$가 $||x-y||_\infty < \delta$와 $\prod_{i=1}^n [x_i, y_i] \subseteq K$를 만족하면 $||\Phi(y) - \Psi_x(y)||_\infty < [\varepsilon/(M+1)]||x-y||_\infty$이다. 이러한 δ에 대해 임의의 $r < \delta$도 고정하되, WLOG, 필요하다면 r을 더 작게 하여 $B(x, r) \subseteq K$라 하면 임의의 $y \in B(x, r)$에 대해 $||x-y||_\infty < r < \delta$이고 $\prod_{i=1}^n [x_i, y_i] \subseteq B(x, r) \subseteq K$가 되어 $||\Phi(y) - \Psi_x(y)||_\infty < [\varepsilon/(M+1)]||x-y||_\infty < \varepsilon r/(M+1)$이다. 이상을 종합하면 임의의 $y \in B(x, r)$에 대해 $||(\Psi_x^{-1} \circ \Phi)(y) - y||_\infty = ||(\Psi_x^{-1} \circ \Phi)(y) - (\Psi_x^{-1} \circ \Psi_x)(y)||_\infty \leq M||\Phi(y) - \Psi_x(y)||_\infty < \varepsilon r$이 성립하여 $||(\Psi_x^{-1} \circ \Phi)(y) - x||_\infty \leq ||(\Psi_x^{-1} \circ \Phi)(y) - y||_\infty + ||x-y||_\infty < (1+\varepsilon)r$이므로 $\Phi(y) \in \Psi_x(B(x, (1+\varepsilon)r))$이고, 곧 $\Phi(B(x, r)) \subseteq \Psi_x(B(x, (1+\varepsilon)r))$임을 안다.

다른 포함관계를 보이기 위해 $x \in K$와 $r > 0$, $0 < \varepsilon < 1$을 위에서와 같이 두고 임의의 $y \in S(x, r)$를 생각하면 $||x-y||_\infty = r$이므로 $||(\Psi_x^{-1} \circ \Phi)(y) - x||_\infty \geq ||x-y||_\infty - ||(\Psi_x^{-1} \circ \Phi)(y) - y||_\infty > (1-\varepsilon)r$에서 $\Phi(y) \notin \Psi_x(B(x, (1-\varepsilon)r))$이다. 이는 곧 $\Psi_x(B(x, (1-\varepsilon)r))$과 $\Phi(S(x, r))$이 서로소임을 뜻하므로 집합 $A = \Psi_x(B(x, (1-\varepsilon)r)) \cap \Phi(B(x, r))$, $B = \Psi_x(B(x, (1-\varepsilon)r)) \setminus \Phi(B(x, r))$를 생각하면 $A \sqcup B = \Psi_x(B(x, (1-\varepsilon)r))$이며 A는 열린집합임이 자명하고 $B = \Psi_x(B(x, (1-\varepsilon)r)) \setminus [\Phi(B(x, r)) \cup \Phi(S(x, r))] = \Psi_x(B(x, (1-\varepsilon)r)) \setminus \Phi(\overline{B(x, r)})$에서 B도 열린집합이다.[23] 그러나 $\Psi_x(B(x, (1-\varepsilon)r))$이 연결집합이므로 A, B 중 하나는 공집합인데, $\Psi_x(x) = \Phi(x) \in A$이므로 $B = \emptyset$이 되어 $A = \Psi_x(B(x, (1-\varepsilon)r))$이고, 곧 $\Psi_x(B(x, (1-\varepsilon)r)) \subseteq \Phi(B(x, r))$이 되어 7단계의 증명이 끝난다.

Step 8. 가측집합 $A \subseteq U$에 대해 $\Phi(A)$도 가측이다.

함수 Φ가 미분동형사상이므로 Φ^{-1}가 연속이고, 곧 가측이므로 임의의 Borel 집합 $B \subseteq U$에 대해 $\Phi(B) = (\Phi^{-1})^{-1}(B)$가 가측이다. 한편, 정리 ? ? , ? ? 로부터 임의의 가측집합 $A \subseteq U$에 대해 적당한 Borel 집합 $B, C \subseteq \mathbb{R}^n$가 존재하여 $B \subseteq A \subseteq C$이며 $C \setminus B$가 영집합이고, $C \setminus A \subseteq C \setminus B$이므로 $N := C \setminus A$도 영집합이다. 나아가 WLOG, 필요하다면 B, C를 각각 $B \cap U, C \cap U$로 바꾸어 $B, C \subseteq U$라 해도 된다. 그렇다면 $\Phi(A) = \Phi(C) \setminus \Phi(N)$에서 $\Phi(C)$는 앞선 결과로부터 가측이고, $\Phi(N)$은 위의 보조정리로부터 영집합이 되어 Lebesgue 측도 공간이 완비성을 가진다는 점에서 곧 가측이므로 $\Phi(A)$는 가측이다.

Step 9. 열린집합 $W \subseteq U$에 대해 $\lambda_n(\Phi(W)) = \int_W |\det \mathbf{D}\Phi| \, d\lambda_n$이다.

먼저 $W \subseteq K$인 특별한 경우를 생각하자. 임의의 $\varepsilon > 0$을 택하고 $k \geq 1$를 충분히 크게 잡으면 $\log|\det \mathbf{D}\Phi(\cdot)| : K \to \mathbb{R}$가 균등연속이므로 적당한 $\delta > 0$가 존재하여 임의의

$x, y \in K$가 $||x-y||_\infty < \delta$를 만족하면 $|\log|\det \mathbf{D}\Phi(x)/\det \mathbf{D}\Phi(y)|| = |\log|\det \mathbf{D}\Phi(x)| - \log|\det \mathbf{D}\Phi(y)|| < \varepsilon/k$이고 곧 $1-\varepsilon \le 1-\varepsilon/k \le e^{-\varepsilon/k} < |\det \mathbf{D}\Phi(x)/\det \mathbf{D}\Phi(y)| < e^{\varepsilon/k} \le 1+\varepsilon$이다. 또한, 7단계의 결과로부터 WLOG, 필요하다면 $\delta > 0$를 더 작게 하여 임의의 열린 ball $B(x,r) \subseteq K$이 $r < \delta$를 만족하면 $\Psi_x(B(x,(1-\varepsilon)r)) \subseteq \Phi(B(x,r)) \subseteq \Psi_x(B(x,(1+\varepsilon)r))$이도록 할 수 있다. 나아가, $\mathscr{P}(W)$에 속하는 적당한 서로소인 열린 ball의 열 $\{B(x_i,r_i)\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $r_i < \delta$이고 $\lambda_n(W \setminus \bigsqcup_{i=1}^\infty B(x_i,r_i)) = 0$이며,[24] 위의 보조정리로부터 $\lambda_n(\Phi(W) \setminus \bigsqcup_{i=1}^\infty \Phi(B(x_i,r_i))) = \lambda_n(\Phi(W \setminus \bigsqcup_{i=1}^\infty B(x_i,r_i))) = 0$이다. 이상을 종합하면 Affine 변환이 good이라는 점과 따름정리 ? ? , 정리 ? ? 로부터 각 $i \in \mathbb{N}$에 대해

$$\begin{aligned}
\lambda_n(\Psi_{x_i}(B(x_i,(1+\varepsilon)r_i))) &= \int_{B(x_i,(1+\varepsilon)r_i)} |\det \mathbf{D}\Psi_{x_i}|\, d\lambda_n \\
&= |\det \mathbf{D}\Phi(x_i)|\lambda_n(B(x_i,(1+\varepsilon)r_i)) \\
&= |\det \mathbf{D}\Phi(x_i)|(1+\varepsilon)^n \lambda_n(B(x_i,r_i)) \\
&\le (1+\varepsilon)^{n+1} \int_{B(x_i,r_i)} |\det \mathbf{D}\Phi|\, d\lambda_n
\end{aligned}$$

이고, 이로부터

$$\begin{aligned}
\lambda_n(\Phi(W)) &= \lambda_n\left(\bigsqcup_{i=1}^\infty \Phi(B(x_i,r_i))\right) \\
&= \sum_{i=1}^\infty \lambda_n(\Phi(B(x_i,r_i))) \\
&\le \sum_{i=1}^\infty \lambda_n(\Psi_{x_i}(B(x_i,(1+\varepsilon)r_i))) \\
&\le (1+\varepsilon)^{n+1} \sum_{i=1}^\infty \int_{B(x_i,r_i)} |\det \mathbf{D}\Phi|\, d\lambda_n \\
&= (1+\varepsilon)^{n+1} \int_{\bigsqcup_{i=1}^\infty B(x_i,r_i)} |\det \mathbf{D}\Phi|\, d\lambda_n \\
&= (1+\varepsilon)^{n+1} \int_W |\det \mathbf{D}\Phi|\, d\lambda_n
\end{aligned}$$

이다. 이제 $(1-\varepsilon)^{n+1}\int_W |\det \mathbf{D}\Phi|\, d\lambda_n \le \lambda_n(\Phi(W))$임을 이와 비슷하게 보일 수 있으므로 곧 $\lambda_n(\Phi(W)) = \int_W |\det \mathbf{D}\Phi|\, d\lambda_n$임을 안다.

이제 일반적인 $W \subseteq K$에 대해 $\mathscr{P}(U)$에 속하는 적당한 서로소인 열린 ball의 열 $\{B_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $\overline{B_i} \subseteq U$이고 $\lambda_n(U \setminus \bigsqcup_{i=1}^\infty B_i) = 0$이다.[25] 따라서 위의 보조정리로부터 $\lambda_n(\Phi(W) \setminus \bigsqcup_{i=1}^\infty \Phi(W \cap B_i)) = \lambda_n(\Phi(W \setminus \bigsqcup_{i=1}^\infty B_i)) \le \lambda_n(\Phi(U \setminus \bigsqcup_{i=1}^\infty B_i)) = 0$이고, 각 $i \in \mathbb{N}$에 대해 $W \cap B_i \subseteq \overline{B_i} \subseteq U$이므로 앞선 결과에서 $K \subseteq U$가 임의의 compact한 집합이었음을 상기한다면

$$
\begin{aligned}
\lambda_n(\Phi(W)) &= \lambda_n\left(\bigsqcup_{i=1}^{\infty}\Phi(W\cap B_i)\right) \\
&= \sum_{i=1}^{\infty}\lambda_n(\Phi(W\cap B_i)) \\
&= \sum_{i=1}^{\infty}\int_{W\cap B_i}|\det \mathbf{D}\Phi|\,d\lambda_n \\
&= \int_{\bigsqcup_{i=1}^{\infty}(W\cup B_i)}|\det \mathbf{D}\Phi|\,d\lambda_n \\
&= \int_W |\det \mathbf{D}\Phi|\,d\lambda_n
\end{aligned}
$$

이다.

Step 10. 닫힌집합 $F\subseteq U$에 대해 $\lambda_n(\Phi(F)) = \int_F |\det \mathbf{D}\Phi|\,d\lambda_n$이다.

만약 $\Phi(F)$가 유계라면 적당한 열린집합 $W\subseteq U$가 존재하여 $\Phi(W)$가 유계이며 $\Phi(F)\subseteq \Phi(W)$이고,[26] 9단계의 결과로부터 $\lambda_n(\Phi(W\setminus F)) = \int_{W\setminus F}|\det \mathbf{D}\Phi|\,d\lambda_n < \infty$이고 $\lambda_n(\Phi(W)) = \int_W |\det \mathbf{D}\Phi|\,d\lambda_n < \infty$이므로

$$
\begin{aligned}
\lambda_n(\Phi(F)) &= \lambda_n(\Phi(W)) - \lambda_n(\Phi(W\setminus F)) \\
&= \int_W |\det \mathbf{D}\Phi|\,d\lambda_n - \int_{W\setminus F}|\det \mathbf{D}\Phi|\,d\lambda_n \\
&= \int_F |\det \mathbf{D}\Phi|\,d\lambda_n
\end{aligned}
$$

이다. 이제 $\Phi(F)$가 유계가 아닌 경우에는 집합열 $\{F_i\}$를 $F_i := F\cap\Phi^{-1}(B(i))$로 두면 각 $i\in\mathbb{N}$에 대해 $\Phi(F_i) = \Phi(F)\cap B(i)$에서 $\Phi(F_i)$는 유계이고, $F_i\uparrow F$임이 자명하다. 한편, $\Phi(F_i)\uparrow\Phi(F)$임도 분명하므로 앞선 결과와 정리 ? ? 의 i, ? ? 의 i로부터 $\lambda_n(\Phi(F)) = \lim_{i\to\infty}\lambda_n(\Phi(F_i)) = \lim_{i\to\infty}\int_{F_i}|\det \mathbf{D}\Phi|\,d\lambda_n = \int_F |\det \mathbf{D}\Phi|\,d\lambda_n$이다.

증명을 끝내기 위해, 임의의 가측집합 $A\subseteq U$를 생각하면 Lebesgue 측도가 regular하므로

$$
\begin{aligned}
\lambda_n(\Phi(A)) &= \inf\{\lambda_n(W)\in\overline{\mathbb{R}}_0^+ : \Phi(A)\subseteq W\subseteq\mathbb{R}^n, W\text{는 열린집합}\} \\
&= \inf\{\lambda_n(\Phi(W))\in\overline{\mathbb{R}}_0^+ : A\subseteq W\subseteq U, W\text{는 열린집합}\} \\
&= \inf\left\{\int_W |\det \mathbf{D}\Phi|\,d\lambda_n\in\overline{\mathbb{R}}_0^+ : A\subseteq W\subseteq U, W\text{는 열린집합}\right\} \\
&\geq \int_A |\det \mathbf{D}\Phi|\,d\lambda_n
\end{aligned}
$$

이고 비슷하게

$$
\lambda_n(\Phi(A)) = \sup\{\lambda_n(K)\in\overline{\mathbb{R}}_0^+ : K\subseteq\Phi(A), K\text{는 compact 집합}\}
$$

$$
\begin{aligned}
&= \sup\{\lambda_n(\Phi(K)) \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K\text{는 compact 집합}\} \\
&= \sup\left\{ \int_K |\det \mathbf{D}\Phi|\, d\lambda_n \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K\text{는 compact 집합} \right\} \\
&\leq \int_A |\det \mathbf{D}\Phi|\, d\lambda_n
\end{aligned}
$$

이 되어 곧 $\lambda_n(\Phi(A)) = \int_A |\det \mathbf{D}\Phi|\, d\lambda_n$이다. 이제 1단계와 8단계로부터 임의의 $\mathscr{C}^1$급 미분동형사상이 good임을 안다.

한편, V에서 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$의 경우에는 이를 각각 양과 음의 부분으로 나누어 앞선 결과를 적용하면 $(f \circ \Phi)|\det \mathbf{D}\Phi|$는 U에서 적분가능하며 $\int_V f\, d\lambda_n = \int_U (f \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n$임을 쉽게 알 수 있고, $(f \circ \Phi)|\det \mathbf{D}\Phi|$가 U에서 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

여기서 한 발 더 나아가 Sard의 정리와 INFT로부터 변수변환 공식에서의 Φ의 조건을 $\mathscr{C}^1$급 전단사함수로 완화할 수 있다.

Theorem 1.173 (Sard) 열린집합 $U \subseteq \mathbb{R}^n$에서 정의된 $\mathscr{C}^1$급 단사함수 $f : U \to \mathbb{R}^n$의 모든 임계점의 집합 C에 대해 $f(C)$는 영집합이다.

Proof 먼저 $U = (0,1)^n$인 특별한 경우를 생각하자. 가정으로부터 f가 $\mathscr{C}^1$급이므로 곧 Lipschitz 연속이며, 이때의 Lipschitz 상수를 $L > 0$이라 하고 임의의 $x \in U$에 대해 함수 $T_x : \mathbb{R}^n \to \mathbb{R}^n$를 $T_x : y \mapsto f(x) + \mathbf{D}f(x)(y - x)$로 두자. 증명은 크게 2단계로 이루어진다.

Step 1. 임의의 $x, y \in U$에 대해 $||f(y) - T_x(y)|| = o(||x - y||)$이다.

서로다른 $x, y \in U$를 임의로 택하면 MVT에서 x와 y를 잇는 선분 위에 적당한 $z_1, \cdots, z_n \in U$이 존재하여 각 $i \leq n$에 대해 $f_i(y) - T_x^i(y) = f_i(y) - f_i(x) - \mathbf{D}f_i(x)(y - x) = [\mathbf{D}f_i(z_i) - \mathbf{D}f_i(x)](y - x)$이다. 이는 각 $i \leq n$에 대해 $|f_i(y) - T_x^i(y)| = |[\mathbf{D}f_i(z_i) - \mathbf{D}f_i(x)](y - x)| \leq ||\mathbf{D}f_i(z_i) - \mathbf{D}f_i(x)||_{\mathrm{op}}||x - y||$임을 뜻하므로 $||f(y) - T_x(y)||/||x - y|| \leq \sum_{i=1}^n |f_i(y) - T_x^i(y)|/||x - y|| \leq \sum_{i=1}^n ||\mathbf{D}f_i(z_i) - \mathbf{D}f_i(x)||_{\mathrm{op}}$이고, 곧 f가 $\mathscr{C}^1$급이라는 사실로부터 $||f(y) - T_x(y)|| = o(||x - y||)$임을 안다.

Step 2. 임의의 $\varepsilon > 0$에 대해 적당한 $R > 0$이 존재하여 임의의 $0 < r < R$에 대해 다음이 성립한다: 임의의 $x \in C$에 대해 적당한 box $B \subseteq \mathbb{R}^n$가 존재하여 $\lambda_n(B) = \varepsilon r^n$이고 $f(B(x, r)) \subseteq B$이다.

1단계의 결과로부터 적당한 $R > 0$이 존재하여 $0 < r < R$이면 $||x - y|| < r$인 임의의 $x, y \in U$에 대해 $||f(y) - T_x(y)|| \leq (\varepsilon/L^{n-1})||x - y|| < \varepsilon r/L^{n-1}$이며 $||f(x) - f(y)|| < L||x - y|| < Lr$이다. 이제 임의의 $x \in C$를 택하면 적당한 $n - 1$차원 부분공간 $\mathsf{V} < \mathbb{R}^n$가 존재하여 $\mathrm{im}\, T_x \leq \mathsf{V}$이고, 곧 $||x - y|| < r$인 임의의 $y \in U$에 대해 $d(f(y), \mathsf{V}) \leq ||f(y) - T_x(y)|| < \varepsilon r/L^{n-1}$이다. 따라서 $f(x)$를 중심으로 하고 각 모서리의 길이가 Lr로 같은 $n - 1$차원 box $B' \subseteq \mathsf{V}$와 이를

V에 수직인 방향으로 $\pm\varepsilon r/2L^{n-1}$씩 평행이동시켜 얻는 n차원 box $B\subseteq\mathbb{R}^n$를 생각하면 임의의 $y\in B(x,r)\cap U$에 대해 $f(y)\in B$임이 자명하여 곧 $f(B(x,r))\subseteq B$이고 이때 $\lambda_n(B)=\varepsilon r^n$이다. (이는 적당한 Affine 변환에 변수변환 공식을 적용한 결과로, 직관적으로도 자명하다. 이제부터는 이 정도의 변수변환 공식의 가벼운 응용은 특별한 언급 없이 사용하도록 한다.)

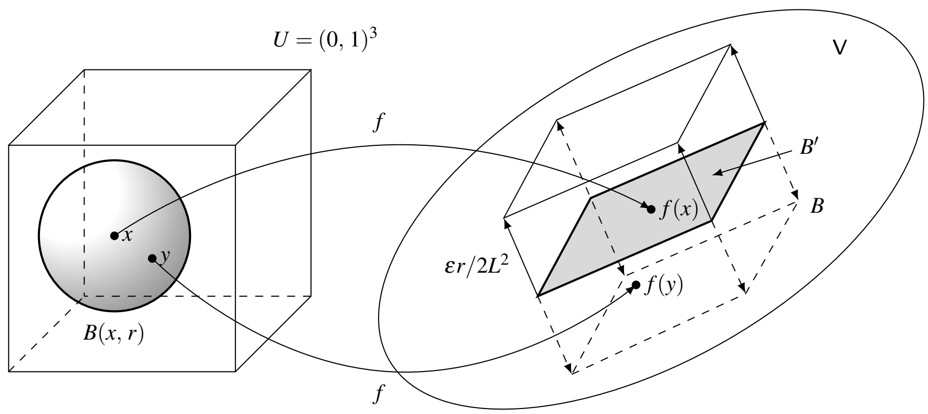


Figure 1.13 Sard의 정리의 증명에서의 집합 B'와 B.

증명을 끝내기 위해 임의의 $\varepsilon>0$을 택하면 2단계의 결과로부터 적당한 $0<r<1$이 존재하여 다음이 성립한다: 임의의 $x\in C$에 대해 적당한 box $B\subseteq\mathbb{R}^n$가 존재하여 $\lambda_n(B)=\varepsilon r^n/(\sqrt{n}+1)^n$이고 $f(B(x,r))\subseteq B$이다. 이제 $N=\lceil\sqrt{n}/r\rceil$과 임의의 $j_1,\cdots,\ j_n\le N$에 대해 $B_{j_1\cdots j_n}=\prod_{i=1}^n((j_i-1)/N,\ j_i/N)$라 하고, 이 중에 C의 원소를 적어도 하나 포함하는 것만을 택하여 $B_1,\cdots,B_k$ $(k\le N^n)$라 하자. 그렇다면 각 $j\le k$에 대해 $x_j\in B_j\cap C$를 택할 수 있으므로 적당한 box $\widetilde{B}_j\subseteq\mathbb{R}^n$가 존재하여 $\lambda_n(\widetilde{B}_j)=\varepsilon r^n/(\sqrt{n}+1)^n$이고 $B_j\subseteq B(x_j,\sqrt{n}/N)\subseteq B(x_j,r)$에서 $f(B_j)\subseteq f(B(x_j,r))\subseteq\widetilde{B}_j$이다. 이제 $f(C)\subseteq\bigcup_{j=1}^k f(B_j)\cup\bigcup_{j_1,\cdots,j_n=1}^N f(\partial B_{j_1\cdots j_n})\subseteq\bigcup_{j=1}^k\widetilde{B}_j\cup\bigcup_{j_1,\cdots,j_n=1}^N f(\partial B_{j_1\cdots j_n})$인데, 보조정리 ? ? , ? ? 에서 $\bigcup_{j_1,\cdots,j_n=1}^N f(\partial B_{j_1\cdots j_n})$가 영집합이므로

$$\begin{aligned}\rho_n^*(f(C))&\le\lambda_n\left(\bigcup_{j=1}^k\widetilde{B}_j\cup\bigcup_{j_1,\cdots,j_n=1}^N f(\partial B_{j_1\cdots j_n})\right)\\&\le\lambda_n\left(\bigcup_{j=1}^k\widetilde{B}_j\right)\\&\le\sum_{j=1}^k\lambda_n(\widetilde{B}_j)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{k} \frac{\varepsilon r^n}{(\sqrt{n}+1)^n} \\
&\leq N^n \frac{\varepsilon r^n}{(\sqrt{n}+1)^n} \\
&< \left(\frac{\sqrt{n}}{r}+1\right)^n \frac{\varepsilon r^n}{(\sqrt{n}+1)^n} \\
&= \left(\frac{\sqrt{n}+r}{\sqrt{n}+1}\right)^n \varepsilon \\
&< \varepsilon
\end{aligned}
$$

에서 $\rho_n^*(f(C)) = 0$이다. 이는 곧 임의의 $S \subseteq \mathbb{R}^n$에 대해 $\rho_n^*(S \cap f(C)) + \rho_n^*(S \cap [f(C)]^c) \leq \rho_n^*(f(C)) + \rho_n^*(S \cap [f(C)]^c) = \rho_n^*(S \cap [f(C)]^c) \leq \rho_n^*(S)$임을 뜻하여 명제 ? ? 로부터 $f(C)$가 영집합임을 안다.

이제 f를 적당히 scaling하고 평행이동함으로써 이상의 결론이 U가 임의의 유계인 열린 box인 경우에도 성립함을 알 수 있다. 한편, 일반적인 열린집합 $U \subseteq \mathbb{R}^n$를 생각하면 이는 적당한 유계인 열린 box $B_1, B_2, \cdots$에 대해 $U = \bigcup_{i=1}^k B_i$로 표현되므로 (각주 ? ? 참조.) $f(C) = f(\bigcup_{i=1}^k (C \cap B_i)) = \bigcup_{i=1}^k f(C \cap B_i) = \bigcup_{i=1}^k f|_{B_i}(C)$에서 $f(C)$가 영집합임을 안다. (여기서 k는 유한할 수도, ∞일 수도 있다.) □

Corollary 1.174 (Transformation formula) 열린집합 $U, V \subseteq \mathbb{R}^n$ 사이에서 정의된 $\mathscr{C}^1$급 전단사함수 $\Phi : U \to V$와 음이 아닌 가측함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$에 대해 $(f \circ \Phi)|\det \mathbf{D}\Phi| : U \to \overline{\mathbb{R}}_0^+$는 가측이고 $\int_V f\, d\lambda_n = \int_U (f \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n$이다. 한편, V에서 적분가능한 (혹은 $(f \circ \Phi)|\det \mathbf{D}\Phi|$가 U에서 적분가능한) $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며, 이때 $(f \circ \Phi)|\det \mathbf{D}\Phi|$는 U에서 적분가능하다 (혹은 f는 V에서 적분가능하다).

Proof 함수 Φ의 모든 임계점의 집합 $C = (\det \mathbf{D} f)^{-1}(0)$를 생각하면 $\det \mathbf{D} f : U \to \mathbb{R}$가 연속이므로 이는 U에 대해 닫혀있다. 이로부터 $W = U \setminus C$는 열려있으며 INFT에서 $\Phi(W)$도 열려있고 $\Phi|_W : W \to \Phi(W)$는 $\mathscr{C}^1$급 미분동형사상이다. 그렇다면 Sard의 정리로부터 $\Phi(C)$가 영집합이고, 앞서 증명한 변수변환 공식으로부터 $(f \circ \Phi)|\det \mathbf{D}\Phi| = (f \circ \Phi|_W)\mathbf{1}_W$는 가측이며 $\int_V f\, d\lambda_n = \int_{V \setminus \Phi(C)} f\, d\lambda_n = \int_{\Phi(W)} f\, d\lambda_n = \int_W (f \circ \Phi|_W)|\det \mathbf{D}\Phi|_W|\, d\lambda_n = \int_U (f \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n$이다.

한편, V에서 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$의 경우에는 이를 각각 양과 음의 부분으로 나누어 앞선 결과를 적용하면 $(f \circ \Phi)|\det \mathbf{D}\Phi|$는 U에서 적분가능하며 $\int_V f\, d\lambda_n = \int_U (f \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n$임을 쉽게 알 수 있고, $(f \circ \Phi)|\det \mathbf{D}\Phi|$가 U에서 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

이제 변수변환 공식으로부터 쉽게 얻어지는 따름정리들을 간단히 소개하는 것으로 이번 절을 마무리하자. 첫 번째로 소개할 따름정리는 미적분학에서 배우는 직교좌표계에서 극좌표계, 원통좌표계, 구면좌표계로의 변수변환 공식이다.

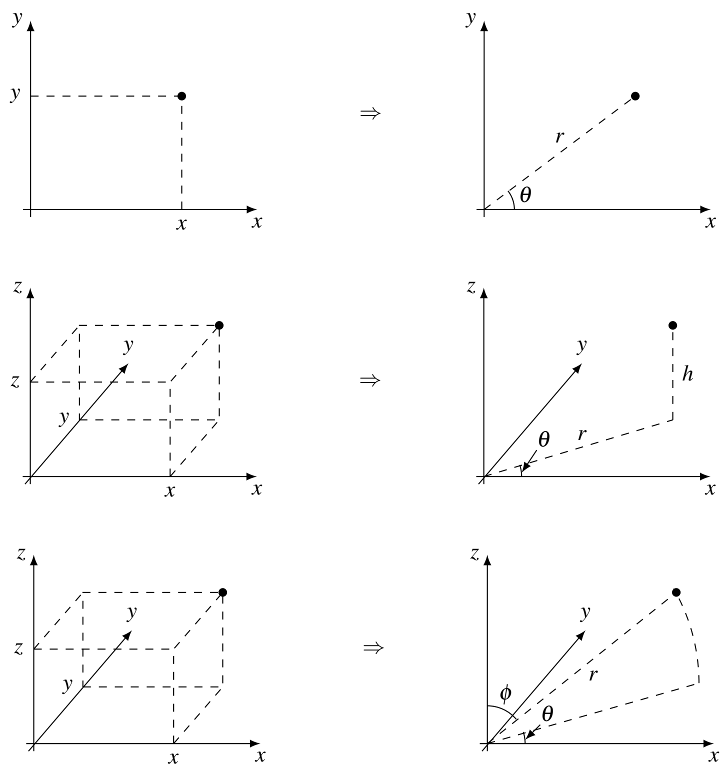


Figure 1.14 기본적인 직교좌표계에서 극좌표계(위), 원통좌표계(중간), 구면좌표계(아래)로의 변환.

Corollary 1.175 음이 아닌 가측함수 $f : \mathbb{R}^2 \to \overline{\mathbb{R}}_0^+$, $g : \mathbb{R}^3 \to \overline{\mathbb{R}}_0^+$에 대해 다음이 성립한다.

i. (극좌표계 변환)

$$\int_{\mathbb{R}^2} f(x,y)\,d\lambda_2(x,y) = \int_{\mathbb{R}_0^+ \times [0,2\pi)} f(r\cos\theta, r\sin\theta)\,r\,d\lambda_2(r,\theta)$$

ii. (원통좌표계 변환)

$$\int_{\mathbb{R}^3} g(x,y,z)\,d\lambda_3(x,y,z) = \int_{\mathbb{R}_0^+ \times [0,2\pi) \times \mathbb{R}} g(r\cos\theta, r\sin\theta, h)\,r\,d\lambda_3(r,\theta,h)$$

iii. (구면좌표계 변환)

$$\int_{\mathbb{R}^3} g(x,y,z)\,d\lambda_3(x,y,z) \\ = \int_{\mathbb{R}_0^+\times[0,\pi]\times[0,2\pi)} g(r\sin\theta\cos\phi,\, r\sin\theta\sin\phi,\, r\cos\theta)r^2\sin\theta\,d\lambda_3(r,\theta,\phi)$$

한편, 적분가능한 $f:\mathbb{R}^2\to\overline{\mathbb{R}}$, $g:\mathbb{R}^3\to\overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립한다.

PROOF 함수

$$\Phi_1:\mathbb{R}^+\times(0,2\pi)\to\mathbb{R}^2\setminus(\mathbb{R}_0^+\times\{0\})$$
$$\Phi_2:\mathbb{R}^+\times(0,2\pi)\times\mathbb{R}\to\mathbb{R}^3\setminus(\mathbb{R}_0^+\times\{0\}\times\mathbb{R})$$
$$\Phi_3:\mathbb{R}^+\times(0,\pi)\times(0,2\pi)\to\mathbb{R}^3\setminus(\mathbb{R}_0^+\times\{0\}\times\mathbb{R})$$

을 각각

$$\Phi_1:(r,\theta)\mapsto(r\cos\theta,\, r\sin\theta)$$
$$\Phi_2:(r,\theta,h)\mapsto(r\cos\theta,\, r\sin\theta,\, h)$$
$$\Phi_3:(r,\theta,\phi)\mapsto(r\sin\theta\cos\phi,\, r\sin\theta\sin\phi,\, r\cos\theta)$$

로 정의하면 이들이 $\mathscr{C}^1$급 미분동형사상임을 쉽게 확인할 수 있고,

$$\det\mathbf{D}\Phi_1(r,\theta)=\det\begin{bmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{bmatrix}=r$$
$$\det\mathbf{D}\Phi_2(r,\theta,h)=\det\begin{bmatrix}\cos\theta & -r\sin\theta & 0\\ \sin\theta & r\cos\theta & 0\\ 0 & 0 & 1\end{bmatrix}=r$$
$$\det\mathbf{D}\Phi_3(r,\theta,\phi)=\det\begin{bmatrix}\sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi\\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi\\ \cos\theta & -r\sin\theta & 0\end{bmatrix}=r^2\sin\theta$$

이므로 변수변환 공식과 보조정리 ? ? 로부터 정리가 자명하다. □

다음은 변수변환 공식의 Borel 측도공간 버전이다.

Corollary 1.176 열린집합 $U,V\subseteq\mathbb{R}^n$ 사이에서 정의된 $\mathscr{C}^1$급 전단사함수 $\Phi:U\to V$와 음이 아닌 Borel 함수 $f:\mathbb{R}^n\to\overline{\mathbb{R}}_0^+$에 대해 합성 $f\circ\Phi:U\to\overline{\mathbb{R}}_0^+$는 Borel이고 $\int_V f\,d\mu_n=\int_U(f\circ\Phi)|\det\mathbf{D}\Phi|\,d\mu_n$이다. 한편, V에서 μ_n−적분가능한 (혹은 $(f\circ\Phi)|\det\mathbf{D}\Phi|$가 U에서 μ_n−적분가능한) $f:\mathbb{R}^n\to\overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며, 이때 $(f\circ\Phi)|\det\mathbf{D}\Phi|$는 U에서 μ_n−적분가능하다 (혹은 f는 V에서 μ_n−적분가능하다).

Proof 이는 변수변환 공식과 정리 ? ? , ? ? 로부터 자명하다. □

세 번째는 보조정리 ? ? 의 일반화로, 직관적으로도 꽤 자명하다.

Corollary 1.177 모든 부분공간 $\mathsf{V} < \mathbb{R}^n$는 영집합이다.

Proof 영집합의 정의로부터 $\dim \mathsf{V} = n - 1$인 경우만 생각하면 되고, 이때에 V는 $\mathbb{R}^n$에서의 초평면이므로 만약 이가 $\mathbb{R}^n$의 좌표축 중 어느 하나와 직교한다면 보조정리 ? ? 로부터 정리가 자명하다. 만약 V와 직교하는 좌표축이 존재하지 않는다면, 적당한 가역인 선형사상 $T : \mathbb{R}^n \to \mathbb{R}^n$가 존재하여 $T(\mathsf{V})$는 $\mathbb{R}^n$의 좌표축 중 어느 하나와 직교하는 초평면이고 $|\det T| = 1$이므로[27]변수변환 공식과 앞선 결과로부터 $\lambda_n(T(\mathsf{V})) = \int_{\mathbb{R}^n} \mathbf{1}_{T(\mathsf{V})}\, d\lambda_n = \int_{\mathbb{R}^n} (\mathbf{1}_{T(\mathsf{V})} \circ T) |\det T|\, d\lambda_n = \int_{\mathbb{R}^n} \mathbf{1}_{\mathsf{V}}\, d\lambda_n = \lambda_n(\mathsf{V})$에서 V가 영집합임을 안다. □

마지막 따름정리는 선형사상의 행렬식과 측도로써 정의되는 넓이의 개념을 아름답게 연결해준다.

Corollary 1.178 가측집합 $A \subseteq \mathbb{R}^n$와 선형사상 $T : \mathbb{R}^n \to \mathbb{R}^n$에 대해 $T(A)$도 가측이고, $\lambda_n(T(A)) = |\det T| \lambda_n(A)$이다.

Proof 만약 T가 가역이라면 변수변환 공식으로부터 $\lambda_n(A) = \int_{T(A)} |\det T^{-1}|\, d\lambda_n = \lambda_n(T(A))/|\det T|$가 되어 정리가 자명하다. 한편, T가 가역이 아니라면 $\operatorname{im} T < \mathbb{R}^n$이므로 따름정리 ? ? 로부터 $\lambda_n(T(A)) \le \lambda_n(\operatorname{im} T) = 0 = |\det T| \lambda_n(A)$이다. □

1.13 Fundamental Theorem of Calculus

이번 절에서는 FTC로 돌아간다. FTC는 적분과 미분을 이어주는 다리 역할을 하는 정리로서 미적분학의 근간을 이루는 중요한 정리이다. 먼저, 우리가 1학년 미적분학 시간에 배운 기본적인 FTC를 보자.

Theorem (Fundamental theorem of calculus) 다음이 성립한다.

i. 적분가능한 연속함수 $f : \mathbb{R} \to \mathbb{R}$에 대해 함수 $F : \mathbb{R} \to \mathbb{R}$를 $F : x \mapsto \int_{-\infty}^{x} f\, d\lambda_1$로 두면 이는 f의 역도함수이다. 즉, $F' = f$이다.
ii. 함수 $F : \mathbb{R} \to \mathbb{R}$가 $\mathscr{C}^1$급이고 그 도함수가 적분가능하면 $f := F'$에 대해 $F(x) = \int_{-\infty}^{x} f\, d\lambda_1$이다.

쉽게 생각하면, i은 적분한 뒤 미분하면 원래의 함수를 얻는다는 의미이고, ii는 반대로 미분한 뒤 적분해도 원래의 함수를 얻는다는 의미로, 이 둘은 서로 필요충분과 비슷한 관계이다. 우리는 하도 많이 쓰는 까닭에 자명하게까지 느껴지는 이 FTC에 한 가지 간단하지

만, 중요한 질문을 던진다. 만약 여기서 f가 연속이 아니면 어떻게 될까? 지금부터 우리는 이에 대해 답을 찾아나선다.

먼저 i을 보면, 함수 F를 정의할 수는 있어야 하므로 f가 적분가능하다는 조건은 필수적이다. 그러나 f가 적분가능하다는 조건만으로는 F의 미분가능성을 보장할 수가 없다. 실제로 적분가능하지만 F가 특정 점에서 미분불가능한 f를 쉽게 찾을 수 있다.[28]하지만 이렇게 허무하게 끝날 것이었으면 애초에 시작하지도 않았을 것이다. 놀랍게도, 비록 F가 모든 점에서 미분가능하지는 않지만, 이가 거의 대부분의 점에서 '미분 비슷한 무언가'를 할 수 있음을 보일 수 있다. 다만, 이를 증명하기 위해 준비해야 할 것들이 꽤나 많다. 먼저 비교적 간단한 적분가능한 함수의 근사부터 시작하자.

Theorem 1.179 적분가능한 $f:\mathbb{R}^n\to\overline{\mathbb{R}}$와 임의의 $\varepsilon>0$에 대해 다음이 성립한다.

i. 단순함수 $g:\mathbb{R}^n\to\mathbb{R}$가 존재하여 이는 적당한 $a_1,\cdots,a_k\in\mathbb{R}$와 유계인 $B_1,\cdots,B_k\in\mathscr{S}_n$에 대해 $g=\sum_{i=1}^k a_i\mathbf{1}_{B_i}$로 쓸 수 있으며 $\int_{\mathbb{R}^n}|f-g|\,d\lambda_n<\varepsilon$이다.
ii. 적분가능한 연속함수 $g:\mathbb{R}^n\to\mathbb{R}$가 존재하여 $\int_{\mathbb{R}^n}|f-g|\,d\lambda_n<\varepsilon$이다.

Proof i. 먼저 f가 음이 아닌 특별한 경우를 생각하면 정리 ? ? 로부터 적당한 음이 아닌 단순함수 $f_i:\mathbb{R}^n\to\mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f-f_i\downarrow 0$이고 각 $i\in\mathbb{N}$에 대해 $f-f_i\le f$이다. 그렇다면 DCT로부터 $\int_{\mathbb{R}^n}|f-f_i|\,d\lambda_n\to 0$이고, 곧 임의의 $\varepsilon>0$을 택하면 충분히 큰 $i_0\in\mathbb{N}$에 대해 $\int_{\mathbb{R}^n}|f-f_{i_0}|\,d\lambda_n<\varepsilon/2$이다.

표기의 편의를 위해 $\widetilde{f}:=f_{i_0}$로 두자. 함수 $\widetilde{f}$의 표준형을 $\sum_{j=1}^l a_j\mathbf{1}_{A_j}$라 하고, WLOG, 모든 a_j가 0이 아니라 하면 정리 ? ? , ? ? 로부터 각 $j\le l$에 대해 적당한 $B_j,C_j\in\mathscr{B}_n$가 존재하여 $B_j\subseteq A_j\subseteq C_j$이고 $C_j\setminus B_j$가 영집합이므로 $A_j\setminus B_j$도 영집합이다. 나아가, 집합족 $\mathscr{S}_n$이 $\mathscr{B}_n$의 생성자이고 각 $j\le l$에 대해 $\mu_n(B_j)<\infty$임이 자명하므로(만약 어떤 $j_0\le l$에 대해 $\mu_n(B_{j_0})=\infty$이면 $\lambda_n(A_{j_0})=\infty$이고, 곧 $\int_{\mathbb{R}^n}f\,d\lambda_n\ge\int_{\mathbb{R}^n}\widetilde{f}\,d\lambda_n=\sum_{j=1}^l a_j\lambda_n(A_j)\ge a_{j_0}\lambda(A_{j_0})=\infty$가 되어 f가 적분가능하다는 가정에 모순된다.) 정리 ? ? 의 ii로부터 서로소인 $B_{j1},\cdots,B_{jm_j}\in\mathscr{S}_n$이 존재하여 $\mu_n(B_j\triangle\bigsqcup_{k=1}^{m_j}B_{jk})<\varepsilon/2a_jl$이다. 이제 단순함수 $g=\sum_{j=1}^l a_j\mathbf{1}_{\bigsqcup_{k=1}^{m_j}B_{jk}}=\sum_{j=1}^l\sum_{k=1}^{m_j}a_j\mathbf{1}_{B_{jk}}$, $h=\sum_{j=1}^l a_j\mathbf{1}_{B_j}$를 생각하면 이상으로부터

$$
\begin{aligned}
\int_{\mathbb{R}^n}|g-h|\,d\lambda_n&=\int_{\mathbb{R}^n}\left|\sum_{j=1}^l a_j(\mathbf{1}_{\bigsqcup_{k=1}^{m_j}B_{jk}}-\mathbf{1}_{B_j})\right|d\lambda_n\\
&\le\int_{\mathbb{R}^n}\sum_{j=1}^l a_j\mathbf{1}_{B_j\triangle\bigsqcup_{k=1}^{m_j}B_{jk}}\,d\lambda_n\\
&=\sum_{j=1}^l a_j\lambda_n\left(B_j\triangle\bigsqcup_{k=1}^{m_j}B_{jk}\right)\\
&<\sum_{j=1}^l\frac{\varepsilon}{2l}
\end{aligned}
$$

$$= \frac{\varepsilon}{2}$$

이고 $\widetilde{f} = h$ (ae.)이므로 따름정리 ? ? 의 ii로부터 $\int_{\mathbb{R}^n} |f-g|\,d\lambda_n \le \int_{\mathbb{R}^n} |f-\widetilde{f}|\,d\lambda_n + \int_{\mathbb{R}^n} |\widetilde{f}-g|\,d\lambda_n < \varepsilon/2 + \int_{\mathbb{R}^n} |g-h|\,d\lambda_n < \varepsilon$에서 정리가 성립한다.

한편, 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해서는 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 된다.

ii. 함수 f가 적분가능하므로 정리 ? ? 의 ii로부터 거의 어디서나 유한하고, WLOG, 필요하다면 영집합에서의 함숫값을 0으로 바꾸어 f가 항상 유한하다고 해도 된다. 그렇다면 i로부터 임의의 $\varepsilon > 0$을 택하면 단순함수 $h : \mathbb{R}^n \to \mathbb{R}$가 존재하여 이는 적당한 $a_1, \cdots, a_k \in \mathbb{R}$와 유계인 $B_1, \cdots, B_k \in \mathscr{S}_n$에 대해 $h = \sum_{i=1}^{k} a_i \mathbf{1}_{B_i}$로 쓸 수 있으며 $\int_{\mathbb{R}^n} |f-h|\,d\lambda_n < \varepsilon/2$이다. 이제 각 $i \le k$와 임의의 $j \in \mathbb{N}$에 대해 B_i의 모서리를 조금씩 줄여 (단, 이때 $2/j$ 이상으로는 줄이지 않는 것으로 하자.) 만든 semi-open box를 B_i'라 하고 $B_i \setminus B_i'$에서 h의 함숫값인 a_i를 0과 선형으로 보간하여 얻는 함수를 $g_j : \mathbb{R}^n \to \mathbb{R}$라 하면 $|g_j - h| \to 0$ (ae.)이고(각 $i \le k$에 대해 $j \to \infty$이면 $B_i' \uparrow B_i$이므로 함수열 $\{h_j\}$는 $\bigcup_{i=1}^{k} \partial B_i$를 제외한 모든 곳에서 g로 수렴한다. 그런데 주석 ? ? 에서와 같은 이유로 $\lambda_n(\bigcup_{i=1}^{k} \partial B_i) = 0$이므로 결국 $|g - h_j| \to 0$ (ae.)이다.) $|g_j - h| \le |g_j| + |h| \le 2|h|$이므로 DCT에서 $\int_{\mathbb{R}^n} |g_j - h|\,d\lambda_n \to 0$이다. 이는 곧 충분히 큰 $j_0 \in \mathbb{N}$를 택하면 $\int_{\mathbb{R}^n} |g_{j_0} - h|\,d\lambda_n < \varepsilon/2$임을 뜻하므로 $g := g_{j_0}$라 하면 $\int_{\mathbb{R}^n} |f-g|\,d\lambda_n \le \int_{\mathbb{R}^n} |f-h|\,d\lambda_n + \int_{\mathbb{R}^n} |g-h|\,d\lambda_n < \varepsilon$이다. 이때 g가 연속임이 분명하고 $\int_X |g|\,d\lambda_n \le \int_X |f-g|\,d\lambda_n + \int_X |f|\,d\lambda_n < \infty$에서 이가 적분가능하므로 정리가 성립한다. □

다음으로 필요한 것은 $\sigma-$대수와 측도의 pushfowarding이다. 이는 pushfoward라는 단어의 사전적 뜻 그대로 고정된 $\sigma-$대수나 측도에 대해 어떤 함수를 앞으로 '밀어넣어' 새로운 $\sigma-$대수나 측도를 얻는 방법이다. 엄밀한 정의를 보면 그 느낌을 더 잘 알 수 있을 것이다.

Definition 1.180 측도공간 $(X, \mathscr{A}, \mu)$와 집합 Y 사이에서 정의된 함수 $\varphi : X \to Y$를 생각하자. 이때 φ에 대한 $\mathscr{A}$의 **pushfoward $\sigma-$대수(− $\sigma-$algebra)**를 $\varphi_*\mathscr{A}$로 쓰고 $\varphi_*\mathscr{A} = \{A \subseteq Y : \varphi^{-1}(A) \in \mathscr{A}\}$로 정의한다. 비슷하게, φ에 대한 μ의 **pushfoward 측도(− measure)**를 $\varphi_*\mu : \varphi_*\mathscr{A} \to \overline{\mathbb{R}}_0^+$로 쓰고 $\varphi_*\mu : A \mapsto \mu(\varphi^{-1}(A))$로 정의한다.

물론, 이때의 정의가 well-define되는지 확인해줘야 한다.

Proposition 1.181 측도공간 $(X, \mathscr{A}, \mu)$와 집합 Y 사이에서 정의된 함수 $\varphi : X \to Y$에 대해 $(Y, \varphi_*\mathscr{A}, \varphi_*\mu)$는 측도공간을 이룬다. 따라서 pushfoward $\sigma-$대수와 측도는 well-define된다.

Proof 먼저 $\varphi_*\mathscr{A}$가 $\sigma-$대수임을 보이자. 이를 위해서는 정의 ? ? 의 조건을 하나하나 따져보면 된다. 그 중 몇 개만 연습삼아 살펴보면, $\emptyset \in \mathscr{A}$에서 $\emptyset = \varphi^{-1}(\emptyset) \in \varphi_*\mathscr{A}$에서 $\varphi_*\mathscr{A}$가

비어있지 않음이 자명하다. 또한 임의의 $A \in \varphi_* \mathscr{A}$에 대해 $\varphi^{-1}(A) \in \mathscr{A}$이므로 $\varphi^{-1}(A^c) = [\varphi^{-1}(A)]^c \in \mathscr{A}$에서 $A^c \in \varphi_* \mathscr{A}$이다. 이제 남은 조건들도 이와 같이 쉽게 따져볼 수 있다.

다음으로, $\varphi_* \mu$가 측도임을 보이자. 우선 $\varphi_* \mu(\emptyset) = \mu(\varphi^{-1}(\emptyset)) = \mu(\emptyset) = 0$임은 자명하고, $\varphi_* \mathscr{A}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 $\varphi_* \mu(\bigsqcup_{i=1}^{\infty} A_i) = \mu(\varphi^{-1}(\bigsqcup_{i=1}^{\infty} A_i)) = \mu(\bigsqcup_{i=1}^{\infty} \varphi^{-1}(A_i)) = \sum_{i=1}^{\infty} \mu(\varphi^{-1}(A_i)) = \sum_{i=1}^{\infty} \varphi_* \mu(A_i)$이므로 $\varphi_* \mu$가 측도임을 알고, 증명이 끝난다. □

위의 정의에서 알 수 있듯이 pushfowarding은 φ에게 아무런 조건을 요구하지 않는다. 연속일 필요도, 전단사여야 할 필요도 없고, 심지어 가측일 필요도 없다. 앞서 배운 음이 아닌 가측함수의 적분으로 새로운 측도를 유도하는 것과 비교하면 pushfowarding은 보다 일반적인 상황에서도 적용할 수 있는 기법인 셈이다. 이러한 pushfowarding은 나중에 확률변수를 배우며 다시 등장할 것이다. 여기서는 pushfowarding에 대한 측도론적인 성질들을 조금 더 살펴보자.

Theorem 1.182 가측공간 $(X, \mathscr{A})$와 집합 Y 사이에서 정의된 함수 $\varphi : X \to Y$와 Y와 가측공간 $(Z, \mathscr{B})$ 사이에서 정의된 함수 $f : Y \to Z$에 대해 f가 $\varphi_* \mathscr{A} / \mathscr{B}$−가측일 필요충분조건은 $f \circ \varphi$가 $\mathscr{A} / \mathscr{B}$−가측인 것이다.

Proof 충분조건임을 보이기 위해 임의의 $A \in \mathscr{B}$를 생각하면 가정으로부터 $f^{-1}(A) \in \varphi_* \mathscr{A}$이므로 $(f \circ \varphi)^{-1}(A) = \varphi^{-1}(f^{-1}(A)) \in \mathscr{A}$에서 $f \circ \varphi$가 $\mathscr{A} / \mathscr{B}$−가측이 되어 충분조건임이 보여진다. 이제 필요조건임도 이와 비슷하게 보일 수 있다. □

Theorem 1.183 (Transformation formula) 측도공간 $(X, \mathscr{A}, \mu)$와 집합 Y 사이에서 정의된 함수 $\varphi : X \to Y$에 대해 음이 아닌 함수 $f : Y \to \overline{\mathbb{R}}_0^+$가 $\varphi_* \mathscr{A} / \mathscr{B}_1$−가측이라 하자. 그렇다면 합성 $f \circ \varphi : X \to \overline{\mathbb{R}}_0^+$는 가측이고 $\int_Y f \, d\varphi_* \mu = \int_X f \circ \varphi \, d\mu$이다. 한편, $\varphi_* \mu$−적분가능한(혹은 $f \circ \varphi$가 μ−적분가능한 가측함수) $f : Y \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며 이때 $f \circ \varphi$는 μ−적분가능하다(혹은 f는 $\varphi_* \mu$−적분가능하다).

Proof 먼저 f가 음이 아닌 단순함수인 경우를 생각하여 이의 표준형을 $f = \sum_{i=1}^{k} a_i \mathbf{1}_{A_i}$라 하자. 그렇다면 $f \circ \varphi = \sum_{i=1}^{k} a_i \mathbf{1}_{\varphi^{-1}(A_i)}$에서 이가 가측이고 $\int_Y f \, d\varphi_* \mu = \sum_{i=1}^{k} a_i \varphi_* \mu(A_i) = \sum_{i=1}^{k} a_i \mu(\varphi^{-1}(A_i)) = \int_X \sum_{i=1}^{k} a_i \mathbf{1}_{\varphi^{-1}(A_i)} \, d\mu = \int_X f \circ \varphi \, d\mu$임이 분명하다. 이제 f를 음이 아닌 $\varphi_* \mathscr{A} / \mathscr{B}_1$−가측함수라 하면 정리 ? ? 로부터 음이 아닌 단순함수 $f_i : Y \to \mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i \uparrow f$이다. 여기서 각 $i \in \mathbb{N}$에 대해 $f_i \circ \varphi$가 가측이므로 정리 ? ? 의 ii에서 $f \circ \varphi$도 가측이고 MCT에서 $\int_Y f \, d\varphi_* \mu = \lim_{i \to \infty} \int_Y f_i \, d\varphi_* \mu = \lim_{i \to \infty} \int_Y f_i \circ \varphi \, d\mu = \int_Y f \circ \varphi \, d\mu$이다.

한편, $\varphi_* \mu$−적분가능한 $f : Y \to \overline{\mathbb{R}}$에 대해서도 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 $f \circ \varphi$가 μ−적분가능하며 $\int_Y f \, d\varphi_* \mu = \int_X f \circ \varphi \, d\mu$임을 쉽게 알 수 있고, $f \circ \varphi$가 μ−적분가능한 가측함수 $f : Y \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

마지막 성질은 앞서 살펴본 변수변환 공식과 묘하게 닮은 부분이 있기에 두 정리 모두 변수변환 공식이라 불린다. 물론, 이러한 유사함은 기분탓만은 아니다. 더 파보면 분명 무언가 더 있겠지만 우리의 갈 길이 먼 까닭에 pushfowarding에 대해서는 이쯤으로 정리하자. [29]

이번 절에서 pushfowarding이 필요한 이유는 바로 다음의 유명한 부등식 때문이다. 보통 확률론에서 쓰이는 이 부등식은 측도론에서도 유용하게 쓰이는 경우가 종종 있다. 나중에 확률론을 배우며 다른 부등식들도 만나볼 수 있을 것이다.

Lemma 1.184 측도공간 $(X, \mathscr{A}, \mu)$와 집합 Y 사이에서 정의된 함수 $\varphi : X \to Y$와 집합 $A \in \varphi_*\mathscr{A}$에 대해 $\mu(\varphi^{-1}(A)) = \int_X \mathbf{1}_A \circ \varphi \, d\mu$이다.

Proof 이는 $\int_X \mathbf{1}_A \circ \varphi \, d\mu = \int_X \mathbf{1}_A \, d\varphi_*\mu = \varphi_*\mu(A) = \mu(\varphi^{-1}(A))$로부터 자명하다. $\square$

Theorem 1.185 (Markov's inequality) 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f : X \to \overline{\mathbb{R}}$와 임의의 $M, p > 0$에 대해

$$\mu\{|f(x)| \geq M\} \leq \frac{1}{M^p} \int_X |f|^p \, d\mu$$

가 성립한다. [30]

Proof 부등식 $\mathbf{1}_{(-M,M)^c} \circ f = \mathbf{1}_{(-1,1)^c} \circ (f/M) \leq (|f|/M)^p \mathbf{1}_{(-1,1)^c} \circ (f/M) \leq (|f|/M)^p$와 위의 보조정리로부터 $\mu\{|f(x)| \geq M\} = \int_X \mathbf{1}_{(-M,M)^c} \circ f \, d\mu \leq \int_X (|f|/M)^p \, d\mu$가 성립한다. $\square$

이로써 준비는 끝났다. 이제 본격적으로 FTC의 i을 확장해보자. 먼저 '미분 비슷한 무언가' 를 엄밀히 도입하기 위한 framework이 필요한데, 곧 증명할 Hardy-Littlewood의 정리가 그 역할을 해 줄 것이다.

Definition 1.186 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 이의 **(Hardy-Littlewood) maximal function**을 $Mf : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$로 쓰고

$$Mf : x \mapsto \sup_{r>0} \frac{1}{\lambda_n(B(x, r))} \int_{B(x,r)} |f| \, d\lambda_n$$

로 정의한다.

Proposition 1.187 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 이의 maximal function Mf는 가측이다.

Proof 만약 $(Mf)(x) = 0$인 $x \in \mathbb{R}^n$가 존재하면 이는 곧 임의의 $i \in \mathbb{N}$에 대해 $\int_{B(x,i)} |f| \, d\lambda_n = 0$임을 뜻하여 정리 ? ? 의 i로부터 $\int_{\mathbb{R}^n} |f| \, d\lambda_n = \lim_{i\to\infty} \int_{B(x,i)} |f| \, d\lambda_n = 0$이고, 곧 정리 ? ? 의 ii로부터 $f = 0$ (ae.)이다. 그런데 이 경우에는 $Mf = 0$이 되어 이가 명백히 가측이므로 $Mf > 0$이라 가정하자.

이제 임의의 $x > 0$에 대해 $(Mf)^{-1}((x,\infty])$가 가측이면 $(Mf)^{-1}([-\infty,x]) = [(Mf)^{-1}((x,\infty])]^c$에서 $(Mf)^{-1}([-\infty,x])$도 가측이므로 $(Mf)^{-1}((x,\infty])$가 가측임을 보이는 것으로 충분하다. 이를 위해 임의의 $x > 0$를 고정하고 $(Mf)^{-1}((x,\infty])$가 열려있음을 보이자. 여기서 임의의 $y \in (Mf)^{-1}((x,\infty])$도 고정하면 적당한 $r > 0$이 존재하여

$$\frac{1}{\lambda_n(B(y,r))}\int_{B(y,r)} |f|\,d\lambda_n > x$$

이다. 한편, $r_i \downarrow r$인 임의의 $\{r_i\}$에 대해 정리 ? ? 의 ii로부터 $\lambda_n(B(y,r_i)) \downarrow \lambda_n(B(y,r))$이므로 $\lim_{s\downarrow r}\lambda_n(B(y,s)) = \lambda_n(B(y,r))$이고, 곧

$$\lim_{s\downarrow r}\frac{1}{\lambda_n(B(y,s))}\int_{B(y,r)} |f|\,d\lambda_n = \frac{1}{\lambda_n(B(y,r))}\int_{B(y,r)} |f|\,d\lambda_n > x$$

에서 적당한 $s > r$가 존재하여

$$\frac{1}{\lambda_n(B(y,s))}\int_{B(y,r)} |f|\,d\lambda_n > x$$

이다. 또한, 임의의 $z \in B(y,s-r)$에 대해 $w \in B(y,r)$이면 $||z-w|| \le ||z-y|| + ||y-w|| < s$이므로 $B(y,r) \subseteq B(z,s)$이고, 곧 Lebesgue 측도의 이동 불변성으로부터

$$\begin{aligned}
x &< \frac{1}{\lambda_n(B(y,s))}\int_{B(y,r)} |f|\,d\lambda_n \\
&\le \frac{1}{\lambda_n(B(y,s))}\int_{B(z,s)} |f|\,d\lambda_n \\
&= \frac{1}{\lambda_n(B(z,s))}\int_{B(z,s)} |f|\,d\lambda_n \\
&= (Mf)(z)
\end{aligned}$$

가 되어 $z \in (Mf)^{-1}((x,\infty])$, 즉 $B(y,s-r) \subseteq (Mf)^{-1}((x,\infty])$에서 증명이 끝난다. □

Lemma 1.188 (Vitali covering lemma) 집합 $\mathbb{R}^n$에 속하는 열린 ball의 모임 $\mathscr{C}$에 대해 $\bigcup\mathscr{C}$가 유계라면 $\mathscr{C}$에서 가산개의 서로소인 열린 ball $B(x_1,r_1), B(x_2,r_2),\cdots$를 적당히 택하여 $\bigcup\mathscr{C} \subseteq \bigcup_{i=1}^k B(x_i,5r_i)$이도록 할 수 있다. (여기서 k는 유한할 수도, ∞일 수도 있다.)

Proof 먼저 $d_1 = \sup\{r \in \mathbb{R}_0^+ : B(x,r) \in \mathscr{C}\} < \infty$에 대해 $r > d_1/2$인 열린 ball $B(x,r) \in \mathscr{C}$을 (가능하다면) 택하여 이를 $B(x_1,r_1)$이라 하자. 비슷하게 $d_2 = \sup\{r \in \mathbb{R}_0^+ : B(x,r) \in \mathscr{C}, B(x,r)\cap B(x_1,r_1) = \emptyset\} < \infty$에 대해 $r > d_2/2$인 열린 ball $B(x,r) \in \mathscr{C}$을 (가능하다면) 택하여 이를 $B(x_2,r_2)$라 하고, 이를 가능할 때까지 반복하여 가산개의 열린 ball $B(x_1,r_1), B(x_2,r_2),\cdots$를 얻는다. 이제 이들이 모두 서로소임은 그 구성으로부터 자명하므로 $\bigcup\mathscr{C} \subseteq \bigcup_i^k B(x_i,5r_i)$만 보이면 된다. (여기서 k는 유한할 수도, ∞일 수도 있는데, 이제

부터는 $k = \infty$라 생각하고 논의를 진행하도록 한다. k가 유한한 경우에도 비슷하게 하면 된다.)

만약 어떤 $B(x,r) \in \mathscr{C}$가 존재하여 모든 $i \in \mathbb{N}$에 대해 $B(x,r) \cap B(x_i, r_i) = \emptyset$이라면 위의 구성으로부터 각 $i \in \mathbb{N}$에 대해 $r \le d_i$이고, 곧 $r_i > d_i/2 \ge r/2$에서 $\lambda_n(\bigsqcup_i^\infty B(x_i, r_i)) = \sum_{i=1}^\infty \lambda_n(B(x_i, r_i)) \ge \sum_{i=1}^\infty \lambda_n(B(x_i, r/2)) = \infty$가 되어 $\bigcup \mathscr{C}$가 유계라는 가정에 모순되므로 임의의 $B(x,r) \in \mathscr{C}$에 대해 $B(x,r) \cap B(x_i, r_i) \neq \emptyset$인 $i \in \mathbb{N}$가 적어도 하나 존재한다. 이제 임의의 $B(x,r) \in \mathscr{C}$를 고정하고 $B(x,r) \cap B(x_i, r_i) \neq \emptyset$인 $i \in \mathbb{N}$ 중 가장 작은 것을 택하여 $i_0 \in \mathbb{N}$라 하면 $B(x,r) \cap \bigsqcup_{i=1}^{i_0-1} B(x_i, r_i) = \emptyset$이고 곧 위의 구성으로부터 $r \le d_{i_0} < 2r_{i_0}$이다. 또한 i_0의 정의로부터 $B(x,r) \cap B(x_{i_0}, r_{i_0}) \neq \emptyset$이므로 임의의 $y \in B(x,r)$에 대해 $||x_{i_0} - y|| \le ||x_{i_0} - x|| + ||x - y|| < 2r + r_{i_0} < 5r_{i_0}$가 되어 $y \in B(x_{i_0}, 5r_{i_0})$, 즉 $B(x,r) \subseteq B(x_{i_0}, 5r_{i_0})$에서 $\bigcup \mathscr{C} \subseteq \bigcup_i^\infty B(x_i, 5r_i)$임을 안다. □

Lemma 1.189 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$와 임의의 $M > 0$에 대해 $\lambda_n\{(Mf)(x) > M\} \le 5^n ||f||_1/M$이다.

Proof 임의의 $M > 0$과 임의의 $i \in \mathbb{N}$를 고정하면 임의의 $x \in (Mf)^{-1}((M, \infty]) \cap B(i) =: A_i$에 대해 적당한 $r_x > 0$가 존재하여

$$\frac{1}{\lambda_n(B(x, r_x))} \int_{B(x,r_x)} |f| \, d\lambda_n > M$$

에서 $\lambda_n(B(x,r_x)) < \int_{B(x,r_x)} |f|/M \, d\lambda_n \le ||f||_1/M$이므로 $x \in A_i$의 선택과는 무관하게 적당한 $R > 0$에 대해 $r_x \le R$이고, 곧 $\bigcup_{x \in A_i} B(x, r_x) \subseteq \bigcup_{x \in A_i} B(x, R) \subseteq B(R+i)$에서 $\bigcup_{x \in A_i} B(x, r_x)$는 유계이다. 그렇다면 Vitali covering lemma로부터 $\{B(x, r_x)\}_{x \in A_i}$에 속하는 서로소인 열린 ball의 열 $\{B(x_j, r_{x_j})\}$가 존재하여 $\bigcup_{x \in A_i} B(x, r_x) \subseteq \bigcup_{j=1}^\infty B(x_j, 5r_{x_j})$이다. 이로부터

$$\begin{aligned}
\lambda_n(A_i) &\le \lambda_n\left(\bigcup_{x \in A_i} B(x, r_x)\right) \\
&\le \lambda_n\left(\bigcup_{j=1}^\infty B(x_j, 5r_{x_j})\right) \\
&\le \sum_{j=1}^\infty \lambda_n(B(x_j, 5r_{x_j})) \\
&= 5^n \sum_{j=1}^\infty \lambda_n(B(x_j, r_{x_j})) \\
&\le \frac{5^n}{M} \sum_{j=1}^\infty \int_{B(x_j, r_{x_j})} |f| \, d\lambda_n \\
&= \frac{5^n}{M} \int_{\bigsqcup_{j=1}^\infty B(x_j, r_{x_j})} |f| \, d\lambda_n
\end{aligned}$$

$$\leq \frac{5^n}{M}||f||_1$$

이므로 정리 ? ? 의 i에서 $\lambda_n\{(Mf)(x) > M\} = \lim_{i\to\infty}\lambda_n(A_i) \leq 5^n||f||_1/M$이다. □

Theorem 1.190 (Hardy-Littlewood) 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$와 거의 대부분의 $x \in \mathbb{R}^n$에 대해 $r \downarrow 0$이면

$$\frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)}|f - f(x)|\,d\lambda_n \to 0$$

이다.

Proof 임의의 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 함수 $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$를

$$f^* : x \mapsto \limsup_{r\downarrow 0}\frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)}|f - f(x)|\,d\lambda_n$$

로 정의하고 이의 성질을 살펴보는 것으로 증명을 시작한다. 함수 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$를 임의의 적분가능한 함수라 하면

$$\begin{aligned}
&\limsup_{r\downarrow 0}\frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)}|(f+g) - (f+g)(x)|\,d\lambda_n \\
&\quad\leq \limsup_{r\downarrow 0}\frac{1}{\lambda_n(B(x,r))}\left[\int_{B(x,r)}|f - f(x)|\,d\lambda_n + \int_{B(x,r)}|g - g(x)|\,d\lambda_n\right] \\
&\quad\leq \limsup_{r\downarrow 0}\frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)}|f - f(x)|\,d\lambda_n + \limsup_{r\downarrow 0}\frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)}|g - g(x)|\,d\lambda_n
\end{aligned}$$

에서 $(f+g)^* \leq f^* + g^*$이다. 나아가, 만약 g가 $x_0 \in \mathbb{R}^n$에서 연속이라면 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 $||x - x_0|| < \delta$이면 $|g(x) - g(x_0)| < \varepsilon$이므로 $r < \delta$이면

$$\frac{1}{\lambda_n(B(x_0,r))}\int_{B(x_0,r)}|g - g(x_0)|\,d\lambda_n \leq \varepsilon$$

에서 $g^*(x_0) = 0$이다. 따라서 g가 연속이면 $g^* = (-g)^* = 0$이 되어 $(f-g)^* \leq f^* + (-g)^* = f^* \leq (f-g)^* + g^* = (f-g)^*$에서 $(f-g)^* = f^*$이다. 한편,

$$\begin{aligned}
\frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)}|f - f(x)|\,d\lambda_n &\leq \frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)}|f| + |f(x)|\,d\lambda_n \\
&= \frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)}|f|\,d\lambda_n + |f(x)| \\
&\leq (Mf)(x) + |f(x)|
\end{aligned}$$

에서 $f^* \leq Mf + |f|$도 성립한다.

이상의 성질들을 정리하면 다음과 같다.

i. 만약 $x_0 \in \mathbb{R}^n$에서 g가 연속이면 $g^*(x_0) = 0$이다.
ii. 만약 g가 연속이면 $(f-g)^* = f^*$이다.
iii. $f^* \leq Mf + |f|$.

이제 임의의 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$를 고정하고 $f^* = 0$ (ae.)임을 보이자. 이를 위해 임의의 $\varepsilon > 0$과 임의의 $M > 0$을 택하면 정리 ? ? 의 ii로부터 적당한 적분가능한 연속함수 $g : \mathbb{R}^n \to \mathbb{R}$가 존재하여 $||f-g||_1 < M\varepsilon/2(5^n+1)$이다. 한편, $(f-g)^* \leq M(f-g) + |f-g|$에서 $(f-g)^* > M$이면 $M(f-g) > M/2$이거나 $|f-g| > M/2$이므로 $[(f-g)^*]^{-1}((M,\infty]) \subseteq [M(f-g)]^{-1}((M/2,\infty]) \cup (|f-g|)^{-1}((M/2,\infty])$이고, 위의 보조정리와 Markov의 부등식으로부터

$$\begin{aligned}
\lambda_n&\left\{(M(f-g))(x) \geq \frac{M}{2} \text{ or } |(f-g)(x)| \geq \frac{M}{2}\right\} \\
&\leq \lambda_n\left\{(M(f-g))(x) \geq \frac{M}{2}\right\} + \lambda_n\left\{|f-g|(x) \geq \frac{M}{2}\right\} \\
&\leq \frac{2\cdot 5^n}{M}||f-g||_1 + \frac{2}{M}||f-g||_1 \\
&< \varepsilon
\end{aligned}$$

이다. 이는 곧 $[M(f-g)]^{-1}((M/2,\infty]) \cup (|f-g|)^{-1}((M/2,\infty])$가 영집합임을 함의하므로 $[(f-g)^*]^{-1}((M,\infty])$도 영집합이고, $(f-g)^* = f^*$에서 $(f^*)^{-1}((M,\infty])$도 영집합이다. 이제 M이 임의의 양수임을 상기한다면 $\{x \in \mathbb{R}^n : f^*(x) \neq 0\} = \bigcup_{i=1}^{\infty}(f^*)^{-1}((1/i,\infty])$에서 $f^* = 0$ (ae.)이고, 증명은 이로써 충분하다. □

Definition 1.191 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$와 $x \in \mathbb{R}^n$에 대해 $r \downarrow 0$일 때

$$\frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)} |f - f(x)|\, d\lambda_n \to 0$$

이면 이때의 $x \in \mathbb{R}^n$를 f의 **Lebesgue 점(− point)**라 한다.

Corollary 1.192 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 이가 $x \in \mathbb{R}^n$에서 연속이라 하면 x는 Lebesgue 점이다.

PROOF 이는 위의 정리의 증명에서 밝힌 f^*의 성질 i로부터 자명하다. □

Hardy-Littlewood의 정리는 적분가능한 함수에 대해 모든 Lebesgue 점을 모은 집합은 영집합임을 함의한다. 뭔가 우리가 원하는 결과가 얻어지는 듯한 느낌이 든다. 그러나 통상적인 미분과 Hardy-Littlewood의 정리를 자연스럽게 연결하기에는 아직 조금 부족하다.

Definition 1.193 집합족 $\mathscr{P}(\mathbb{R}^n)$에 속하는 가측인 집합열 $\{A_i\}$와 점 $x \in \mathbb{R}^n$에 대해 만약 적당한 상수 $C > 0$와 0으로 수렴하는 양의 수열 $\{r_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $A_i \subseteq B(x, r_i)$

이고 $\lambda_n(B(x, r_i)) \le C\lambda_n(A_i)$이면 이때의 집합열 $\{A_i\}$가 x로 **regular하게 수렴한다(converge regularly)**고 한다.

Theorem 1.194 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$와 이의 Lebesgue 점 $x \in \mathbb{R}^n$, 그리고 x로 regular하게 수렴하는 $\mathscr{P}(\mathbb{R}^n)$에 속하는 가측인 집합열 $\{A_i\}$에 대해

$$\frac{1}{\lambda_n(A_i)} \int_{A_i} |f - f(x)|\, d\lambda_n \to 0$$

이다.

Proof 가정으로부터 적당한 $C > 0$와 0으로 수렴하는 양의 수열 $\{r_i\}$가 존재하여 $A_i \subseteq B(x, r_i)$이고 $\lambda_n(B(x, r_i)) \le C\lambda_n(A_i)$이므로 임의의 $i \in \mathbb{N}$에 대해

$$\frac{1}{\lambda_n(A_i)} \int_{A_i} |f - f(x)|\, d\lambda_n \le \frac{C}{\lambda_n(B(x, r_i))} \int_{B(x, r_i)} |f - f(x)|\, d\lambda_n$$

이다. 그런데 Hardy-Littlewood의 정리로부터 위 식의 우변이 $i \to \infty$이면 0으로 수렴하므로 증명이 끝난다. □

Corollary 1.195 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$와 이의 Lebesgue 점 $x \in \mathbb{R}^n$, 그리고 x로 regular하게 수렴하는 $\mathscr{P}(\mathbb{R}^n)$에 속하는 가측인 집합열 $\{A_i\}$에 대해

$$\frac{1}{\lambda_n(A_i)} \int_{A_i} f\, d\lambda_n \to f(x)$$

이다.

Proof 임의의 $i \in \mathbb{N}$에 대해 $|\int_{A_i} f\, d\lambda_n - f(x)\lambda_n(A_i)| = |\int_{A_i} f - f(x)\, d\lambda_n| \le \int_{A_i} |f - f(x)|\, d\lambda_n$이므로 따름정리는 자명하다. □

마침내 다음 정리를 얻는다. 이는 일반적인 n차원에서 적분가능한 함수에 대한 정리이므로 우리가 원했던 결과보다 더 강력한 결과이다.

Definition 1.196 열린집합 $U \subseteq \mathbb{R}^n$에서 정의된 함수 $f : U \to \mathbb{R}$와 한 점 $x \in U$에 대해 극한 $\lim_{h\downarrow 0} \Delta_{[x-h\mathbf{1}, x+h\mathbf{1}]} f/(2h)^n$이 수렴하면 f는 x에서 **symmetric differentiable**하다고 하고, 이때의 극한값을 $f^{(s)}(x)$로 쓴다. 나아가, 만약 f가 모든 $x \in U$에서 symmetric differentiable하면, 이때의 f를 **symmetric differentiable**하다고 한다.

Corollary 1.197 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 함수 $F : \mathbb{R}^n \to \mathbb{R}$를

$$F : x \mapsto \int_{(-\infty, x]} f\, d\lambda_n$$

로 두면 $F^{(s)} = f$ (ae.)이다. 나아가, $x_0 \in \mathbb{R}^n$의 근방에서 $\partial^{\mathbf{1}}F$이 존재하고 이가 x_0에서 연속이면 $F^{(s)}(x_0) = \partial^{\mathbf{1}}F(x_0)$이다. (단, 여기서 F가 $x \in \mathbb{R}$에서 symmetric differentiable하지 않은 경우 $F^s(x) = 0$으로 생각한다.)

Proof 간결한 논의를 위해 $n = 2$인 경우만 생각하자. ($n = 1$인 경우를 포함한 일반적인 경우에도 이와 비슷하게 보일 수 있다.) 이제 임의의 Lebesgue 점 $x \in \mathbb{R}^2$와 0으로 수렴하는 임의의 수열 $\{h_i\}$를 택하여 $B_x(h_i) = [x_1 - h_i, x_1 + h_i] \times [x_2 - h_i, x_2 + h_i]$라 하면 $\{B_x(h_i)\}$가 x로 regular하게 수렴함을 쉽게 보일 수 있다. 그렇다면 따름정리 ? ? 로부터

$$\begin{aligned}
\frac{\Delta_{B_x(h_i)}F}{(2h_i)^2} &= \frac{1}{(2h_i)^2}\Bigg[\left(\int_{(-\infty,x_1+h_i]\times(-\infty,x_2+h_i]} f\,d\lambda_2 - \int_{(-\infty,x_1-h_i]\times(-\infty,x_2+h_i]} f\,d\lambda_2\right) \\
&\qquad - \left(\int_{(-\infty,x_1+h_i]\times(-\infty,x_2-h_i]} f\,d\lambda_2 - \int_{(-\infty,x_1-h_i]\times(-\infty,x_2-h_i]} f\,d\lambda_2\right)\Bigg] \\
&= \frac{1}{(2h_i)^2}\int_{[x_1-h_i,x_1+h_i]\times[x_2-h_i,x_2+h_i]} f\,d\lambda_2 \\
&= \frac{1}{\lambda_2(B_x(h_i))}\int_{B_x(h_i)} f\,d\lambda_2 \\
&\to f(x)
\end{aligned}$$

이므로 $F^{(s)}(x) = f(x)$이고, Hardy-Littlewood의 정리로부터 $\mathbb{R}^n$의 거의 대부분의 점들이 이러한 Lebesgue 점이므로 $F^{(s)} = f$ (ae.)임을 안다.

한편, $x_0 \in \mathbb{R}^n$의 근방에서 $\partial^{\mathbf{1}}F$가 존재하고 이가 x_0에서 연속이라 하자. 또한, 편의를 위해 편미분이 $\partial x_1 \partial x_2$의 순서로 주어진 경우만 생각하자. (다른 경우에 대해서도 이와 비슷하게 보일 수 있다.) 그렇다면 따름정리 ? ? 에서 x_0는 Lebesgue 점이므로 앞선 결과로부터 F가 x_0에서 symmetric differentiable하고, MVT와 L'Hôspital의 법칙으로부터

$$\begin{aligned}
&\lim_{h\downarrow 0}\frac{\Delta_{[x_0^1-h,x_0^1+h]\times[x_0^2-h,x_0^2+h]}F}{(2h)^2} \\
&= \lim_{h\downarrow 0}\frac{1}{2h}\left[\frac{F(x_0^1+h, x_0^2+h) - F(x_0^1+h, x_0^2-h)}{2h} - \frac{F(x_0^1-h, x_0^2+h) - F(x_0^1-h, x_0^2-h)}{2h}\right] \\
&= \lim_{h\downarrow 0}\frac{1}{2h}\left[\frac{\partial F(x_0^1+h, y_h)}{\partial x_2} - \frac{\partial F(x_0^1-h, z_h)}{\partial x_2}\right] \\
&= \lim_{h\downarrow 0}\frac{1}{2}\left[\frac{\partial^2 F(x_0^1+h, y_h)}{\partial x_1 \partial x_2} - \frac{\partial^2 F(x_0^1-h, z_h)}{\partial x_1 \partial x_2}\right] \\
&= \frac{\partial^2 F}{\partial x_1 \partial x_2}(x_0)
\end{aligned}$$

가 되어 증명이 끝난다. 여기서 y_h, z_h는 모두 $x_0^2 - h$와 $x_0^2 + h$를 잇는 선분 위에 존재하는 점들로 MVT에서 그 존재가 보장된다. □

비록 위의 정리에서는 F가 symmetric differentiable하지 않은 $x \in \mathbb{R}$에서의 $F^{(s)}(x)$의 값을 0이라는 dummy value로 두었지만, Hardy-Littlewood의 정리로부터 이런 점의 집합이 영집합을 이루므로 정리의 성립여부는 이때의 dummy value의 선택과는 무관하다. 이러한 사실은 뒤따르는 FTC의 일반화에 대해서도 마찬가지이다. 이제 위의 정리에서 $n = 1$인 경우를 기존의 FTC의 i에 대입하면 다음을 얻는다.

Theorem (Fundamental theorem of calculus) i. 적분가능한 함수 $f : \mathbb{R} \to \mathbb{R}$에 대해 함수 $F : \mathbb{R} \to \mathbb{R}$를 $F : x \mapsto \int_{-\infty}^{x} f\, d\lambda_1$로 두면 $F^{(s)} = f$ (ae.)이다.

여기서의 극한은 꽤나 익숙하다. 고등학교 문제집에서 한 번쯤은 보았을 이 극한은 정말 '미분 비슷한 무언가' 이지만 항상 미분과 그 값이 같은 것도 아니고, 이 극한의 존재성이 미분가능성을 함의하는 것도 아니다. 하지만, 이는 f의 연속성을 포기한 대가이므로 이 정도에 만족하도록 하자.

이제 기존의 FTC의 ii를 살펴보는 것으로 이번 절의 후반부를 시작한다. 앞서 i을 일반화시킨 것과 비슷하게, F가 $\mathscr{C}^1$급이라는 조건을 F가 거의 어디서나 symmetric differentiable 하다는 조건으로 바꾸어보자. 아쉽게도 아직은 한참 부족하다. 생각해보면, 이는 굉장히 순진한 시도였다. 만약 F가 정말 $F(x) = \int_{-\infty}^{x} f\, d\lambda_1$을 만족한다면 이는 정리 ? ? 의 iv와 $F(x) = \int_{-\infty}^{x} f_+\, d\lambda_1 - \int_{-\infty}^{x} f_-\, d\lambda_1 = (\lambda_1)_{f_+}((-\infty, x]) - (\lambda_1)_{f_-}((-\infty, x])$에서 반드시 오른쪽 연속인데, 우리는 F가 거의 어디서나 symmetric differentiable해야 할 것을 요구했을 뿐, 이의 연속성에 대해서는 생각하지 않았기 때문이다.

그렇다면 조금 더 머리를 써서 $F^{(s)} = f$ (ae.)일 뿐만 아니라 F가 연속일 것까지 요구해보자. 그러나 아직도 부족하다. 즉, $F^{(s)} = f$ (ae.)이고 F가 연속이지만 $F(x) \neq \int_{-\infty}^{x} f\, d\lambda_1$인 반례가 존재한다.[31]어쩌면 f의 연속성을 포기한 대가가 너무나도 커서 이를 대체할 좋은 조건이 없는 것은 아닐까? 다행히, 우리의 질문이 이렇게 재미없게 결론나지는 않는다. 지금 우리에게 필요한 적당한 조건은 바로 F의 절대연속성이다.

Definition 1.198 함수 $f : \mathbb{R}^n \to \mathbb{R}$를 생각하자. 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 임의의 유계이고 서로소인 $B_1, \cdots, B_k \in \mathscr{S}_n$에 대해 $\sum_{i=1}^{k} \mu_n(B_i) < \delta$이면 $\sum_{i=1}^{k} |\Delta_{B_i} f| < \varepsilon$을 만족하면 이때의 f를 **절대연속(absolutely continuous)**이라 한다.

이전 절에서 Radon-Nikodým 도함수를 도입하며 측도의 절대연속을 정의했는데, 그 용어가 그대로 쓰이는 것은 단순한 우연의 일치가 아니다. 측도의 절대연속성과 함수의 절대연속성은 서로 긴밀히 연결되어 있다.

Theorem 1.199 (Borel-Cantelli) 측도공간 $(\mathbb{R}^n, \mathscr{B}_n, \mu)$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $\sum_{i=1}^{\infty} \mu(A_i) < \infty$이면 $\mu(\limsup_{i\to\infty} A_i) = 0$이다.

Proof 임의의 $\varepsilon > 0$을 택하면 충분히 큰 $i_0 \in \mathbb{N}$에 대해 $\sum_{i=i_0}^{\infty} \mu(A_i) < \varepsilon$인데, 정의로부터 $\limsup_{i\to\infty} A_i \subseteq \bigcup_{i=i_0}^{\infty} A_i$이므로 $\mu(\limsup_{i\to\infty} A_i) \leq \mu(\bigcup_{i=i_0}^{\infty} A_i) \leq \sum_{i=i_0}^{\infty} \mu(A_i) < \varepsilon$에서 $\mu(\limsup_{i\to\infty} A_i) = 0$임을 안다. □

Theorem 1.200 $\sigma-$유한 측도공간 $(\mathbb{R}^n, \mathscr{B}_n, \mu)$에 대해 함수 $F_\mu : \mathbb{R}^n \to \mathbb{R}_0^+$를 $F_\mu : x \mapsto \mu(\prod_{i=1}^{n}(-\infty, x_i])$로 두자. 그렇다면 F_μ가 절대연속일 필요충분조건은 $\mu \ll \mu_n$인 것이다.

Proof 충분조건임을 보이기 위해서는 $\mu_n(N) = 0$인 임의의 $N \in \mathscr{B}_n$을 고정하고 $\mu(N) = 0$임을 보이면 충분하다. 이제 임의의 $\varepsilon > 0$을 택하면 정리 ? ? 의 ii와 가정으로부터 임의의 $\delta > 0$가 존재하여 임의의 유계이고 서로소인 $B_1, \cdots, B_k \in \mathscr{S}_n$에 대해 $\sum_{i=1}^{k} \mu_n(B_i) < \delta$이면 $\sum_{i=1}^{k} \mu(B_i) = \sum_{i=1}^{k} \Delta_{B_i} F_\mu < \varepsilon$이다. 한편, 정리 ? ? 의 i과 ? ? 로부터 $\mathscr{S}_n$에 속하는 적당한 서로소인 집합열 $\{B_i\}$가 존재하여 $N \subseteq \bigsqcup_{i=1}^{\infty} B_i$이고 $\sum_{i=1}^{\infty} \mu_n(B_i) = \mu_n(\bigsqcup_{i=1}^{\infty} B_i) - \mu_n(N) = \mu_n(\bigsqcup_{i=1}^{\infty} B_i \setminus N) < \delta$이다. 이상으로부터 임의의 $k \in \mathbb{N}$에 대해 $\sum_{i=1}^{k} \mu_n(B_i) < \delta$이므로 $\sum_{i=1}^{k} \mu(B_i) < \varepsilon$이고, 곧 $\sum_{i=1}^{\infty} \mu(B_i) \leq \varepsilon$이 되어 $\mu(N) \leq \mu(\bigsqcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i) \leq \varepsilon$에서 $\mu(N) = 0$임을 안다.

다음으로, 필요조건임을 보이기 위해 F_μ가 절대연속이 아니면 $\mu \ll \mu_n$이 아님을 보이자. 정리 ? ? 의 ii와 가정으로부터 적당한 $\varepsilon > 0$이 존재하여 임의의 $i \in \mathbb{N}$에 대해 유계이고 서로소인 $B_{i1}, \cdots, B_{ik_i} \in \mathscr{S}_n$가 존재하여 $\mu_n(\bigsqcup_{j=1}^{k_i} B_{ij}) = \sum_{j=1}^{k_i} \mu_n(B_{ij}) < 1/i^2$이고 $\mu(\bigsqcup_{j=1}^{k_i} B_{ij}) = \sum_{j=1}^{k_i} \mu(B_{ij}) = \sum_{j=1}^{k_i} \Delta_{B_{ij}} F_\mu \geq \varepsilon$이다. 이로부터 집합열 $\{B_i\}$를 $B_i := \bigsqcup_{j=1}^{k_i} B_{ij}$로 정의하면 $\sum_{i=1}^{\infty} \mu_n(B_i) = \sum_{i=1}^{\infty} \mu_n(\bigsqcup_{j=1}^{k_i} B_{ij}) < \sum_{i=1}^{\infty} 1/i^2 < \infty$이므로 정리 ? ? 와 Borel-Cantelli의 정리로부터 $\mu_n(\limsup_{i\to\infty} B_i) = 0$이지만 $\mu(\limsup_{i\to\infty} B_i) \geq \limsup_{i\to\infty} \mu(B_i) = \limsup_{i\to\infty} \mu(\bigsqcup_{j=1}^{k_i} B_{ij}) \geq \varepsilon$이 되어 $\mu \ll \mu_n$이 아님을 안다. □

위의 정리는 놀라운 발견이라기보다 이러한 관계가 성립하도록 함수의 절대연속의 정의가 디자인되었다고 보는 것이 맞을 것 같다. (균등연속이나 Lipschitz 연속과 같은 다른 함수의 연속성의 정의에 비해 절대연속의 정의는 다소 복잡하고 기하학적 직관도 부족하다.) 아무튼 이로부터 우리가 원하던 결론을 얻을 수 있다.

Theorem 1.201 함수 $F : \mathbb{R}^n \to \mathbb{R}$가 다음을 만족한다고 하자.

i. 함수 F는 거의 어디서나 symmetric differentiable하다.
ii. 함수 F는 절대연속이고 오른쪽 연속이며 각 변수에 대해 증가한다.
iii. 유계인 semi-open box $B \subseteq \mathbb{R}^n$에 대해 $\Delta_B F \geq 0$이다.
iv. 각 $i \leq n$에 대해 $\lim_{x_i \to -\infty} F(x) = 0$이다.
v. $\lim_{x_1, \cdots, x_n \to \infty} F(x) < \infty$.

그렇다면 $f := F^{(s)}$에 대해 $F = \int_{(-\infty, x]} f \, d\lambda_n$이다. (단, 여기서 F가 $x \in \mathbb{R}^n$에서 symmetric differentiable하지 않은 경우 $f(x) = 0$으로 생각한다.)

Proof 먼저 정리 ? ? 로부터 $\mathscr{B}_n$ 위의 적당한 측도 μ가 존재하여 $F(x) = \mu(\prod_{i=1}^n(-\infty, x_i])$이고, 조건 v와 정리 ? ? 의 vii로부터 이때의 μ는 유한하다. 그렇다면 조건 ii와 정리 ? ? 로부터 $\mu \ll \mu_n$이 되어 Radon-Nikodým 도함수 $g := d\mu/d\mu_n$가 존재하고 정리 ? ? 로부터 $F(x) = \mu(\prod_{i=1}^n(-\infty, x_i]) = \int_{\prod_{i=1}^n(-\infty,x_i]} g\,d\mu_n = \int_{\prod_{i=1}^n(-\infty,x_i]} g\,d\lambda_n$이다. 그런데 우리가 앞서 일반화한 FTC i로부터 $h \downarrow 0$이면 $\lim_{h\downarrow 0} \Delta_{B_x(h)}F/(2h)^n \to g(x)$ (ae.)이므로 $f = g$ (ae.)가 되어 정리 ? ? 의 ii로부터 $F = \int_{\prod_{i=1}^n(-\infty,x_i]} f\,d\lambda_n$이다. □

위의 정리에서 $n = 1$인 경우를 기존의 FTC에 대입하면 다음을 얻는다.

Proposition 1.202 절대연속함수 $f : \mathbb{R} \to \mathbb{R}$는 균등연속이다.

Proof 거의 자명하다. 함수 f가 절대연속이므로 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 임의의 유계인 반열린구간 $I = (x, y) \subseteq \mathbb{R}$에 대해 $|x-y| = \mu_1(I) < \delta$이면 $|f(x) - f(y)| = |\Delta_I f| < \varepsilon$이고, 이로부터 f가 균등연속임이 분명하다. □

Theorem (Fundamental theorem of calculus) ii. 함수 $F : \mathbb{R} \to \mathbb{R}$가 다음을 만족한다고 하자.

a. 함수 F는 거의 어디서나 symmetric differentiable하다.
b. 함수 F는 절대연속인 증가함수이다.
c. $\lim_{x\to-\infty} F(x) = 0$.
d. $\lim_{x\to\infty} F(x) < \infty$.

그렇다면 $f := F^{(s)}$에 대해 $F = \int_{-\infty}^x f\,d\lambda_1$이다. (단, 여기서 F가 $x \in \mathbb{R}$에서 symmetric differentiable하지 않은 경우 $f(x) = 0$으로 생각한다.)

정리 ? ? 에서의 조건들이 $n = 1$인 경우에는 많이 간단해진 것을 볼 수 있다. 그렇지만 뭔가 만족스럽지 못한 것도 사실이다. 일반적인 $n \in \mathbb{N}$에 대해 절대연속이라는 개념이 그 이름이 무색하게 오른쪽 연속성조차 함의하지 못하고, FTC의 ii의 확장이라고는 했지만 F에는 여전히 많은 조건들이 붙어있다. 특히 F가 증가함수인 경우로 한정된 것은 누가 보더라도 불필요하게 강하게 주어진 감이 있다. 하지만 나중에 필요한 정리나 개념은 모두 소개했고, 보다 나은 결과를 얻기 위해서는 지금까지의 논의와는 다른 별개의 논의가 필요한데, 이는 하나의 절에 담아내기에는 벅찬 내용이다. 이에 우리는 아쉬움을 뒤로 하고 여기서 멈추어 다음의 FTC를 얻은 것에 만족하기로 한다. [32]

Theorem 1.203 (Fundamental theorem of calculus) 다음이 성립한다.

i. 적분가능한 함수 $f : \mathbb{R} \to \mathbb{R}$에 대해 함수 $F : \mathbb{R} \to \mathbb{R}$를 $F : x \mapsto \int_{-\infty}^x f\,d\lambda_1$로 두면 $F^{(s)} = f$ (ae.)이다.
ii. 함수 $F : \mathbb{R} \to \mathbb{R}$가 다음을 만족한다고 하자.

a. 함수 F는 거의 어디서나 symmetric differentiable하다.
b. 함수 F는 절대연속인 증가함수이다.
c. $\lim_{x\to-\infty} F(x) = 0$.
d. $\lim_{x\to\infty} F(x) < \infty$.

그렇다면 $f := F^{(s)}$에 대해 $F = \int_{-\infty}^{x} f\, d\lambda_1$이다. (단, 여기서 F가 $x \in \mathbb{R}$에서 symmetric differentiable하지 않은 경우 $f(x) = 0$으로 생각한다.)

1.14 Convolution

이번 절에서는 적분으로 정의되는 유명한 연산인 합성곱에 대한 엄밀한 이론 전개를 해 보기로 한다. 이미 대부분의 독자들이 알고 있겠지만 합성곱은 두 함수를 '곱하는' 특유의 방식으로 신호처리에서 두 신호를 합성하는 방법의 일종으로 널리 쓰인다. 우리도 이와 비슷하게 이후 확률론을 배우면서 두 확률변수를 더하는 과정에서 합성곱을 사용하게 될 것이다. 함수의 합성곱과 관련하여 필요한 이론의 전개가 이루어진 후에는 이와 유사하게 측도의 합성곱을 정의하고, 함수의 합성곱과 측도의 합성곱이 어떻게 연결되어 있는지를 살펴보도록 하자.

Proposition 1.204 가측함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$와 임의의 $x \in \mathbb{R}^n$에 대해 함수 $t \mapsto f(x-t)$는 가측이다.

Proof 함수 $f_x : \mathbb{R}^n \to \mathbb{R}$와 미분동형사상 $\Phi_x : \mathbb{R}^n \to \mathbb{R}^n$를 각각 $f_x : t \mapsto f(x-t)$, $\Phi_x : t \mapsto x-t$로 두면 변수변환 공식으로부터 $f_x = f \circ \Phi_x$가 가측이므로 이 명제는 자명하다. □

Definition 1.205 가측함수 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 집합

$$E(f,g) = \{x \in \mathbb{R}^n : \text{함수 } t \mapsto f(x-t)g(t)\text{가 적분가능하지 않다.}\}$$

를 생각하자. 이제 f와 g의 **합성곱(convolution)**을 $f * g : \mathbb{R}^n \to \mathbb{R}$로 쓰고

$$f * g : x \mapsto \begin{cases} \int_{\mathbb{R}^n} f(x-t)g(t)\, d\lambda_n(t) & x \notin E(f,g)\text{인 경우} \\ 0 & \text{ow.} \end{cases}$$

로 정의한다.

명제 ? ? 에서 함수 $t \mapsto f(x-t)g(t)$가 가측이므로 합성곱은 well-define되며 이로써 $f * g$를 아주 일반적으로 정의할 수 있다. 이제 합성곱의 기본적인 성질들을 알아보자.

Theorem 1.206 가측함수 $f, \widetilde{f}, g, \widetilde{g} : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $f = \widetilde{f}$ (ae.)이고 $g = \widetilde{g}$ (ae.)이면 $E(f, g) = E(\widetilde{f}, \widetilde{g})$이고 $f * g = \widetilde{f} * \widetilde{g}$이다.

Proof 임의의 $x \in \mathbb{R}^n$에 대하여 함수 $f_x, \widetilde{f}_x : \mathbb{R}^n \to \mathbb{R}$와 미분동형사상 $\Phi_x : \mathbb{R}^n \to \mathbb{R}^n$를 각각 $f_x : t \mapsto f(x-t), \widetilde{f}_x : t \mapsto \widetilde{f}(x-t), \Phi_x : t \mapsto x - t$로 두면 명제 ? ? 로부터 f_x와 $\widetilde{f}_x$가 가측이다. 이제 집합 $M = \{x \in \mathbb{R}^n : f(x) \neq \widetilde{f}(x)\}, N = \{x \in \mathbb{R}^n : g(x) \neq \widetilde{g}(x)\}$이 영집합이므로 임의의 $x \in \mathbb{R}^n$를 택하여 집합 $A_x = \{t \in \mathbb{R}^n : (f_x g)(t) \neq (\widetilde{f}_x \widetilde{g})(t)\}$를 생각하면 $A_x \subseteq \{t \in \mathbb{R}^n : x - t \in M\} \cup N = \Phi_x^{-1}(M) \cup N$이다. 그런데 Φ_x^{-1}가 미분동형사상이고 곧 Lipschitz 연속이므로 보조정리 ? ? 로부터 $\Phi_x^{-1}(M)$이 영집합이 되어 A_x도 영집합이고, 이로부터 $f_x g = \widetilde{f}_x \widetilde{g}$ (ae.)임을 안다. 이는 임의의 $x \in \mathbb{R}^n$에 대해 $f_x g$와 $\widetilde{f}_x \widetilde{g}$의 적분가능성이 일치함을 의미하므로 $E(f, g) = E(\widetilde{f}, \widetilde{g})$가 성립하고, 이어서 $f * g = \widetilde{f} * \widetilde{g}$도 성립한다. □

Theorem 1.207 가측함수 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 집합 $E(f, g)$는 Borel이고 함수 $f * g$도 Borel이다.

Proof 정리 ? ? , ? ? 의 ii로부터 적당한 Borel 함수 $\widetilde{f}, \widetilde{g} : \mathbb{R}^n \to \overline{\mathbb{R}}$가 존재하여 $f = \widetilde{f}$ (ae.)이고 $g = \widetilde{g}$ (ae.)이다. 이제 함수 $F : \mathbb{R}^{2n} \to \overline{\mathbb{R}}$를 $F : (x, t) \mapsto \widetilde{f}(x-t)\widetilde{g}(t)$로 두면 명제 ? ? 의 증명에서와 비슷하게 하여 이가 Borel임을 쉽게 보일 수 있다. 이로부터 함수 $G : \mathbb{R}^n \to \overline{\mathbb{R}}$를 $G : x \mapsto \int_{\mathbb{R}^n} |F(x, t)| \, d\mu(t)$로 정의하면 Fubini의 정리로부터 이는 Borel이고, 정리 ? ? 로부터 $E(f, g) = E(\widetilde{f}, \widetilde{g}) = G^{-1}(\infty) =: E$도 Borel이다. 증명을 끝내기 위해 함수 $\widetilde{F}^{\pm} : \mathbb{R}^{2n} \to \overline{\mathbb{R}}$를

$$\widetilde{F}^{\pm} : (x, t) \mapsto \begin{cases} F_{\pm}(x, t) & x \notin E \\ 0 & \text{ow.} \end{cases}$$

로 두면 $\widetilde{F}^{\pm}$는 Borel이고, 다시 정리 ? ? 로부터 임의의 $x \in \mathbb{R}^n$에 대해 $(f * g)(x) = (\widetilde{f} * \widetilde{g})(x) = \int_{\mathbb{R}^n} \widetilde{F}^{+}(x, t) \, d\mu(t) - \int_{\mathbb{R}^n} \widetilde{F}^{-}(x, t) \, d\mu(t)$이므로 Fubini의 정리로부터 $f * g$가 Borel임을 안다. □

이건 사실 조금 신기한 결과이다. 위의 정리에 따르면 비록 f, g 각각은 가측인 함수에 그쳐 다루기 힘든 '이상한' 함수일 수 있지만, $f * g$는 Borel이 되어 그 합성곱은 반드시 '예쁜' 함수로 얻어진다. 이어서 살펴볼 것은 합성곱의 교환법칙과 분배법칙이다.

Theorem 1.208 가측함수 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$와 $c \neq 0$에 대해 다음이 성립한다.

i. $E(f, g) = E(cf, g) = E(f, cg)$이고 $c(f * g) = (cf) * g = f * (cg)$이다.
ii. $E(f, g) = E(g, f)$이고 $f * g = g * f$이다.

Proof i. 우선 $E(f,g) = E(cf,g) = E(f,cg) =: E$임은 정의로부터 자명하고, 임의의 $x \in \mathbb{R}^n \setminus E$에 대해 $[c(f*g)](x) = c\int_{\mathbb{R}^n} f(x-t)g(t)\,d\lambda_n(t) = \int_{\mathbb{R}^n} cf(x-t)g(t)\,d\lambda_n(t) = [(cf)*g](x)$이므로 $c(f*g) = (cf)*g$이다. 이제 이와 비슷하게 $c(f*g) = f*(cg)$임도 보일 수 있다.

ii. 임의의 $x \in \mathbb{R}^n$에 대하여 함수 $f_x, g_x : \mathbb{R}^n \to \overline{\mathbb{R}}$와 미분동형사상 $\Phi_x : \mathbb{R}^n \to \mathbb{R}^n$를 각각 $f_x : t \mapsto f(x-t), g_x : t \mapsto g(x-t), \Phi_x : t \mapsto x-t$로 정의하면 명제 ? ? 로부터 f_x와 g_x가 가측이다. 이제 임의의 $x \in \mathbb{R}^n$에 대해 $g_x f = (f_x g)\circ\Phi_x$이므로 변수변환 공식으로부터 $f_x g$와 $g_x f$의 적분가능성이 일치하여 $E(f,g) = E(g,f) =: E$이다. 나아가 임의의 $x \in \mathbb{R}^n \setminus E$에 대해 $(f*g)(x) = \int_{\mathbb{R}^n} f_x g\,d\lambda_n = \int_{\mathbb{R}^n}[(f_x g)\circ\Phi_x]|\det \mathbf{D}\Phi_x|\,d\lambda_n = \int_{\mathbb{R}^n} g_x f\,d\lambda_n = (g*f)(x)$이므로 $f*g = g*f$가 성립한다. □

Theorem 1.209 가측함수 $f,g,h : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $E(f,g+h) \subseteq E(f,g)\cup E(f,h)$이고 $\mathbb{R}^n \setminus [E(f,g)\cup E(f,h)]$에서 $f*(g+h) = f*g + f*h$이다. 단, 이때 $g+h$를 계산하는 과정에서 $\infty - \infty$와 같은 부정형이 등장하여 그 결과가 well-define되지 않는 경우 이를 $c \in \mathbb{R}$와 같은 dummy-value로 두는 관례를 전제하자.

Proof 임의의 $x \in \mathbb{R}^n$에 대하여 함수 $f_x : \mathbb{R}^n \to \mathbb{R}$를 $f_x : t \mapsto f(x-t)$로 정의하면 명제 ? ? 로부터 f_x가 가측이다. 이제 $|f_x(g+h)| \le |f_x g| + |f_x h|$이므로 정리 ? ? 의 i로부터 $E(f,g+h) \subseteq E(f,g)\cup E(f,h) =: E$임을 안다. 한편, 임의의 $x \in \mathbb{R}^n \setminus E$에 대해

$$\begin{aligned}
[f*(g+h)](x) &= \int_{\mathbb{R}^n} f(x-t)(g+h)(t)\,d\lambda_n(t) \\
&= \int_{\mathbb{R}^n} f(x-t)g(t)\,d\lambda_n(t) + \int_{\mathbb{R}^n} f(x-t)h(t)\,d\lambda_n(t) \\
&= (f*g)(x) + (f*h)(x)
\end{aligned}$$

이므로 $\mathbb{R}^n \setminus E$에서 $f*(g+h) = f*g + f*h$가 성립한다. □

다음으로, 합성곱의 결합법칙을 보자. 방금 살펴본 합성곱의 교환법칙과 분배법칙과는 달리, 결합법칙은 조금 복잡하다.

Theorem 1.210 가측함수 $f,g : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $|f*g| \le |f|*|g|$이다.

Proof 임의의 $x \in \mathbb{R}^n$에 대해 만약 $x \in E(f,g) = E(|f|,|g|)$이면 $(f*g)(x) = (|f|*|g|)(x) = 0$이 되어 정리가 자명하므로 $x \notin E(f,g)$라 하자. 그렇다면 $|(f*g)(x)| = |\int_{\mathbb{R}^n} f(x-t)g(t)\,d\lambda_n(t)| \le \int_{\mathbb{R}^n} |f(x-t)g(t)|\,d\lambda_n(t) = (|f|*|g|)(x)$가 되어 증명이 끝난다. □

Lemma 1.211 Borel 함수 $f,g,h : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $E(f,g), E(g,h)$가 영집합이면 $E(|f|,|g|*|h|) = E(|f|*|g|,|h|)$이다.

PROOF 임의의 $x \in \mathbb{R}^n$를 택하여 함수 $F_x : \mathbb{R}^{2n} \to \overline{\mathbb{R}}$를 각각 $F_x : (s,t) \mapsto f(x-t)g(t-s)h(s)$로 두면 명제 ? ? 의 증명에서와 비슷하게 하여 이가 Borel임을 쉽게 보일 수 있다. 그렇다면 $\int_{\mathbb{R}^n} |F_x(s,t)|\, d\lambda_n(s) = |f(x-t)| \int_{\mathbb{R}^n} |g(t-s)h(s)|\, d\lambda_n(s) = |f(x-t)|(|g|*|h|)(t)$가 임의의 $t \in \mathbb{R}^n \setminus E(|g|,|h|)$에 대해 성립하는데, 가정으로부터 $E := E(|g|,|h|) = E(g,h)$가 영집합이므로 Fubini의 정리에서

$$\begin{aligned}
||F_x||_1 &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |F_x(s,t)|\, d\lambda_n(s) \right] d\lambda_n(t) \\
&= \int_{\mathbb{R}^n} |f(x-t)| \left[\int_{\mathbb{R}^n} |g(t-s)h(s)|\, d\lambda_n(s) \right] d\lambda_n(t) \\
&= \int_{\mathbb{R}^n \setminus E} |f(x-t)| \left[\int_{\mathbb{R}^n} |g(t-s)h(s)|\, d\lambda_n(s) \right] d\lambda_n(t) \\
&= \int_{\mathbb{R}^n \setminus E} |f(x-t)|(|g|*|h|)(t)\, d\lambda_n(t) \\
&= \int_{\mathbb{R}^n} |f(x-t)|(|g|*|h|)(t)\, d\lambda_n(t)
\end{aligned}$$

가 되어 $x \in E(|f|, |g|*|h|)$와 $||F_x||_1 = \infty$가 서로 동치이다. 이와 비슷하게 $x \in E(|f|*|g|, |h|)$와 $||F_x||_1 = \infty$가 서로 동치임도 보일 수 있으므로 $E(|f|, |g|*|h|) = E(|f|*|g|, |h|)$임을 안다. □

Theorem 1.212 가측함수 $f, g, h : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $E(f,g), E(g,h)$가 영집합이면 $E := E(|f|, |g|*|h|) = E(|f|*|g|, |h|)$이고 $\mathbb{R}^n \setminus E$에서 $f*(g*h) = (f*g)*h$이다.

PROOF 정리 ? ? , ? ? 의 ii로부터 적당한 Borel 함수 $\widetilde{f}, \widetilde{g}, \widetilde{h} : \mathbb{R}^n \to \overline{\mathbb{R}}$가 존재하여 $f = \widetilde{f}$ (ae.), $g = \widetilde{g}$ (ae.), $h = \widetilde{h}$ (ae.)이고 정리 ? ? 와 위의 보조정리로부터 $E(|f|, |g|*|h|) = E(|\widetilde{f}|, |\widetilde{g}|*|\widetilde{h}|) = E(|\widetilde{f}|*|\widetilde{g}|, |\widetilde{h}|) = E(|f|*|g|, |h|)$이다. 이제 임의의 $x \in \mathbb{R}^n \setminus E$를 택하여 함수 $F_x : \mathbb{R}^{2n} \to \overline{\mathbb{R}}$를 $F_x : (s,t) \mapsto \widetilde{f}(x-t)\widetilde{g}(t-s)\widetilde{h}(s)$로 두면 명제 ? ? 의 증명에서와 비슷하게 하여 이가 Borel임을 쉽게 보일 수 있다. 여기서 만약 $x \in E(\widetilde{f}, \widetilde{g}*\widetilde{h})$이면 정리 ? ? 로부터 $\infty = \int_{\mathbb{R}^n} |\widetilde{f}(x-t)||(\widetilde{g}*\widetilde{h})(t)|\, d\lambda_n(t) \le \int_{\mathbb{R}^n} |\widetilde{f}(x-t)|(|\widetilde{g}|*|\widetilde{h}|)(t)\, d\lambda_n(t) < \infty$의 모순이 발생하므로 $x \notin E(\widetilde{f}, \widetilde{g}*\widetilde{h})$이고, 곧 Fubini의 정리, 정리 ? ? 와 $E(g,h) = E(\widetilde{g}, \widetilde{h})$가 영집합이라는 점으로부터

$$\begin{aligned}
\int_{\mathbb{R}^{2n}} F_x\, d\lambda_{2n} &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} F_x(s,t)\, d\lambda_n(s) \right] d\lambda_n(t) \\
&= \int_{\mathbb{R}^n} \widetilde{f}(x-t) \left[\int_{\mathbb{R}^n} \widetilde{g}(t-s)\widetilde{h}(s)\, d\lambda_n(s) \right] d\lambda_n(t) \\
&= \int_{\mathbb{R}^n \setminus E(\widetilde{g},\widetilde{h})} \widetilde{f}(x-t) \left[\int_{\mathbb{R}^n} \widetilde{g}(t-s)\widetilde{h}(s)\, d\lambda_n(s) \right] d\lambda_n(t) \\
&= \int_{\mathbb{R}^n \setminus E(\widetilde{g},\widetilde{h})} \widetilde{f}(x-t)(\widetilde{g}*\widetilde{h})(t)\, d\lambda_n(t)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^n} \widetilde{f}(x-t)(\widetilde{g}*\widetilde{h})(t)\,d\lambda_n(t) \\
&= [\widetilde{f}*(\widetilde{g}*\widetilde{h})](x) \\
&= [f*(g*h)](x)
\end{aligned}
$$

이다. 한편, 비슷한 방법으로 $\int_{\mathbb{R}^{2n}} F_x\,d\lambda_{2n} = [(f*g)*h](x)$임을 보일 수 있으므로 $\mathbb{R}^n \setminus E$에서 $f*(g*h) = (f*g)*h$임을 안다. □

비록 앞서 얻은 분배법칙과 결합법칙은 훌륭한 결과이지만, 각각이 성립하기 위한 조건들 때문에 실용성이 떨어지는 것이 사실이다. 이러한 조건들을 생략하기 위해서는 합성곱을 취하는 함수들의 종류를 조금 제한할 필요가 있다. 즉, 지금까지는 임의의 두 가측함수에 합성곱을 취하는 것을 생각했다면, 이를 조금 제한하여 합성곱을 취해 보자는 것이다. 물론, 이때 너무 제한적으로 그 종류를 제한하면 또다시 실용성을 잃고 말겠지만, 다행히 다음 정리는 적당히 적분가능한 함수로 제한하는 것만으로 우리의 목적을 달성할 수 있음을 말해준다.

Theorem 1.213 (Young's convolution inequality) 임의의 $1/p+1/q = 1+1/r$인 $p,q,r \geq 1$과 임의의 $f \in L^p(\lambda_n)$, $g \in L^q(\lambda_n)$에 대해 $E(f,g)$는 영집합이고 $||f*g||_r \leq ||f||_p||q||_q$이다. 따라서 $f*g \in L^r$이다.

PROOF 함수 $h : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$를 $h : x \mapsto \int_{\mathbb{R}^n} |f(x-t)g(t)|\,d\lambda_n(t)$로 두면 이는 명제 ? ? 로부터 well-define되고 $E(f,g) = E(|f|,|g|) = h^{-1}(\infty)$이다. 한편, 정리 ? ? 로부터 $|f*g| \leq h$이므로 $||h||_r \leq ||f||_p||g||_q$만 보일 수 있으면 $||f*g||_r \leq ||f||_p||g||_q$는 자명해지고 $h \in L^r$이 되어 정리 ? ? 의 ii로부터 h가 거의 어디서나 유한하므로 증명이 끝난다. 이를 위해 임의의 $x \in \mathbb{R}^n$를 택하여 함수 $f_x : \mathbb{R}^n \to \overline{\mathbb{R}}$를 $f_x : t \mapsto f(x-t)$로 두고 $\lambda = p(1-1/q)$라 하면 명제 ? ? 로부터 f_x는 가측이고 $0 \leq \lambda < 1$이다. 먼저 $\lambda > 0$인 경우를 생각하여 $s = p/\lambda$라 하면 $1/q+1/s = 1$이므로 Hölder의 부등식으로부터 $h(x) = \int_{\mathbb{R}^n} |f_x|^{\lambda}|f_x|^{1-\lambda}|g|\,d\lambda_n \leq |||f_x|^{1-\lambda}|g|||_q|||f_x|^{\lambda}||_s$이고, 곧

$$
\begin{aligned}
[h(x)]^q &\leq \left[\int_{\mathbb{R}^n} |f_x|^{(1-\lambda)q}|g|^q\,d\lambda_n\right]\left(\int_{\mathbb{R}^n} |f_x|^{\lambda s}\,d\lambda_n\right)^{q/s} \\
&= \left[\int_{\mathbb{R}^n} |f_x|^{(1-\lambda)q}|g|^q\,d\lambda_n\right]\left(\int_{\mathbb{R}^n} |f_x|^{p}\,d\lambda_n\right)^{\lambda q/p} \\
&= \left[\int_{\mathbb{R}^n} |f_x|^{(1-\lambda)q}|g|^q\,d\lambda_n\right]||f_x||_p^{\lambda q}
\end{aligned}
$$

이며 이는 $\lambda = 0$인 경우에도 성립함을 쉽게 알 수 있다. 그렇다면 $||f_x||_p = ||f||_p$임은 자명하므로 Minkowski의 적분부등식으로부터 $s' = r/q$에 대해

$$
\begin{aligned}
||h||_r^q &= \left(\int_{\mathbb{R}^n} h^r \, d\lambda_n \right)^{q/r} \\
&= \left(\int_{\mathbb{R}^n} h^{qs'} \, d\lambda_n \right)^{1/s'} \\
&\leq ||f||_p^{\lambda q} \left\{ \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |f(x-t)|^{(1-\lambda)q} |g(t)|^q \, d\lambda_n(t) \right]^{s'} d\lambda_n(x) \right\}^{1/s'} \\
&\leq ||f||_p^{\lambda q} \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |f(x-t)|^{(1-\lambda)qs'} |g(t)|^{qs'} \, d\lambda_n(x) \right]^{1/s'} d\lambda_n(t) \\
&= ||f||_p^{\lambda q} \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |f(x)|^{(1-\lambda)qs'} \, d\lambda_n(x) \right]^{1/s'} |g(t)|^q \, d\lambda_n(t) \\
&= ||f||_p^{\lambda q} ||f||_p^{(1-\lambda)q} ||g||_q^q \\
&= (||f||_p ||g||_q)^q
\end{aligned}
$$

에서 $||h||_r \leq ||f||_p ||g||_q$임을 안다. 여기서 마지막에서 두 번째 등호는 $(1-\lambda)qs' = (1-\lambda)r = p$에서 성립한다. □

Corollary 1.214 적분가능한 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $E(f, g)$는 영집합이고 $f * g$도 적분가능하다.

Proof 이는 Young의 합성곱 부등식에서 $p = q = r = 1$인 특별한 경우이다. □

따라서 f, g가 모두 적분가능한 경우에는 분배법칙과 결합법칙의 조건들을 생략하여 다음과 같이 간단히 쓸 수 있다. 이와 더불어 L^1 공간은 합성곱 $*$에 대해 닫혀있다는 이상의 결론은 기억해둘만한 가치가 있다.

Corollary 1.215 적분가능한 $f, g, h : \mathbb{R}^n \to \overline{\mathbb{R}}$와 $c \in \mathbb{R}$에 대해 다음이 성립한다.

i. (교환법칙) $c(f * g) = (cf) * g = f * (cg)$.
ii. (교환법칙) $f * g = g * f$.
iii. (결합법칙) $f * (g * h) = (f * g) * h$ (ae.).
iv. (분배법칙) $f * (g + h) = f * g + f * h$ (ae.).

Proof 이는 정리 ? ? , ? ? , ? ? 에 따름정리 ? ? 를 적용한 결과로 자명하다. □

이제 측도의 합성곱을 정의하는 것으로 이번 절의 후반부를 시작한다.

Proposition 1.216 Borel 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 함수 $F : \mathbb{R}^{2n} \to \overline{\mathbb{R}}$를 $F : (x, y) \mapsto f(x + y)$로 정의하면 이는 Borel이다.

PROOF 함수 $g:\mathbb{R}^{2n}\to\mathbb{R}^n$를 $g:(x,y)\mapsto x+y$로 두면 이가 연속이므로 곧 Borel이고, 정리 ? ? 에서 $F=f\circ g$는 Borel이다. □

Definition 1.217 Borel $\sigma-$대수 $\mathscr{B}_n$ 위의 측도 μ,ν에 대해 μ와 ν의 **합성곱(convolution)**을 $\mu*\nu:\mathscr{B}_n\to\overline{\mathbb{R}}_0^+$로 쓰고 $\mu*\nu:A\mapsto\int_{\mathbb{R}^{2n}}\mathbf{1}_A(x+y)\,d(\mu\otimes\nu)(x,y)$로 정의한다.

Proposition 1.218 Borel $\sigma-$대수 $\mathscr{B}_n$ 위의 측도 μ,ν에 대해 함수 $\mu*\nu$는 측도이다.

PROOF 우선 $(\mu*\nu)(\emptyset)=0$임은 분명하고, $\mathscr{B}_n$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 MCT로부터

$$\begin{aligned}(\mu*\nu)\left(\bigsqcup_{i=1}^{\infty}A_i\right)&=\int_{\mathbb{R}^{2n}}\mathbf{1}_{\bigsqcup_{i=1}^{\infty}A_i}(x+y)\,d(\mu\otimes\nu)(x,y)\\&=\int_{\mathbb{R}^{2n}}\sum_{i=1}^{\infty}\mathbf{1}_{A_i}(x+y)\,d(\mu\otimes\nu)(x,y)\\&=\sum_{i=1}^{\infty}\int_{\mathbb{R}^{2n}}\mathbf{1}_{A_i}(x+y)\,d(\mu\otimes\nu)(x,y)\\&=\sum_{i=1}^{\infty}(\mu*\nu)(A_i)\end{aligned}$$

이므로 $\mu*\nu$는 측도이다. □

명제 ? ? 로부터 $\mathbf{1}_A(x+y)$가 Borel이므로 측도의 합성곱은 well-define된다. 다음 순서는 측도의 합성곱의 기본적인 성질들을 살피는 것이다. 다행히 측도의 합성곱의 성질들은 함수의 합성곱의 성질들보다 쉽게 얻어진다.

Theorem 1.219 Borel $\sigma-$대수 $\mathscr{B}_n$ 위의 측도 μ,ν,ξ에 대해 $\xi=\mu*\nu$일 필요충분조건은 임의의 음이 아닌 Borel 함수 $f:\mathbb{R}^n\to\overline{\mathbb{R}}_0^+$에 대해 $\int_{\mathbb{R}^n}f\,d\xi=\int_{\mathbb{R}^{2n}}f(x+y)\,d(\mu\otimes\nu)(x,y)$인 것이다.

PROOF 충분조건임을 보이기 위해 먼저 f가 단순함수인 경우를 생각하여 이의 표준형을 $f=\sum_{i=1}^k a_i\mathbf{1}_{A_i}$라 하면 $\int_{\mathbb{R}^n}f\,d\xi=\sum_{i=1}^k a_i\xi(A_i)=\sum_{i=1}^k a_i\int_{\mathbb{R}^{2n}}\mathbf{1}_{A_i}(x+y)\,d(\mu\otimes\nu)(x,y)=\int_{\mathbb{R}^{2n}}f(x+y)\,d(\mu\otimes\nu)(x,y)$이다. 이제 일반적인 음이 아닌 Borel 함수 f를 생각하면 정리 ? ? 로부터 적당한 단순함수 $f_i:\mathbb{R}^n\to\mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i\uparrow f$이므로 MCT로부터 $\int_{\mathbb{R}^n}f\,d\xi=\lim_{i\to\infty}\int_{\mathbb{R}^n}f_i\,d\xi=\lim_{i\to\infty}\int_{\mathbb{R}^{2n}}f_i(x+y)\,d(\mu\otimes\nu)(x,y)=\int_{\mathbb{R}^{2n}}f(x+y)\,d(\mu\otimes\nu)(x,y)$가 되어 충분조건임이 보여진다.

반대로, 필요조건임을 보이는 것은 쉽다. 임의의 $A\in\mathscr{B}_n$에 대해 $\mathbf{1}_A$이 Borel 함수이므로 $\xi(A)=\int_{\mathbb{R}^n}\mathbf{1}_A\,d\xi=\int_{\mathbb{R}^{2n}}\mathbf{1}_A(x+y)\,d(\mu\otimes\nu)(x,y)=(\mu*\nu)(A)$에서 필요조건임이 분명하다. □

Theorem 1.220 Borel $\sigma-$대수 $\mathscr{B}_n$ 위의 측도 μ, ν, ξ와 $c \geq 0$에 대해 다음이 성립한다.

i. (교환법칙) $c(\mu * \nu) = (c\mu) * \nu = \mu * (c\nu)$.
ii. (교환법칙) $\mu * \nu = \nu * \mu$.
iii. (결합법칙) $\mu * (\nu * \xi) = (\mu * \nu) * \xi$.
iv. (분배법칙) $\mu * (\nu + \xi) = \mu * \nu + \mu * \xi$.

Proof 임의의 $A \in \mathscr{B}_n$를 택하자.

i. Fubini의 정리로부터

$$\begin{aligned}
[c(\mu * \nu)](A) &= c\int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu \otimes \nu)(x, y) \\
&= c\int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x)\right] d\nu(y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d(c\mu)(x)\right] d\nu(y) \\
&= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d[(c\mu) \otimes \nu](x, y) \\
&= [(c\mu) * \nu](A)
\end{aligned}$$

에서 $c(\mu * \nu) = (c\mu) * \nu$이고 이와 비슷하게 $c(\mu * \nu) = \mu * (c\nu)$도 보일 수 있다.

ii. Fubini의 정리로부터

$$\begin{aligned}
(\mu * \nu)(A) &= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu \otimes \nu)(x, y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x)\right] d\nu(y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(y)\right] d\nu(x) \\
&= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\nu \otimes \mu)(x, y) \\
&= (\nu * \mu)(A)
\end{aligned}$$

에서 $\mu * \nu = \nu * \mu$이다.

iii. Fubini의 정리와 정리 ? ? 로부터

$$\begin{aligned}
[\mu * (\nu * \xi)](A) &= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d[\mu \otimes (\nu * \xi)](x, y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x)\right] d(\nu * \xi)(y) \\
&= \int_{\mathbb{R}^{2n}} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y+z)\, d\mu(x)\right] d(\nu \otimes \xi)(y, z)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^n} \left\{ \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y+z)\, d\mu(x) \right] d\nu(y) \right\} d\xi(z) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y+z)\, d(\mu \otimes \nu)(x, y) \right] d\xi(z) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+z)\, d(\mu * \nu)(x) \right] d\xi(z) \\
&= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+z)\, d[(\mu * \nu) \otimes \xi](x, z) \\
&= [(\mu * \nu) * \xi](A)
\end{aligned}
$$

에서 $\mu * (\nu * \xi) = (\mu * \nu) * \xi$이다.

iv. Fubini의 정리로부터

$$
\begin{aligned}
[\mu * (\nu + \xi)](A) &= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d[\mu \otimes (\nu + \xi)](x, y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x) \right] d(\nu + \xi)(y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x) \right] d\nu(y) + \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x) \right] d\xi(y) \\
&= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu \otimes \nu)(x, y) + \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu \otimes \xi)(x, y) \\
&= (\mu * \nu)(A) + (\mu * \xi)(A)
\end{aligned}
$$

에서 $\mu * (\nu + \xi) = \mu * \nu + \mu * \xi$이다. □

한편, 함수의 합성곱과 측도의 합성곱은 다음 정리에 의해 서로 연결되어 있다.

Theorem 1.221 음이 아닌 적분가능한 Borel 함수 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$에 대해 $(\mu_n)_f * (\mu_n)_g = (\mu_n)_{f*g}$이다.

Proof 함수 $F : \mathbb{R}^{2n} \to \overline{\mathbb{R}}_0^+$를 $F : (x, y) \mapsto f(x - y)$로 두면 명제 ? ? 의 증명에서와 비슷하게 하여 이가 Borel임을 쉽게 보일 수 있다. 이제 임의의 $A \in \mathscr{B}_n$를 택하면 Fubini의 정리로부터

$$
\begin{aligned}
[(\mu_n)_f * (\mu_n)_g](A) &= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d[(\mu_n)_f \otimes (\mu_n)_g](x, y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d(\mu_n)_f(x) \right] d(\mu_n)_g(y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x)g(y)\mathbf{1}_A(x+y)\, d\mu_n(x) \right] d\mu_n(y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x-y)g(y)\mathbf{1}_A(x)\, d\mu_n(x) \right] d\mu_n(y)
\end{aligned}
$$

$$= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x-y)g(y)\mathbf{1}_A(x)\, d\mu_n(y) \right] d\mu_n(x)$$

인데, 정리 ? ? 와 따름정리 ? ? 로부터 $E(f,g)$가 영집합이고 $f*g$도 음이 아닌 적분가능한 Borel 함수이므로 $[(\mu_n)_f * (\mu_n)_g](A) = \int_{\mathbb{R}^n} \mathbf{1}_A(f*g)\, d\mu_n = \int_{\mathbb{R}^n} \mathbf{1}_A\, d(\mu_n)_{f*g} = (\mu_n)_{f*g}(A)$이다. □

나중에 사용할 정리 하나를 증명하는 것으로 이번 절을 끝마치도록 하자.

Theorem 1.222 Borel $\sigma-$대수 $\mathscr{B}_n$ 위의 측도 μ, ν와 임의의 $A \in \mathscr{B}_n$에 대해 함수 $x \mapsto \mu(A-x)$는 Borel이고 $(\mu * \nu)(A) = \int_{\mathbb{R}^n} \mu(A-x)\, d\nu(x)$이다.

Proof Fubini의 정리로부터

$$\begin{aligned}
(\mu*\nu)(A) &= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu\otimes\nu)(y,x) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(y) \right] d\nu(x) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_{A-x}(y)\, d\mu(y) \right] d\nu(x) \\
&= \int_{\mathbb{R}^n} \mu(A-x)\, d\nu(x)
\end{aligned}$$

이고 이때 함수 $x \mapsto \mu(A-x)$가 Borel임이 분명하다. □

1.15 Epilogue

본 장의 마지막을 장식할 이번 절에서는 별개의 절로 구성할 정도의 내용은 아니지만, 그래도 놓치기는 아까운 주제들에 대해 가볍게 알아보도록 하겠다. 다루게 될 주제는 측도의 분해, 셈측도, 그리고 벡터함수와 복소함수의 적분으로 크게 3가지이다.

1.15.1 Lebesgue's decomposition theorem

우리는 주어진 임의의 측도를 보다 단순하거나 특정한 구조를 가지는 측도들의 합으로 분해할 수 있도록 해 주는 정리들을 통틀어 흔히 '분해정리'라 부른다. 이런 분해정리들은 다분히 이론적인 도구로, 마치 인수분해를 통해 정수의 성질을 더 깊이 탐구할 수 있는 것처럼 분해정리들로 측도를 적당히 분해함으로써 측도론의 깊이있는 이론을 전개할 수 있다. 이렇게 측도론에서 큰 의의를 가지는 정리들이지만 분해정리가 본격적으로 필요할 정도의 추

상적인 측도론은 지나치다는 생각에 아쉽게도 이 책에서는 분해정리가 빛을 보지 못한 것이 사실이다. 그럼에도 불구하고, 이를 에필로그에서라도 뒤늦게 소개하는 까닭은 이가 확률론에서 그 자체로 나름 함의하는 바가 있고, 그 함의가 어쩌면 고등학교 때부터 궁금했을 수도 있는 '왜 이산확률변수와 연속확률변수에만 관심을 가지는가? '라는 질문에 대한 훌륭한 답변이 되기 때문이다. 그 답변은 확률의 현실적인 응용에서는 이산확률변수와 연속확률변수로 충분하다는 식의 공학도의 마인드가 물씬 풍기는 타협적인 답변보다 훨씬 지적으로 우아하고 만족스러울 것이다.

일단 이번 절은 에필로그인 만큼 그 답변을 듣는 것은 다음 장으로 잠시 미루어두고, 여기서는 분해정리를 증명하고, 내친 김에 이를 Radon-Nikodým 도함수에 간단히 응용하는 정도로 만족하도록 하자. 우리는 singular라는 개념과 함께 앞서 미쳐 살펴보지 못했던 측도의 절대연속에 관한 성질을 살펴보는 것으로 논의를 시작한다.

Definition 1.223 가측공간 $(X, \mathscr{A})$ 위의 측도 μ, ν에 대해 적당한 $A \in \mathscr{A}$가 존재하여 $\mu(A) = \nu(A^c) = 0$이면 이때 μ와 ν는 서로 **singular**하다고 하고 $\mu \perp \nu$로 쓴다.

측도 μ와 ν가 서로 singular하다는 것은, μ와 ν가 측도를 부여하는 데 있어 서로 공통된 정보를 사용하지 않는다는, 따라서 서로 독립적이라는 정도의 의미이다. 즉, 정의로부터 적당한 $A \in \mathscr{A}$가 존재하여 $\mu(A) = \nu(A^c) = 0$인데, 임의의 $B \in \mathscr{A}$를 생각하면 $\mu(B) = \mu(B \cap A) + \mu(B \cap A^c) = \mu(B \cap A^c)$이고 $\nu(B) = \nu(B \cap A) + \nu(B \cap A^c) = \nu(B \cap A)$이므로 μ는 $B \cap A^c$의 정보만으로 B에 측도를 부여하고, ν는 $B \cap A$의 정보만으로 B에 측도를 부여하여 이 둘은 B의 측도를 부여하는 데 있어 서로 공통된 정보를 사용하지 않는다.

한편, 앞서 소개한 절대연속은 이와 정반대의 의미이다. 예컨대 $\mu \ll \nu$라 하고 적당한 $A \in \mathscr{A}$에 대해 $\nu(A^c) = 0$이라 하면 ν는 측도를 부여하는 데 있어 A에 속한 정보만을 사용한다고 생각할 수 있다. 그런데 정의로부터 $\mu(A^c) = 0$이므로 μ 또한 A에 속한 정보만을 사용하고, A 밖의 정보는 사용하지 못한다. 즉, μ는 측도를 부여하는 데 있어 철저하게 ν가 사용하는 정보만 사용할 수 있고, 곧 μ는 ν에 종속적이다.

Proposition 1.224 가측공간 $(X, \mathscr{A})$ 위의 측도 μ, ν, ξ에 대해 다음이 성립한다.

i. 만약 $\mu \perp \xi$이고 $\nu \perp \xi$이면 $(\mu + \nu) \perp \xi$이다.
ii. 만약 $\mu \ll \xi$이고 $\nu \ll \xi$이면 $(\mu + \nu) \ll \xi$이다.
iii. 만약 $\mu \ll \xi$이고 $\nu \perp \xi$이면 $\mu \perp \nu$이다.
iv. 만약 $\mu \ll \xi$이고 $\mu \perp \xi$이면 $\mu = 0$이다.

PROOF i. 가정으로부터 적당한 $A, B \in \mathscr{A}$가 존재하여 $\mu(A) = \nu(B) = \xi(A^c) = \xi(B^c) = 0$이므로 $\mu(A \cap B) \leq \mu(A) = 0, \nu(A \cap B) \leq \nu(B) = 0, \xi((A \cap B)^c) = \xi(A^c \cup B^c) \leq \xi(A^c) + \xi(B^c) = 0$이고, 이로부터 $(\mu + \nu) \perp \xi$이다.

ii. 가정으로부터 임의의 $N \in \mathscr{A}$에 대해 $\xi(N) = 0$이면 $(\mu+\nu)(N) = \mu(N)+\nu(N) = 0$이므로 $(\mu+\nu) \ll \xi$임이 자명하다.

iii. 가정으로부터 적당한 $A \in \mathscr{A}$가 존재하여 $\nu(A) = \xi(A^c) = 0$인데, 다시 가정으로부터 $\mu(A^c) = 0$이 되어 $\mu \perp \nu$이다.

iv. 가정으로부터 적당한 $A \in \mathscr{A}$가 존재하여 $\mu(A) = \xi(A^c) = 0$인데, 다시 가정으로부터 $\mu(A^c) = 0$이므로 $\mu(X) = \mu(A \sqcup A^c) = \mu(A)+\mu(A^c) = 0$에서 $\mu = 0$이다. □

많은 분해정리 중에서 우리가 특히 주목할 것은 Lebesgue의 분해정리이다.

Theorem 1.225 (Lebesgue's decomposition theorem) 기측공간 $(X, \mathscr{A})$ 위의 σ−유한 측도 μ, ν를 생각하면 ν에 대해 절대연속인 $\mathscr{A}$ 위의 측도 μ_{ac}와 ν와 singular한 $\mathscr{A}$ 위의 측도 μ_{s}가 유일하게 존재하여 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{s}}$이다.

Proof 먼저 존재성을 보이자. 이를 위해 측도 $\xi = \mu+\nu$를 생각하면 이는 $\mathscr{A}$ 위의 σ−유한 측도이고 $\mu, \nu \ll \xi$임이 자명하므로 Radon-Nikodým 도함수 $f := d\mu/d\xi$, $g := d\nu/d\xi$가 존재한다. 이제 집합 $A = \{x \in X : g(x) > 0\}$를 생각하여 $\mu_{\mathrm{ac}} := \mu_A$, $\mu_{\mathrm{s}} := \mu_{A^c}$로 두면 이들은 모두 $\mathscr{A}$ 위의 측도이며 $\mu = \mu_{\mathrm{ac}}+\mu_{\mathrm{s}}$이다. 따라서 $\mu_{\mathrm{ac}} \ll \nu$와 $\mu_{\mathrm{s}} \perp \nu$만 보이면 존재성이 보여지는데, $\mu_{\mathrm{s}}(A) = \mu(A \cap A^c) = 0$이고 $\nu(A^c) = \int_{A^c} g\,d\xi = 0$이므로 $\mu_{\mathrm{s}} \perp \nu$임은 쉽게 알 수 있다. 다음으로, $\nu(N) = 0$인 임의의 $N \in \mathscr{A}$을 고정하면 $\int_X g\mathbf{1}_N\,d\xi = \int_N g\,d\xi = \nu(N) = 0$이므로 정리 ? ? 의 ii로부터 $g\mathbf{1}_N = 0$ (ξ−ae.)이다. 이로부터 $g\mathbf{1}_{A\cap N} = 0$ (ξ−ae.)인데 $A \cap N$ 위에서 $g > 0$이므로 $\xi(A \cap N) = 0$이 되어 정리 ? ? 의 i로부터 $\mu_{\mathrm{ac}}(N) = \mu(A \cap N) = \int_{A\cap N} f\,d\xi = 0$이고, 곧 $\mu_{\mathrm{ac}} \ll \nu$임을 안다.

다음으로, 유일성을 보이기 위해 $\mu_{\mathrm{ac}}, \mu'_{\mathrm{ac}} \ll \nu$이고 $\mu_{\mathrm{s}}, \mu'_{\mathrm{s}} \perp \nu$인 $\mathscr{A}$ 위의 측도 $\mu_{\mathrm{ac}}, \mu'_{\mathrm{ac}}$, $\mu_{\mathrm{s}}, \mu'_{\mathrm{s}}$가 존재하여 $\mu = \mu_{\mathrm{ac}}+\mu_{\mathrm{s}} = \mu'_{\mathrm{ac}}+\mu'_{\mathrm{s}}$라 하면 적당한 $A, A' \in \mathscr{A}$가 존재하여 $\mu_{\mathrm{s}}(A) = \mu'_{\mathrm{s}}(A') = \nu(A^c) = \nu(A'^c) = 0$이다. 이제 임의의 $B \in \mathscr{A}$를 고정하면 $\mu_{\mathrm{s}}(B\cap A\cap A') \le \mu_{\mathrm{s}}(A) = 0$, $\mu'_{\mathrm{s}}(B\cap A\cap A') \le \mu'_{\mathrm{s}}(A') = 0$이고, 이로부터 $\mu_{\mathrm{ac}}(B\cap A\cap A') = \mu(B\cap A\cap A') = \mu'_{\mathrm{ac}}(B\cap A\cap A')$임을 안다. 나아가 $\nu(B\cap(A\cap A')^c) \le \nu((A\cap A')^c) = \nu(A^c \cup A'^c) \le \nu(A)+\nu(A'^c) = 0$이므로 $\mu_{\mathrm{ac}}(B\cap(A\cap A')^c) = \mu'_{\mathrm{ac}}(B\cap(A\cap A')^c) = 0$이 되어 $\mu_{\mathrm{s}}(B\cap(A\cap A')^c) = \mu(B\cap(A\cap A')^c) = \mu'_{\mathrm{s}}(B\cap(A\cap A')^c)$임을 안다. 이상을 종합하면

$$
\begin{aligned}
\mu_{\mathrm{ac}}(B) &= \mu_{\mathrm{ac}}(B\cap A\cap A') + \mu_{\mathrm{ac}}(B\cap(A\cap A')^c) \\
&= \mu_{\mathrm{ac}}(B\cap A\cap A') \\
&= \mu'_{\mathrm{ac}}(B\cap A\cap A') \\
&= \mu'_{\mathrm{ac}}(B\cap A\cap A') + \mu'_{\mathrm{ac}}(B\cap(A\cap A')^c) \\
&= \mu'_{\mathrm{ac}}(B)
\end{aligned}
$$

에서 $\mu_{\mathrm{ac}} = \mu'_{\mathrm{ac}}$가 보여지고 비슷하게 $\mu_{\mathrm{s}} = \mu'_{\mathrm{s}}$도 보여지므로 유일성이 증명된다. □

약간의 노력으로 Lebesgue 분해정리를 조금 더 refine할 수 있다.

Definition 1.226 측도공간 $(X, \mathscr{A}, \mu)$와 집합 $A \in \mathscr{A}$에 대해 $\mu(A^c) = 0$이면 이때의 집합 A를 측도 μ의 **지지집합(support)**이라 하고 $\operatorname{supp}\mu = A$로 쓴다. 나아가, μ가 가산인 지지집합을 가지면 이때의 측도 μ를 **이산측도(discrete measure)**라 한다.

Definition 1.227 가측공간 $(X, \mathscr{A})$ 위의 측도 μ, ν를 생각하고 X의 모든 한원소 집합이 $\mathscr{A}$에 속한다고 하자. 만약 $\mu \perp \nu$이고 X의 모든 한원소 집합이 $\mu-$영집합이면 이때 μ는 ν에 대해 **특이연속(singular continuous)**이라 한다.

일반적인 측도는 전체 공간 X의 측도가 X 전반에 걸쳐 연속적으로 퍼져있다면 이산측도는 가산개의 점에만 전체 공간의 측도가 띄엄띄엄 집중되어있는 느낌이다. 흔히 이렇게 측도가 집중된 듯한 점들을 point mass라 부르곤 한다. 한편, μ가 ν에 대해 특이연속이라는 것은 μ와 ν가 서로 독립적일 뿐만 아니라 μ가 point mass를 가지지 않음을 의미한다.

Lemma 1.228 $\sigma-$유한 측도공간 $(X, \mathscr{A}, \mu)$에 대해 X의 모든 한원소 집합이 $\mathscr{A}$에 속하다면 $(X, \mathscr{A}, \mu)$는 가산개의 point mass를 가진다. 즉, 집합 $\{x \in X : \mu\{x\} > 0\}$은 가산이다.

Proof 주어진 집합을 A라 하고 모순을 유도하기 위해 A가 비가산이라 하자. 그렇다면 가정으로부터 $\mathscr{A}$에 속하는 적당한 집합열 $\{B_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $\mu(B_i) < \infty$이고 $\bigcup_{i=1}^{\infty} B_i = X$이므로 $A \cap B_{i_0}$가 비가산인 $i_0 \in \mathbb{N}$가 적어도 하나 존재하여 집합열 $\{C_i\}$를 $C_i := \{x \in B_{i_0} : \mu\{x\} > 1/i\}$로 두면 이는 증가하는 집합열이 되어 $C_i \uparrow A \cap B_{i_0}$이다. 그렇다면 다시 적당한 $i_1 \in \mathbb{N}$이 존재하여 C_{i_1}이 비가산이다. 이제 여기서 서로다른 가산개의 원소를 뽑아 수열 $\{x_i\}$를 구성하면 $\mu\{x_1, x_2, \cdots\} = \sum_{i=1}^{\infty} \mu\{x_i\} > \sum_{i=1}^{\infty} 1/i_1 = \infty$인 동시에 $\mu\{x_1, x_2, \cdots\} \leq \mu(C_{i_1}) \leq \mu(B_{i_0}) < \infty$가 되어 모순이 발생한다. 증명은 이로써 충분하다. □

Theorem 1.229 (Lebesgue's decomposition theorem) 가측공간 $(X, \mathscr{A})$에 대해 $\mathscr{A}$ 위의 $\sigma-$유한 측도 μ, ν를 생각하고 X의 모든 한원소 집합이 $\mathscr{A}$에 속한다고 하자. 그렇다면 ν에 대해 절대연속인 $\mathscr{A}$ 위의 측도 μ_{ac}, ν에 대해 특이연속인 $\mathscr{A}$ 위의 측도 μ_{sc}, $\mathscr{A}$ 위의 이산측도 μ_{pp}가 유일하게 존재하여 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{sc}} + \mu_{\mathrm{pp}}$이다.

Proof 먼저 존재성을 보이자. 앞서 보인 Lebesgue의 분해정리로부터 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{s}}$이고 $\mu_{\mathrm{ac}} \ll \nu, \mu_{\mathrm{s}} \perp \nu$인 $\mathscr{A}$ 위의 측도 $\mu_{\mathrm{ac}}, \mu_{\mathrm{s}}$가 유일하게 존재한다. 여기서 μ가 $\sigma-$유한하므로 μ_{s}도 $\sigma-$유한하여 집합 $A = \{x \in X : \mu_{\mathrm{s}}\{x\} > 0\}$를 생각하면 위의 보조정리로부터 이는 가산이다. 이제 측도 $\mu_{\mathrm{pp}} = (\mu_{\mathrm{s}})_A, \mu_{\mathrm{sc}} = (\mu_{\mathrm{s}})_{A^c}$를 생각하면 이는 $\mathscr{A}$ 위의 측도이고 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{sc}} + \mu_{\mathrm{pp}}$가 되어 μ_{sc}가 ν에 대해 특이연속이고 μ_{pp}가 이산측도라는 점만 보이면 존재성이 보여진다. 우선 $\mu_{\mathrm{s}} \perp \nu$에서 적당한 $B \in \mathscr{A}$가 존재하여 $\mu_{\mathrm{s}}(B) = \nu(B^c) = 0$이므로 $\mu_{\mathrm{sc}}(B) = \mu_{\mathrm{s}}(B \cap A^c) \leq \mu_{\mathrm{s}}(B) = 0$에서 $\mu_{\mathrm{sc}} \perp \nu$임을 안다. 나아가 임의의 $x \in X$에 대해 만

약 $x \in A$라면 $\mu_{\text{sc}}\{x\} = \mu_{\text{s}}(\{x\} \cap A^c) = 0$이고 반대로 $x \notin A$라도 $\mu_{\text{sc}}\{x\} = \mu_{\text{s}}\{x\} = 0$이므로 곧 어느 경우에나 $\mu_{\text{sc}}\{x\} = 0$이 되어 μ_{sc}는 ν에 대해 특이연속이다. 한편, $\mu_{\text{pp}}(A^c) = \mu_{\text{s}}(A^c \cap A) = 0$에서 가산인 A가 μ_{pp}의 지지집합이 되어 이가 이산측도임도 쉽게 알 수 있다.

다음으로, 유일성을 보이기 위해서는 ν에 대해 특이연속인 $\mathscr{A}$ 위의 측도 $\mu_{\text{sc}}, \mu'_{\text{sc}}$와 $\mathscr{A}$ 위의 이산측도 $\mu_{\text{pp}}, \mu'_{\text{pp}}$에 대해 $\mu_{\text{s}} = \mu_{\text{sc}} + \mu_{\text{pp}} = \mu'_{\text{sc}} + \mu'_{\text{pp}}$이면 $\mu_{\text{sc}} = \mu'_{\text{sc}}, \mu_{\text{pp}} = \mu'_{\text{pp}}$임을 보이면 충분하다. 여기서 μ_{pp}와 μ'_{pp}의 가산 지지집합을 각각 $A, A' \in \mathscr{A}$라 하면 $\mu_{\text{pp}}((A \cup A')^c) = \mu_{\text{pp}}(A^c \cap A'^c) \leq \mu_{\text{pp}}(A^c) = 0$이고 비슷하게 $\mu'_{\text{pp}}((A \cup A')^c) = 0$이므로 WLOG, 필요하다면 A와 A'를 $A \cup A'$로 바꾸어 μ_{pp}와 μ'_{pp}가 공통의 가산 지지집합 $A \in \mathscr{A}$를 가진다고 해도 된다. 이제 임의의 $B \in \mathscr{A}$를 고정하면 $\mu_{\text{pp}}(B \cap A^c) \leq \mu_{\text{pp}}(A^c) = 0, \mu'_{\text{pp}}(B \cap A^c) \leq \mu'_{\text{pp}}(A^c) = 0$이다. 나아가, A의 가산개의 원소를 나열하여 $x_1, x_2, \cdots$라 하면 $\mu_{\text{sc}}(B \cap A) \leq \mu_{\text{sc}}(A) = \sum_{i=1}^{k} \mu_{\text{sc}}\{x_i\} = 0$이고 (여기서 k는 유한할 수도 있고, ∞일 수도 있다.) 비슷하게 $\mu'_{\text{sc}}(B \cap A) = 0$이므로 곧 $\mu_{\text{pp}}(B \cap A) = \mu_{\text{s}}(B \cap A) = \mu'_{\text{pp}}(B \cap A)$가 되어 이상으로부터

$$
\begin{aligned}
\mu_{\text{pp}}(B) &= \mu_{\text{pp}}(B \cap A) + \mu_{\text{pp}}(B \cap A^c) \\
&= \mu_{\text{pp}}(B \cap A) \\
&= \mu'_{\text{pp}}(B \cap A) \\
&= \mu'_{\text{pp}}(B \cap A) + \mu'_{\text{pp}}(B \cap A^c) \\
&= \mu'_{\text{pp}}(B)
\end{aligned}
$$

에서 $\mu_{\text{pp}} = \mu'_{\text{pp}}$임이 보여진다. 한편, 가정으로부터 적당한 $C, C' \in \mathscr{A}$가 존재하여 $\mu_{\text{sc}}(C) = \mu'_{\text{sc}}(C') = \nu(C^c) = \nu(C'^c) = 0$인데, 앞선 결과로부터 $\mu_{\text{sc}}(A \cup C) \leq \mu_{\text{sc}}(A) + \mu_{\text{sc}}(C) = 0$이고, 비슷하게 $\mu'_{\text{sc}}(A \cup C') = 0$이므로 WLOG, 필요하다면 C, C'를 각각 $A \cup C, A \cup C'$로 바꾸어 $A \subseteq C, C'$라 해도 된다. 그렇다면 $\mu_{\text{sc}}(B \cap C \cap C') \leq \mu_{\text{sc}}(C) = 0, \mu'_{\text{sc}}(B \cap C \cap C') \leq \mu'_{\text{sc}}(C') = 0$이고 $\mu'_{\text{pp}}(B \cap (C \cap C')^c) = \mu_{\text{pp}}(B \cap (C \cap C')^c) \leq \mu_{\text{pp}}(A^c) = 0$이므로 $\mu_{\text{sc}}(B \cap (C \cap C')^c) = \mu_{\text{s}}(B \cap (C \cap C')^c) = \mu'_{\text{sc}}(B \cap (C \cap C')^c)$이다. 이상을 종합하면

$$
\begin{aligned}
\mu_{\text{sc}}(B) &= \mu_{\text{sc}}(B \cap C \cap C') + \mu_{\text{sc}}(B \cap (C \cap C')^c) \\
&= \mu_{\text{sc}}(B \cap C \cap C') \\
&= \mu'_{\text{sc}}(B \cap C \cap C') \\
&= \mu'_{\text{sc}}(B \cap C \cap C') + \mu'_{\text{sc}}(B \cap (C \cap C')^c) \\
&= \mu'_{\text{sc}}(B)
\end{aligned}
$$

에서 $\mu_{\text{sc}} = \mu'_{\text{sc}}$도 보여지므로 유일성이 증명된다. □

따라서 Lebesgue의 분해정리는 임의의 측도 μ를 기준 측도 ν에 종속적인 측도 하나, ν와 서로 독립이고 point mass 없는 측도 하나, point mass로만 이루어진 측도 하나로 분해시켜주며, 이들을 각각 **절대연속성분(absolutely continuous component)**, **특이연속성분(singular continuous component)**, **순수 점 성분(pure point component)**이라 한다. 이제 예고했던 바와 같이 Lebesgue의 분해정리를 간단히 응용하여 Radon-Nikodým 도함수를 일반화시켜보자.

Definition 1.230 가측공간 $(X,\mathscr{A})$ 위의 $\sigma-$유한 측도 μ,ν에 대해 μ가 $\mu_{\mathrm{ac}} \ll \nu$이고 $\mu_{\mathrm{s}} \perp \nu$인 $\mathscr{A}$ 위의 측도 $\mu_{\mathrm{ac}},\mu_{\mathrm{s}}$로써 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{s}}$와 같이 Lebesgue 분해된다고 하자. 이때 μ의 ν에 대한 **Radon-Nikodým 도함수(− derivative)**를 $d\mu_{\mathrm{ac}}/d\nu$로 정의하고 $d\mu/d\nu$로 쓴다.

위의 정의에서 만약 $\mu \ll \nu$이면 Lebesgue 분해의 유일성으로부터 $\mu_{\mathrm{ac}} = \mu, \mu_{\mathrm{s}} = 0$이므로 위의 정의에서의 Radon-Nikodým 도함수가 정의 ? ? 에서의 Radon-Nikodým 도함수와 같아진다. 한편, 이렇게 일반화된 Radon-Nikodým 도함수에 대해 따름정리 ? ? 는 다소 불완전한 형태로만 성립한다.

Corollary 1.231 가측공간 $(X,\mathscr{A})$ 위의 $\sigma-$유한 측도 μ,ν와 임의의 음이 아닌 가측함수 $f:X\to\overline{\mathbb{R}}_0^+$에 대해 $\int_X f\,d\mu \geq \int_X f(d\mu/d\nu)\,d\nu$이다.

Proof 측도 μ가 $\mu_{\mathrm{ac}} \ll \nu$이고 $\mu_{\mathrm{s}} \perp \nu$인 $\mathscr{A}$ 위의 측도 $\mu_{\mathrm{ac}},\mu_{\mathrm{s}}$로써 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{s}}$와 같이 Lebesgue 분해된다고 하자. 그렇다면 $\int_X f\,d\mu = \int_X f\,d(\mu_{\mathrm{ac}} + \mu_{\mathrm{s}}) = \int_X f\,d\mu_{\mathrm{ac}} + \int_X f\,d\mu_{\mathrm{s}} \geq \int_X f\,d\mu_{\mathrm{ac}} = \int_X f(d\mu/d\nu)\,d\nu$이다. □

1.15.2 Counting measure

여태껏 우리는 Lebesgue 측도와 Borel 측도에 지대한 관심을 기울여왔는데, 확률론에서는 이만큼이나 중요한 역할을 하는 '셈측도'라는 측도가 하나 더 있다. 이렇게나 중요한 셈측도를 Lebesgue 측도와 달리 본 장의 끝에서야 소개하는 것은 이가 매우 단순한 측도이기 때문이다. 별도의 구성 과정도 필요없고, 관련된 성질들도 우리가 이미 알고 있는 것을 측도론의 언어로 번역한 것에 불과하다.

Definition 1.232 가측공간 $(X,\mathscr{A})$에 대해 다음과 같이 정의되는 함수 $\#:\mathscr{A}\to\overline{\mathbb{R}}_0^+$를 $\mathscr{A}$ 위의 **셈측도(counting measure)**라 한다.

$$\#:A\mapsto\begin{cases}|A| & A\text{가 유한한 경우}\\ \infty & \text{ow.}\end{cases}$$

Proposition 1.233 가측공간 $(X,\mathscr{A})$에 대해 셈측도 #는 측도이다. 따라서 $(X,\mathscr{A},\#)$는 측도공간을 이룬다.

PROOF 우선 $\#(\emptyset)=0$임은 분명하다. 이제 $\mathscr{A}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$를 생각하자. 만약 $\{A_i\}$가 공집합이 아닌 집합을 무한히 많이 포함하거나 적어도 하나의 무한집합을 포함한다면 $\bigsqcup_{i=1}^{\infty}A_i$가 무한집합이므로 $\#(\bigsqcup_{i=1}^{\infty}A_i)=|\bigsqcup_{i=1}^{\infty}A_i|=\infty=\sum_{i=1}^{\infty}|A_i|=\sum_{i=1}^{\infty}\#(A_i)$임이 자명하다. 반대로 $\{A_i\}$가 오직 유한개의 유한집합으로만 이루어져있다면 적당한 $i_0\in\mathbb{N}$에 대해 $i>i_0$이면 A_i가 공집합이고 $i\le i_0$이면 A_i는 유한집합이므로 $\bigsqcup_{i=1}^{\infty}A_i=\bigsqcup_{i=1}^{i_0}A_i$가 유한집합이 되어 $\#(\bigsqcup_{i=1}^{\infty}A_i)=\#(\bigsqcup_{i=1}^{i_0}A_i)=|\bigsqcup_{i=1}^{i_0}A_i|=\sum_{i=1}^{i_0}|A_i|=\sum_{i=1}^{i_0}\#(A_i)=\sum_{i=1}^{\infty}\#(A_i)$이다. 이상으로부터 #가 측도임을 안다. □

보통 특별히 정하지 않는 이상 집합 X에서의 셈측도는 $\mathscr{P}(X)$에서 정의된 것으로 보고, 만약 집합 X도 주어지지 않았다면 $(\mathbb{N},\mathscr{P}(\mathbb{N}))$에서 정의된 것으로 본다.

Proposition 1.234 측도공간 $(X,\mathscr{A},\#)$에서 영집합은 $\emptyset$ 뿐이다.

PROOF 임의의 $N\in\mathscr{A}$에 대해 셈측도의 정의로부터 $\#(N)=0$은 $N=\emptyset$를 뜻하므로 이는 자명하다. □

이로부터 셈측도가 주어진 측도공간에서 거의 어디서나 성립하는 성질은 곧 항상 성립한다.

Theorem 1.235 집합 X를 가산이라 하면 측도공간 $(X,\mathscr{A},\#)$에서 정의된 음이 아닌 가측함수 $f:X\to\overline{\mathbb{R}}_0^+$에 대해 $\int_X f\,d\#=\sum_{x\in X}f(x)$이다. 한편, 적분가능한(혹은 $\sum_{x\in X}f(x)$가 well-define되는[33]가측함수) $f:X\to\overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며 이때 $\sum_{x\in X}f(x)$는 well-define된다(혹은 f는 적분가능하다).

PROOF 간결한 논의를 위해 X가 무한집합인 경우에 대해서만 보이도록 하자. (집합 X가 유한집합인 경우에도 이와 비슷하게 보일 수 있다.) 집합 X의 원소를 나열하여 $\{x_i\}$라 하라 하고 함수열 $\{f_j\}$를 $f_j=\sum_{i=1}^{j}f(x_i)\mathbf{1}_{\{x_i\}}$로 두면 $\int_X f_j\,d\#=\sum_{i=1}^{j}f(x_i)\int_X\mathbf{1}_{\{x_i\}}\,d\#=\sum_{i=1}^{j}f(x_i)\#\{x_i\}=\sum_{i=1}^{j}f(x_i)$이다. 그렇다면 $f_j\uparrow f$에서 MCT로부터 $\int_X f\,d\#=\lim_{j\to\infty}\int_X f_j\,d\#=\lim_{j\to\infty}\sum_{i=1}^{j}f(x_i)=\sum_{x\in X}f(x)$에서 정리가 성립한다.

한편, 적분가능한 $f:X\to\overline{\mathbb{R}}$에 대해서는 우선 정리 ? ? 의 ii, 명제 ? ? 로부터 f가 $\pm\infty$를 그 함숫값으로 갖지 아니하여 $\infty-\infty$와 같은 부정형이 합의 과정에서 등장하지 않고, 정리 ? ? 의 i로부터 $|f|$가 적분가능하여 앞선 결과로부터 $\sum_{x\in X}|f(x)|=\int_X|f|\,d\#<\infty$에서 $\sum_{x\in X}f(x)$가 절대수렴하므로 $\sum_{x\in X}|f(x)|$는 well-define된다. 이제 $Y=f^{-1}([0,\infty)), Z=f^{-1}((-\infty,0))$에 대해 $\int_X f\,d\#=\int_X f_+\,d\#-\int_X f_-\,d\#=\sum_{x\in X}f_+-\sum_{x\in X}f_-=\sum_{x\in Y}f(x)+\sum_{x\in Z}f(x)=\sum_{x\in X}f(x)$이므로 증명이 끝나고, $\sum_{x\in X}f(x)$가 well-define되는 가측함수 $f:X\to\overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

이를 조금 더 친숙한 $\mathbb{N}$에서의 결과로 옮기면 다음과 같다.

Corollary 1.236 함수 $f:\mathbb{N}\to\overline{\mathbb{R}}$에 대해 수열 $\{a_i\}$를 $a_i := f(i)$로 두자. 만약 $\sum_{i=1}^{\infty} a_i$가 양항급수이거나 절대수렴하면 $\int_{\mathbb{N}} f\,d\# = \sum_{i=1}^{\infty} a_i$이다.

Proof 이는 정리 ? ? 와 가측공간 $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$에서 정의된 함수 $f:\mathbb{N}\to\overline{\mathbb{R}}$는 모두 가측이라는 점에서 자명하다. □

따라서 셈측도가 주어진 측도공간에 있어 적분은 합과 같다. 담담하게 말했지만, 이는 정말 흥분되는 결과이다. 아마 다들 합은 이산세계의 적분이고, 적분은 연속세계의 합이라는 수학적 직관을 어느정도 가지고 있을 것인데, 위의 결과는 이러한 직관을 확인시켜주는 동시에 더 이상 합과 적분을 구별하여 다룰 필요가 없음을 뜻한다. 우리가 적분을 통해 얻는 모든 결과는 자연스럽게 합에 대한 결과를 따름정리의 형태로 함의한다.

Corollary 1.237 다음이 성립한다.

i. (MCT) 음이 아닌 수열 $\{a_{ij}\}$에 대해 임의의 $i,j\in\mathbb{N}$에 대해 $a_{ij}\le a_{ij+1}$이고, 적당한 음이 아닌 수열 $\{a_i\}$가 존재하여 각 $i\in\mathbb{N}$에 대해 $j\to\infty$이면 $a_{ij}\uparrow a_i$라 하자. 그렇다면 $j\to\infty$일때 $\sum_{i=1}^{\infty} a_{ij}\uparrow\sum_{i=1}^{\infty} a_i$이다.
ii. (Fatou의 보조정리) 음이 아닌 수열 $\{a_{ij}\}$에 대해 $\liminf_{j\to\infty}\sum_{i=1}^{\infty} a_{ij}\ge\sum_{i=1}^{\infty}\liminf_{j\to\infty} a_{ij}$이다.
iii. (DCT) 수열 $\{a_{ij}\}$에 대해 적당한 수열 $\{a_i\}$가 존재하여 각 $i\in\mathbb{N}$에 대해 $j\to\infty$이면 $a_{ij}\to a_i$라 하자. 또한, 적당한 음이 아닌 수열 $\{M_i\}$가 존재하여 $\sum_{i=1}^{\infty} M_i$가 수렴하고 모든 $i,j\in\mathbb{N}$에 대해 $|a_{ij}|\le M_i$라 하자. 그렇다면 각 $j\in\mathbb{N}$에 대해 $\sum_{i=1}^{\infty} a_{ij}$와 $\sum_{i=1}^{\infty} a_i$는 절대수렴하고 $j\to\infty$이면 $\sum_{i=1}^{\infty} a_{ij}\to\sum_{i=1}^{\infty} a_i$이다.
iv. (Fubini) 음이 아닌 수열 $\{a_{ij}\}$에 대해 $\sum_{(i,j)\in\mathbb{N}^2} a_{ij}=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij}=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}$이다. 한편, $\sum_{(i,j)\in\mathbb{N}^2} a_{ij}$가 절대수렴하는 수열 $\{a_{ij}\}$에 대해서도 동일한 결과가 성립하며 이때 $\sum_{i=1}^{\infty} a_{ij}, \sum_{j=1}^{\infty} a_{ij}$는 절대수렴하고 수열 $\{b_j\}, \{c_i\}$를 각각 $b_j := \sum_{i=1}^{\infty} a_{ij}$, $c_i := \sum_{j=1}^{\infty} a_{ij}$라 하면 $\sum_{j=1}^{\infty} b_j, \sum_{i=1}^{\infty} c_i$도 절대수렴한다.

Proof 이는 셈측도에 각각 MCT, Fatou의 보조정리, DCT, Fubini의 정리를 적용한 결과로 정리 ? ? 와 가측공간 $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$에서 정의된 함수 $f:\mathbb{N}\to\overline{\mathbb{R}}$는 모두 가측이라는 점에서 자명하다. □

Corollary 1.238 다음이 성립한다.

i. 열린집합 $U\subseteq\mathbb{R}^n$에서 정의된 연속함수 $f_i:U\to\mathbb{R}$의 열 $\{f_i\}$가 다음 조건을 만족한다고 하자.

   a. 임의의 $x\in U$에 대해 $\sum_{i=1}^{\infty} f_i(x)$가 절대수렴한다.

b. 적당한 음이 아닌 수열 $\{M_i\}$가 존재하여 $\sum_{i=1}^{\infty} M_i$가 수렴하고 각 $i \in \mathbb{N}_0$에 대해 $|f_i| \leq M_i$이다.

그렇다면 함수 $x \mapsto \sum_{i=1}^{\infty} f_i(x)$는 연속이다.

ii. (Leibniz의 법칙) 열린집합 $U \subseteq \mathbb{R}$에서 정의된 미분가능한 함수 $f_i : U \to \mathbb{R}$의 열 $\{f_i\}$가 다음 조건을 만족한다고 하자.

a. 임의의 $x \in U$에 대해 $\sum_{i=1}^{\infty} f_i(x)$가 절대수렴한다.
b. 적당한 음이 아닌 수열 $\{M_i\}$가 존재하여 $\sum_{i=1}^{\infty} M_i$가 수렴하고 각 $i \in \mathbb{N}_0$에 대해 $|f_i'| \leq M_i$이다.

그렇다면 함수 $x \mapsto \sum_{i=1}^{\infty} f_i(x)$는 미분가능하고 임의의 $x \in U$에 대해 $\sum_{i=1}^{\infty} f'(x)$가 절대수렴하며 다음이 성립한다.

$$\frac{d}{dx}\sum_{i=1}^{\infty} f_i(x) = \sum_{i=1}^{\infty} f_i'(x)$$

Proof 이는 셈측도에 각각 따름정리 ? ? , Leibniz의 법칙을 적용한 결과로 정리 ? ? 와 가측공간 $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$에서 정의된 함수 $f : \mathbb{N} \to \overline{\mathbb{R}}$는 모두 가측이라는 점에서 자명하다. □

마지막으로, Radon-Nikodým 도함수가 셈측도에 대해 어떻게 적용되는지 살펴보자.

Proposition 1.239 가측공간 $(X, \mathscr{A})$ 위의 임의의 측도 μ와 셈측도 #에 대해 $\mu \ll \#$이다.

Proof 명제 ? ? 로부터 $\#(N) = 0$인 집합 $N \in \mathscr{A}$은 공집합뿐이고, 따라서 $\mu \ll \#$이다. □

Theorem 1.240 측도공간 $(X, \mathscr{A}, \mu)$에 대해 X가 가산이고 이의 모든 한원소 집합이 $\mathscr{A}$에 속한다면 $(d\mu/d\#)(x) = \mu\{x\}$이다.

Proof 임의의 $A \in \mathscr{A}$에 대해 이가 가산이므로 $\mu(A) = \sum_{x\in A} \mu\{x\} = \int_A \mu\{x\}\, d\#$이고, Radon-Nikodým 정리와 명제 ? ? , ? ? 로부터 $(d\mu/d\#)(x) = \mu\{x\}$이다. □

1.15.3 Integration of vector/complex-valued functions

지금까지의 우리는 공역이 $\overline{\mathbb{R}}$인 함수의 적분만 생각해왔는데, 때로는 (주로 표기상의 편의를 위해) 공역이 $\mathbb{R}^n$이거나 심지어 $\mathbb{C}$인 함수 즉, 벡터함수나 복소함수의 적분을 생각해야 하는 경우도 있다. 함숫값으로 벡터는 그렇다 쳐도 복소수를 가진다니 복소해석학을 배우지도 않은 상태에서 이게 가능할까 싶을 수도 있지만, 여기서는 아주 간단한 내용만 다룰 것이므로 걱정할 필요는 없다.

Definition 1.241 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f : X \to \mathbb{R}^n$에 대해 f의 모든 성분이 적분가능하면 이때 f를 **(μ−Lebesgue) 적분가능(− integrable)**하다고 한다. 나아가 적분가능한 $f : X \to \mathbb{R}^n$에 대해 μ에 대한 f의 **(Lebesgue) 적분(− integral)**을 $\int_X f(x)\, d\mu(x)$ 혹은 간단히 $\int_X f\, d\mu$로 쓰고 $\int_X f\, d\mu = (\int_X f_i\, d\mu)_i$로 정의한다.

앞서 정리 ? ? 에서 보였던 적분의 기본적인 성질들도 대부분 비슷하게 성립한다.

Corollary 1.242 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 적분가능한 함수 $f, g : X \to \mathbb{R}^n$에 대해 다음이 성립한다.

i. 임의의 $c \in \mathbb{R}$에 대해 cf가 적분가능하고 $\int_X cf\, d\mu = c \int_X f\, d\mu$이다.
ii. 함수 $f + g$가 적분가능하고 $\int_X f + g\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu$이다.
iii. 임의의 $x \in \mathbb{R}^n$에 대해 $\langle x, f\rangle$가 적분가능하고 $\int_X \langle x, f\rangle\, d\mu = \langle x, \int_X f\, d\mu\rangle$이다.
iv. $||\int_X f\, d\mu|| \leq \int_X ||f||\, d\mu$.

Proof i. ii. 이는 정의와 적분의 기본적인 성질로부터 자명하다.

iii. 각 f_i가 적분가능하므로 $\langle x, f\rangle = \sum_{i=1}^n x_i f_i$가 적분가능함은 분명하고, 곧 $\int_X \langle x, f\rangle\, d\mu = \int_X \sum_{i=1}^n x_i f_i\, d\mu = \sum_{i=1}^n x_i \int_X f_i\, d\mu = \langle x, \int_X f\, d\mu\rangle$이다.

iv. 표기의 편의를 위해 $I = \int_X f\, d\mu$라 하자. 나아가 만약 $I = 0$이면 부등식이 자명하므로 $I \neq 0$이라 하자. 그렇다면 Cauchy-Schwarz 부등식으로부터 $||I||^2 = \langle I, \int_X f\, d\mu\rangle = \int_X \langle I, f\rangle\, d\mu \leq \int_X ||I||||f||\, d\mu = ||I|| \int_X ||f||\, d\mu$가 되어 정리가 성립한다. □

Corollary 1.243 (Fundamental theorem of calculus) 다음이 성립한다.

i. 적분가능한 연속함수 $f : \mathbb{R} \to \mathbb{R}^n$에 대해 함수 $F : \mathbb{R} \to \mathbb{R}^n$를 $F : x \mapsto \int_{-\infty}^x f\, d\lambda_1$로 두면 이는 f의 역도함수이다. 즉, $\nabla F = f$이다.
ii. 함수 $F : \mathbb{R} \to \mathbb{R}^n$가 $\mathscr{C}^1$급이고 그 도함수가 적분가능하면 $f := \nabla F$에 대해 $F(x) = \int_{-\infty}^x f\, d\lambda_1$이다.

Proof 이는 정의와 FTC로부터 자명하다. □

복소함수의 적분도 이와 비슷한 아이디어로 정의할 수 있다. 다만, 우선 $\mathbb{C}$ 위의 '바람직한' σ−대수를 어떻게 잡을지에 대한 질문의 답이 필요하다. 앞서 $\mathbb{R}^n$ 위의 '바람직한' σ−대수인 $\mathscr{B}_n$이나 $\mathscr{M}_n$을 구성하기 위해 얼마나 많은 지면을 할애하여 논의를 전개했는지를 생각해보면 정신이 아득해질 법하지만, 조금만 요령을 부리면 이 질문에 꽤 근사한 답을 할 수 있다. 바로 $\mathbb{C}$와 $\mathbb{R}^2$를 동일시하는 것이다. 즉, 복소수 $x + \mathbf{i}y$를 (x, y)로 생각하고, 이를 통해 우리는 $\mathbb{R}^2$에서의 표준위상, 거리함수, 심지어는 $\mathscr{B}_2$와 $\mathscr{M}_2$로 만들어지는 가측공간의 구조도 $\mathbb{C}$로 자연스럽게 옮겨올 수 있다. 곧, 우리는 가측공간 $(X, \mathscr{A})$에서 정의된 복소함수 $f : X \to \mathbb{C}$에 대해 이의 실수부와 허수부가 모두 가측이면 이때의 f를 가측함수

라 한다. 비슷하게, $f:\mathbb{R}\to\mathbb{C}$의 실수부와 허수부가 모두 연속이거나 미분가능하면 이때의 f를 각각 연속 혹은 미분가능하다고 한다. (보통 표기의 편의를 위해 $f:\mathbb{R}\to\mathbb{C}$의 미분을 $f' = (\mathbf{Re}\, f)' + (\mathbf{Im}\, f)'$로 쓴다.)

Definition 1.244 측도공간 $(X,\mathscr{A},\mu)$에서 정의된 가측함수 $f:X\to\mathbb{C}$에 대해 f의 실수부와 허수부가 적분가능하면 이때 f를 **(μ−Lebesgue) 적분가능(− integrable)**하다고 한다. 나아가 적분가능한 $f:X\to\mathbb{C}$에 대해 μ에 대한 f의 **(Lebesgue) 적분(− integral)**을 $\int_X f(x)\,d\mu(x)$ 혹은 간단히 $\int_X f\,d\mu$로 쓰고 $\int_X f\,d\mu = \int_X \mathbf{Re}\, f\,d\mu + \mathbf{i}\int_X \mathbf{Im}\, f\,d\mu$로 정의한다.

적분의 기본적인 성질이 비슷하게 성립하는 것은 벡터함수의 적분과 마찬가지이다.

Corollary 1.245 측도공간 $(X,\mathscr{A},\mu)$에서 정의된 적분가능한 함수 $f,g:X\to\mathbb{R}^n$에 대해 다음이 성립한다.

i. 임의의 $c\in\mathbb{R}$에 대해 cf가 적분가능하고 $\int_X cf\,d\mu = c\int_X f\,d\mu$이다.
ii. 함수 $f+g$가 적분가능하고 $\int_X f+g\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$이다.
iii. 함수 $\overline{f}$가 적분가능하고 $\int_X \overline{f}\,d\mu = \overline{\int_X f\,d\mu}$이다.
iv. $|\int_X f\,d\mu| \le \int_X |f|\,d\mu$.

Proof 이는 정의와 적분의 기본적인 성질로부터 자명하다. □

Corollary 1.246 (Fundamental theorem of calculus) 다음이 성립한다.

i. 적분가능한 연속함수 $f:\mathbb{R}\to\mathbb{C}$에 대해 함수 $F:\mathbb{R}\to\mathbb{C}$를 $F:x\mapsto\int_{-\infty}^{x} f\,d\lambda_1$로 두면 이는 f의 역도함수이다. 즉, $F'=f$이다.
ii. 함수 $F:\mathbb{R}\to\mathbb{C}$가 $\mathscr{C}^1$급이고 그 도함수가 적분가능하면 $f:=F'$에 대해 $F(x)=\int_{-\infty}^{x} f\,d\lambda_1$이다.

Proof 이는 정의와 FTC로부터 자명하다. □

특히, 복소함수의 적분의 경우 Euler의 공식으로 알려진 $e^{\mathbf{i}\theta}=\cos\theta+\mathbf{i}\sin\theta$로부터 $e^{\mathbf{i}\theta}$의 θ에 대한 미분이 $\mathbf{i}e^{\mathbf{i}\theta}$가 되어 그 계산이 간단해지는 경우가 많다. 예컨대 $\int_0^{\sqrt{2\pi}}\theta e^{\mathbf{i}\theta^2}\,d\lambda_1$를 계산하는 경우에, 이를 $\int_0^{\sqrt{2\pi}}\theta(\cos\theta^2+\mathbf{i}\sin\theta^2)\,d\lambda_1$로 생각하여 계산해도 되지만, $(e^{\mathbf{i}\theta^2}/2\mathbf{i})' = \theta e^{\mathbf{i}\theta^2}$임에 착안하여 $\int_0^{\sqrt{2\pi}}\theta e^{\mathbf{i}\theta^2}\,d\lambda_1 = [e^{\mathbf{i}\theta^2}/2\mathbf{i}]_0^{\sqrt{2\pi}} = 0$로 손쉽게 계산할 수도 있다.

이로써 길고 험했던 측도론 이야기는 막을 내린다. 무려 250개 가까이 되는 생소한 정의, 정리와 씨름하며 여기까지 와 준 독자들에게 정말 감사할 따름이다. 통계학 이야기는 이제서야 본격적으로 시작되지만, 늦었다고 걱정할 필요는 없다. 측도론을 통해 보는 통계학은 분명 여태까지와는 다른 새로운 통계학일 것이다.

Notes

1 [측도론] Lebesgue 측도로서 잴 수 없는 집합을 구성하는 과정을 간단히 실어두도록 한다. 우선 $\mathbb{R}$ 위의 관계 $\sim$을 $x \sim y :\Leftrightarrow x - y \in \mathbb{Q}$로 정의하면 이가 동치관계임이 거의 자명하다. 이제 함수 $f : \mathbb{R}/\sim \to [0,1]$를 각 동치류에서 $[0,1]$에 속하는 representative를 하나 선택하는 선택함수라 하면 이는 well-define된다. (여기서 선택공리가 필요하다.) 그렇다면 집합 $V = f(\mathbb{R}/\sim)$가 바로 우리가 찾는 Lebesgue 측도로서 잴 수 없는 집합으로, 이를 흔히 **Vitali 집합(– set)**이라 한다.

집합 V가 정말 비가측인지를 확인하기 위해 이가 가측이라 가정하고 임의의 $p \in \mathbb{Q}$에 대해 $V_p = V + p$라 하자. 그렇다면 Lebesgue 측도의 이동 불변성으로부터 임의의 $p \in \mathbb{Q}$에 대해 $\lambda_1(V) = \lambda_1(V_p)$이고 V의 구성으로부터 서로다른 $p, q \in \mathbb{Q}$에 대해 V_p, V_q는 서로소이다. (그렇지 않다면 어떤 $x \in \mathbb{R}$에 대해 $x - p, x - q \in V$인데, 이는 $x - p \sim x - q$에서 모순이다.) 이제 $(-1,1]$에 속하는 모든 유리수를 나열하여 $\{p_i\}$와 같이 수열의 형태로 나타내면 $(0,1] \subseteq \bigsqcup_{i=1}^{\infty} V_{p_i} \subseteq (-1,2]$이므로(Hint: 임의의 $x \in (0,1]$에 대해 적당한 $y \in V$가 존재하여 $p := x - y \in \mathbb{Q}$이고 $p \in (-1,1]$이다.) 이상의 결과로부터 $1 \le \lambda_1(\bigsqcup_{i=1}^{\infty} V_{p_i}) = \sum_{i=1}^{\infty} \lambda_1(V_{p_i}) = \sum_{i=1}^{\infty} \lambda_1(V) \le 3$의 모순이 발생하고, 따라서 V는 비가측이다.

조금 더 충격적이지만 재미있는 사실은 양의 측도를 갖는 모든 $\mathbb{R}$의 가측집합은 Vitali 집합과 같은 비가측 집합을 부분집합으로 가진다는 사실이다. **Vitali의 정리**라 불리는 이 정리의 구성적인 증명에는 몇가지 다른 정리의 도움이 필요하다. 먼저 compact한 $K \subseteq \mathbb{R}$와 집합 $D = \{x - y \in \mathbb{R} : x, y \in K\}$에 대해 $\lambda_1(K) > 0$이면 적당한 $r > 0$이 존재하여 $B(r) \subseteq D$임을 보이자. 우선 Lebesgue 측도가 regular하므로 적당한 열린집합 $U \subseteq \mathbb{R}$가 존재하여 $K \subseteq U$이고 $\lambda_1(U) < 2\lambda_1(K)$이다. 이제 $r = d(K, U^c) > 0$라 하고 $|x| < r$이면 $K + x \subseteq U$이다. (그렇지 않다면 $x + y \notin U$인 $y \in K$가 존재하는데, $r \le |x + y - y| = |x| < r$에서 모순이 발생한다.) 여기서 만약 $(K + x) \cap K = \emptyset$이라면 $(K + x) \sqcup K \subseteq U$와 Lebesgue 측도의 이동 불변성으로부터 $\lambda_1(U) \ge \lambda_1(K + x) + \lambda_1(K) = 2\lambda_1(K)$가 되고, 곧 가정에 모순되므로 $(K + x) \cap K \ne \emptyset$이고, 따라서 적당한 $y \in K$가 존재하여 $y - x \in K$에서 $x = y - (y - x) \in D$이다. 그리고 이로부터 $B(r) \subseteq D$이다.

다음으로, 임의의 가측집합 $A \subseteq \mathbb{R}$와 집합 $D = \{x - y \in \mathbb{R} : x, y \in A\}$에 대해 $\lambda_1(A) > 0$이면 적당한 $r > 0$이 존재하여 $B(r) \subseteq D$임을 보이자. 집합열 $\{A \cap B(i)\}$를 생각하면 이는 A로 수렴하는 증가하는 집합열이므로 $\lambda_1(A \cap B(i)) \uparrow \lambda_1(A) > 0$이고, 따라서 충분히 큰 $i_0 \in \mathbb{N}$에 대해 $\lambda_1(A \cap B(i_0)) > 0$이다. 또한, Lebesgue 측도가 regular하므로 적당한 닫힌집합 $K \subseteq A \cap B(i_0) \subseteq A$에 대해 $\lambda_1(K) > \lambda_1(A \cap B(i_0))/2$이다. 그렇다면 여기서의 K는 유계여서 compact하고, 곧 앞선 결과로부터 적당한 $r > 0$이 존재하여 $B(r) \subseteq \{x - y \in \mathbb{R} : x, y \in K\} \subseteq D$이다.

이제 임의의 가측집합 $A \subseteq \mathbb{R}$에 대해 $\lambda_1(A) > 0$이라 하고, 앞서 구성한 Vitali 집합을 $V \subseteq [0,1]$라 하자. 또한, 임의의 $p \in \mathbb{Q}$에 대해 $V_p = V + p$라 하면 서로다른 $p, q \in \mathbb{Q}$에 대해서는 V_p와 V_q가 서로소임을 이미 알고 있고, $\bigsqcup_{p \in \mathbb{Q}} V_p = \mathbb{R}$임을 쉽게 보일 수 있다. (Hint: 임의의 $x \in \mathbb{R}$에 대해 적당한 $y \in V$가 존재하여 $p := x - y \in \mathbb{Q}$이다.) 이제 모순을 유도하기 위해 모든 $p \in \mathbb{Q}$에 대해 $A \cap V_p$가 가측이라 하고 모든 유리수를 나열하여 $\{p_i\}$와 같이 수열의 형태로 나타내면 $A = \bigsqcup_{i=1}^{\infty}(A \cap V_{p_i})$에서 $0 < \lambda_1(A) = \sum_{i=1}^{\infty} \lambda_1(A \cap V_{p_i})$이다. 그런데 이는 어떤 $p \in \mathbb{Q}$에 대해 $\lambda_1(A \cap V_p) > 0$임을 함의하므로 앞선 결과로부터 적당한 $r > 0$이 존재하여 $B(r) \subseteq \{x - y \in \mathbb{R} : x, y \in A \cap V_p\}$의 모순이 발생한다. 따라서 어떤 $p \in \mathbb{Q}$에 대해 $A \cap V_p \subseteq A$는 비가측이다.

2 λ–system은 다음 세 가지 조건을 만족하는 집합족 $\mathscr{L}$로써 동등하게 정의할 수 있다.

i. $X \in \mathscr{L}$.
ii. (monotone difference) 임의의 $A, B \in \mathscr{L}$에 대해 $A \subseteq B$이면 $B \setminus A \in \mathscr{L}$이다.

iii. 집합족 $\mathscr{L}$에 속하는 임의의 증가하는 집합열 $\{A_i\}$에 대해 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{L}$이다.

이와 같은 두 정의가 서로 동등함을 보이기 위해서, 먼저 정의 ? ? 의 조건을 만족하는 집합족 $\mathscr{L}$을 생각하자. 그렇다면 정의 ? ? 의 i과 ii로부터 $\emptyset \in \mathscr{L}$임을 알고, 이를 정의 ? ? 의 iii에 적용하면 $\mathscr{L}$이 유한번의 서로소 합집합에 대해서도 닫혀있음을 쉽게 알 수 있다. 이제 임의의 $A, B \in \mathscr{L}$에 대해 $A \subseteq B$라 하면 $B \setminus A = B \cap A^c = (B^c \sqcup A)^c \in \mathscr{L}$이 되어 위의 조건 ii가 만족된다. 비슷하게, 집합족 $\mathscr{L}$에 속하는 임의의 증가하는 집합열 $\{A_i\}$에 대해 집합열 $\{B_i\}$를 $B_i := A_i \setminus A_{i-1}$ $(B_1 := A_1)$로 정의하면 $\{B_i\}$가 $\mathscr{L}$에 속하는 서로소인 집합열이 되어 $\bigcup_{i=1}^{\infty} A_i = \bigsqcup_{i=1}^{\infty} B_i \in \mathscr{L}$에서 위의 조건 iii도 만족된다.
역으로, 위의 조건을 만족하는 집합족 $\mathscr{L}$을 생각하자. 그렇다면 임의의 $A \in \mathscr{L}$에 대해 $A^c = X \setminus A$에서 정의 ? ? 의 ii가 만족된다. 이제 임의의 서로소인 $A, B \in \mathscr{L}$에 대해 $B \subseteq A^c \in \mathscr{L}$이므로 $A \sqcup B = (A^c \cap B^c)^c = (A^c \setminus B)^c \in \mathscr{L}$이다. 여기에 수학적 귀납법을 적용하면 $\mathscr{L}$이 유한번의 서로소 합집합에 대해 닫혀있음을 알 수 있다. 이를 이용하여 마지막 조건인 정의 ? ? 의 iii을 보이기 위해 집합족 $\mathscr{L}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$를 생각하고 집합열 $\{B_i\}$를 $B_i = \bigsqcup_{j=1}^{i} A_j$로 정의하자. 그렇다면 $\{B_i\}$가 $\mathscr{L}$에 속하는 증가하는 집합열이 되어 $\bigsqcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i \in \mathscr{L}$에서 증명이 끝난다.
본문에서는 λ–system의 정의로 정의 ? ? 의 조건과 위의 세 조건을 특별한 언급 없이 모두 사용하였다.

3 **[측도론]** 임의의 $p, q \in \mathbb{Q}$에 대해 $[p, q) \cap \mathbb{Q}$를 rational semi-open interval이라 하자. 이제 $\mathscr{A}$를 $[0, 1)$에 속하는 모든 rational semi-open interval의 모임이라 하면 이는 $\mathbb{R}$ 위의 semi-algebra이고 셈측도 #는 $A(\mathscr{A})$ 위에서 premeasure가 된다. 또한, 그 구성으로부터 임의의 $A \in A(\mathscr{A})$에 대해 $\#(A) = 0$이거나 $\#(A) = \infty$임이 분명하다. (Hint: 대수 $A(\mathscr{A})$는 $\mathscr{A}$에 속하는 서로소인 집합들의 유한 합집합으로 만들 수 있는 모든 집합들의 모임이다.) 그렇다면 모든 #–가측집합의 모임 $\mathscr{M}$에 대해 #와 2# 모두 $A(\mathscr{A})$ 위의 premeasure #을 $\mathscr{M}$ 위의 측도로 확장한 것이지만 이 둘은 서로 다르다. (한원소 집합이 $\sigma(\mathscr{A})$에 속하므로 $\mathscr{M}$에서 $\# = 2\#$일 수 없다.)

4 **[집합론]** 대표적인 σ–대수 $\mathscr{M}_n$과 $\mathscr{B}_n$의 기수는 각각 $2^{\mathfrak{c}}, \beth_1 (= 2^{\aleph_0})$인데, 생각하는 것과 달리 $\mathscr{M}_n$의 기수를 구하는 것이 $\mathscr{B}_n$의 기수를 구하는 것보다 더 쉽다. Cantor 집합 C를 생각하면 이는 가측이고 영집합이므로 C의 모든 부분집합도 가측인데, $|C| = \mathfrak{c}$이므로 $\mathscr{P}(C) \subseteq \mathscr{M}_n \subseteq \mathscr{P}(\mathbb{R})$에서 $|\mathscr{M}_n| = 2^{\mathfrak{c}}$임을 바로 알 수 있다. 한편, $\mathscr{B}_n$의 기수를 구하기 위해서는 준비가 조금 필요하다.
먼저 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 이가 가산이면 $|\sigma(\mathscr{C})| < \aleph_0$이거나 $|\sigma(\mathscr{C})| \geq \beth_1$임을 보이자. 편의상 $\mathscr{C}$가 무한집합이라 생각하고 이의 원소를 나열하여 $\{A_i\}$와 같이 집합열의 형태로 쓰자. 또한, 임의의 $I \subseteq \mathbb{N}$에 대해 집합 $B_I = \bigcap_{i \in I} A_i \cap \bigcap_{i \notin I} A_i^c$를 생각하면 서로다른 $I, J \subseteq \mathbb{N}$에 대해 B_I와 B_J는 서로소이고 $\bigsqcup_{I \subseteq \mathbb{N}} B_I = X$이다. (Hint: 집합족 $\mathscr{C}$가 그려내는 Venn diagram을 생각해보면 거의 자명하다.) 이는 곧 $\mathscr{D} := \{B_I \subseteq X : I \subseteq \mathbb{N}, B_I \neq \emptyset\}$가 X의 분할임을 뜻한다. 이제 경우를 나누어 $\mathscr{D}$가 유한집합인 경우를 보자. 이 경우에는 $\mathscr{D} = \{B_{I_1}, \cdots, B_{I_k}\}$로 쓸 수 있고 $\sigma(\mathscr{C}) = \sigma(\mathscr{D}) = \{\bigsqcup_{j \in J} B_{I_j} : J \subseteq \{1, \cdots, k\}\}$이므로 임을 쉽게 보일 수 있으므로 $|\sigma(\mathscr{C})| < \aleph_0$이다. (Hint: 임의의 $A_i \in \mathscr{C}$에 대해 A_i와 교차하는 $\mathscr{D}$의 원소를 골라 이들을 서로소 합집합하면 A_i이다. 따라서 $\mathscr{C} \subseteq \sigma(\mathscr{D})$이고, 이제 나머지는 자명하다.) 반대로 $\mathscr{D}$가 무한집합인 경우에는 $\mathscr{D} \subseteq \sigma(\mathscr{C})$에서 $\{\bigsqcup_{j \in J} B_{I_j}\}_{J \subseteq \mathbb{N}} \subseteq \sigma(\mathscr{C})$이고 서로다른 $J, J' \subseteq \mathbb{N}$에 대해 $\bigsqcup_{j \in J} B_{I_j}$와 $\bigsqcup_{j \in J'} B_{I_j}$가 서로소이므로 $|\sigma(\mathscr{C})| \geq \beth_1$이다.
다음으로, $\mathscr{B}_n$의 가산 생성자 $\mathscr{C}$를 하나 적당히 구한다. (예컨대 $\mathbb{R}^n$의 모든 열린집합의 모임이 $\mathscr{B}_n$을 생성하므로 꼭짓점이 $\mathbb{Q}^n$인 열린 box의 모임을 생각하면 이는 $\mathscr{B}_n$의 가산 생성자이다.) 그렇다면 앞선 결과로부터 $|\mathscr{B}_n| \leq \beth_1$만 보이면 되는데, 이를 위해서는 초한귀납법을 써야 한다. 우선 $\mathscr{C}_0 = \mathscr{C}$로 시작하여 임의의 순서수 α에 대해 $\mathscr{C}_{\alpha+1} = \{\bigcup_{i \in I} A_i \cup \bigcup_{j \in J} B_j^c \subseteq \mathbb{R}^n : I, J \subseteq \mathbb{N}, A_i, B_j \in \mathscr{C}_\alpha\}$로 두고 극한 순서수 α에 대해서는 $\mathscr{C}_\alpha = \bigcup_{\beta < \alpha} \mathscr{C}_\beta$로 둔다. 그리고는 $\mathscr{A} = \bigcup_{\alpha < \omega_1} \mathscr{C}_\alpha$라 하여 $\mathscr{A}$가 σ–대수임을 주장한다. 이를 위해서는 이런저런 조건을 확인해봐야 하는데, 가산 합집합에 대해 닫혀있다는 조건

만 확인하면 나머지는 자명하다. 이에 $\mathscr{A}$에 속하는 집합열 $\{A_i\}$를 택하면 $\{\mathscr{C}_\alpha\}$가 증가하는 초한 집합열이므로 $A_i \in \mathscr{C}_{\alpha_i}$인 증가하는 수열 $\{\alpha_i\}$를 구성할 수 있다. 그런데 $\mathscr{A}$의 구성으로부터 각 $i \in \mathbb{N}$에 대해 $\alpha_i < \omega_1$이므로 $\sup_{i\in\mathbb{N}} \alpha_i < \omega_1$에서 $\bigcup_{i=1}^\infty A_i \in \bigcup_{i=1}^\infty \mathscr{C}_{\alpha_i} \subseteq \bigcup_{\alpha<\omega_1} \mathscr{C}_\alpha = \mathscr{A}$가 되어 $\mathscr{A}$가 σ−대수임을 알고, 곧 $\mathscr{B}_n \subseteq \mathscr{A}$에서 $|\mathscr{B}_n| \le |\mathscr{A}| \le \sum_{\alpha<\omega_1} |\mathscr{C}_\alpha| \le \aleph_0 \beth_1 = \beth_1$이다. (이는 임의의 $\alpha < \omega_1$에 대해 $|\mathscr{C}_\alpha| \le \beth_1$이기 때문에 성립하고, 다시 이는 초한귀납법으로 쉽게 보일 수 있다. DIY!)

5 집합 $\mathbb{R}^n$에서 모든 꼭짓점이 유리점인 열린 box를 rational box라 하면, $\mathbb{Q}$가 가산이므로 $\mathbb{Q}^n$도 가산이고, 곧 모든 rational box도 가산개이다. 이제 임의의 열린집합 $U \subseteq \mathbb{R}^n$에 대해 U에 포함되는 rational box만을 모두 택하여 집합족 $\mathscr{U}$를 구성하자. 그렇다면 $U = \bigcup \mathscr{U}$임을 쉽게 보일 수 있고, $\mathscr{U}$가 명백히 가산이므로 $\mathbb{R}^n$의 모든 열린집합은 유계인 열린 box의 가산 합집합으로 표현됨을 안다.

6 [측도론] 각주 ? ? 에서 살펴본 바와 같이 $|\mathscr{B}_n| < |\mathscr{M}_n|$이므로 $\mathscr{B}_n$은 $\mathscr{M}_n$의 진부분집합이다. 이는 곧 가측이지만 Borel은 아닌 집합이 존재한다는 뜻인데, 이를 직접 구성해볼 수 있다. 이번에는 Cantor 함수 $c : [0,1] \to \mathbb{R}$가 중요한 역할을 한다. 우선 함수 $f : [0,1] \to [0,2]$를 $f : x \mapsto x + c(x)$로 정의하면 이는 증가하는 연속함수이므로 Borel이고, 전단사임을 쉽게 보일 수 있다. (Hint: 함수 f가 전사임은 IVT로부터 자명하다. 한편, f가 단사임을 보이기 위해서 $x < y$인 임의의 $x, y \in [0,1]$에 대해 $f(x) = f(y)$라 하면 $c(x) > c(y)$에서 모순이 발생한다.) 또한, Cantor 집합 C에 대해 $\lambda_1(f(C)) = 1$임도 쉽게 보일 수 있으므로(Hint: 닫힌구간 $[0,1]$에 속하는 서로소인 열린구간 $I_1, I_2, \cdots$에 대해 $[0,1] \setminus C = \bigsqcup_{i=1}^\infty I_i$로 쓸 수 있고, 각 I_i에서 c는 상수함수이므로 $\lambda_1(f(I_i)) = \lambda_1(I_i)$이다.) Vitali의 정리로부터 적당한 비가측 집합 $A \subseteq f(C)$가 존재한다. 그렇다면 $f^{-1}(A) \subseteq C$는 영집합이어서 가측이지만 Borel이 아니다.

7 구체적으로 다음과 같이 할 수 있다. 먼저 $\delta = \lambda_n(A) - \sum_{i=1}^\infty \rho_n(A_i) + \varepsilon > 0$로 두고 각 $i \in \mathbb{N}$에 대해 semi-open box A_i의 중심을 고정한 채로 이의 모서리의 길이를 조금씩 늘리되 그 결과로 만들어진 새 semi-open box A_i'이 $\rho_n(A_i') < \rho_n(A_i) + \delta/2^{i+1}$를 만족하도록 한다. 그렇다면 $A_i \subseteq (A_i')^\circ \subseteq A_i'$에서 $A \subseteq \bigcup_{i=1}^\infty A_i \subseteq \bigcup_{i=1}^\infty (A_i')^\circ$이고 $\sum_{i=1}^\infty \rho_n(A_i') \le \sum_{i=1}^\infty [\rho_n(A_i) + \delta/2^{i+1}] = \sum_{i=1}^\infty \rho_n(A_i) + \delta/2 < \lambda_n(A) + \varepsilon$이 성립한다.

8 [위상수학] $\mathbb{R}$에 주어진 표준위상 $\mathscr{T}$는 $\beta = \{(x,\infty)\}_{x\in\mathbb{R}} \cup \{(-\infty, y)\}_{y\in\mathbb{R}}$를 부분기저로 하는 순서위상이다. 따라서 이를 그대로 확장하여 $\overline{\mathbb{R}}$에도 $\widetilde{\beta} = \{(x,\infty]\}_{x\in\mathbb{R}} \cup \{[-\infty, y)\}_{y\in\mathbb{R}}$를 부분기저로 하는 순서위상 $\mathscr{T}_{\text{ext}}$를 줄 수 있고, 이는 곧 $\gamma = \{(x,\infty]\}_{x\in\mathbb{R}} \cup \{[-\infty, y)\}_{y\in\mathbb{R}} \cup \{(x,y)\}_{x,y\in\mathbb{R}}$이 $\mathscr{T}_{\text{ext}}$의 기저라는 것과 동치이므로 $\mathscr{T} \subseteq \mathscr{T}_{\text{ext}}$임이 자명하다. 이제 임의의 $U \in \mathscr{T}_{\text{ext}}$는 적당한 $x, y \in \mathbb{R}$와 $V \subseteq (y,x)$인 $V \in \mathscr{T}$에 대해 $[-\infty, y) \cup V \cup (x, \infty]$로 쓸 수 있고, V는 유계인 열린구간의 가산 합집합으로 쓸 수 있으므로 곧 U는 적당한 $x, y \in \mathbb{R}$에 대해 (x,y), $(x,\infty]$, $[-\infty, y)$로 표현되는 구간들의 가산 합집합으로 표현되고 그 역도 성립한다.

한편, $\overline{\mathbb{R}}$에서의 사칙연산에서는 $\infty - \infty$, ∞/∞, $0 \cdot \infty$와 같은 부정형이 나올 수 있으므로 각별히 주의해야 한다. 다만, 이 중의 일부의 값을 관례를 따라 $0 \cdot \infty := 0$, $0^0 := 1$로 둔다.

9 [위상수학] 일반적으로 위상공간 $(X, \mathscr{T})$에 대해 $\mathscr{T}$가 생성하는 σ−대수를 X 위의 Borel σ−대수로 정의한다. 이 책에서는 위상수학의 내용을 피하기 위해 $\mathbb{R}^n$에 한정하여 Borel σ−대수를 도입하였다.

10 [측도론] 함수 $f : [0,1] \to [0,2]$를 각주 6에서와 같이 두면 이가 전단사인 연속함수로 Cantor 집합 C에 대해 $\lambda_1(f(C)) = 1$임을 이미 알고 있다. 따라서 Vitali의 정리로부터 비가측 집합 $A \subseteq f(C)$를 적어도 하나 택할 수 있는 한편, 집합 $B = f^{-1}(A) \subseteq C$는 C가 영집합이라는 점에서 가측이다. 이제 함수 $g : [0,2] \to \mathbb{R}$를 $g = \mathbf{1}_B \circ f^{-1}$로 두면 f^{-1}가 연속이고 B가 가측이므로 f^{-1}와 $\mathbf{1}_B$ 각각은 가측이지만 $g^{-1}(1) = f(B) = A$가 가측이 아니므로 g는 가측이 아니다.

11 만약 $x \in X$에 대해 $f(x) + g(x)$가 $\infty - \infty$의 부정형이 되면, 이때의 값을 0과 같은 dummy value로 정하는 관례를 전제하였다. 가정으로부터 f와 g가 적분가능하여 이들은 거의 어디서나 유한하고, 따라서 이러한 점의 집합이 영집합을 이루므로 적분의 결과는 이때의 dummy value의 선택과는 무관하다.

12 수렴정리의 조건이 만족되지 않아 극한과 적분이 교환되지 않는 예시를 보자. 단순함수 $f_i : \mathbb{R} \to \mathbb{R}$의 열 $\{f_i\}$를 $f_i := \mathbf{1}_{(i^2-i,i^2)}/2i - \mathbf{1}_{(i^2,i^2+i)}/i$로 두면 이는 $f := 0$로 수렴함이 자명하다. (심지어 균등수렴한다.) 하지만 $\{f_i\}$가 양과 음의 값을 모두 가져 MCT를 쓸 수 없다. 실제로 임의의 $i \in \mathbb{N}$에 대해 $\int_{\mathbb{R}} f_i \, d\lambda_1 = -1/2$이므로 $\lim_{i\to\infty} \int_{\mathbb{R}} f_i \, d\lambda_1 = -1/2 \neq 0 = \int_{\mathbb{R}} f \, d\lambda_1$이다. 같은 이유에서 Fatou의 보조정리도 쓸 수 없다.

그렇다면 $\{f_i\}$ 대신 $\{|f_i|\}$를 생각해보면 어떨까? 이 경우에 $\{|f_i|\}$는 명백히 음이 아닌 함수열이지만 이가 증가하지 않아 MCT는 여전히 쓸 수 없다. 실제로 임의의 $i \in \mathbb{N}$에 대해 $\int_{\mathbb{R}} |f_i| \, d\lambda_1 = 3/2$이므로 $\lim_{i\to\infty} \int_{\mathbb{R}} f_i \, d\lambda_1 = 3/2 \neq 0 = \int_{\mathbb{R}} f \, d\lambda_1$이다.

한편, $\{f_i\}$에 DCT를 쓸 수는 없을까? DCT를 쓰려면 $\{f_i\}$를 지배하는, 즉 각 f_i에 대해 $|f_i| \le g$인 적분가능한 $g : \mathbb{R} \to \overline{\mathbb{R}}_0^+$가 필요하다. 이러한 g가 존재한다고 하면 서로다른 $i, j \in \mathbb{N}$에 대해 (i^2-i, i^2+i)와 (j^2-j, j^2+j)가 서로소이므로(WLOG, 만약 $i < j$라 하면 $j^2 - j - i^2 - i = (j-i)(j+i) - (j+i) > (j+i) - (j+i) = 0$이다.) 모든 $x \in \mathbb{R}$에 대해 $\sum_{i=1}^{\infty} |f_i(x)|$는 well-define되고 $\sum_{i=1}^{\infty} |f_i| \le g$이다. 그러나 이는 MCT에서 $\int_{\mathbb{R}} \sum_{i=1}^{\infty} |f_i| \, d\lambda_1 = \sum_{i=1}^{\infty} \int_{\mathbb{R}} |f_i| \, d\lambda_1 = \infty \le \int_{\mathbb{R}} g \, d\lambda_1$의 모순을 일으키므로 어떠한 적분가능한 함수도 $\{f_i\}$를 지배할 수 없고, 곧 DCT도 쓸 수 없다.

13 만약 $x \in X$에 대해 $\sum_{i=1}^{\infty} f_i(x)$가 진동하면, 이때의 값을 0과 같은 dummy value로 정하는 관례를 전제하였다. 가정으로부터 이러한 점의 집합이 영집합을 이루므로 적분의 결과는 이때의 dummy value의 선택과는 무관하다.

14 만약 $x \in X$에 대해 함수 $y \mapsto f(x,y)$가 미분가능하지 않다면, 이때의 $(\partial f/\partial y)(x,y)$의 값을 0과 같은 dummy value로 정하는 관례를 전제하였다. 가정으로부터 이러한 점의 집합이 영집합을 이루므로 따름정리의 결론은 이때의 dummy value의 선택과는 무관하다.

15 측도공간 $(X, \mathscr{A}, \mu)$와 $(Y, \mathscr{B}, \nu)$에 대해 $\mathscr{A} \times \mathscr{B}$는 semi-algebra이다. 이제 함수 $\rho : \mathscr{A} \times \mathscr{B} \to \overline{\mathbb{R}}_0^+$를 $\rho : A \times B \mapsto \mu(A)\nu(B)$로 두고 이가 $\mathscr{A} \times \mathscr{B}$ 위의 premeasure임을 보이면 정리 ? ? 와 Carathéodory의 확장정리로부터 ρ를 $\mathscr{A} \otimes \mathscr{B}$ 위의 측도로 확장시킬 수 있으므로 존재성이 보여진다. 이를 위해 $\mathscr{A} \times \mathscr{B}$에 속하는 임의의 서로소인 집합열 $\{A_i \times B_i\}$를 생각하여 $\bigsqcup_{i=1}^{\infty} (A_i \times B_i) \in \mathscr{A} \times \mathscr{B}$라 하자. 그렇다면 적당한 $A \in \mathscr{A}, B \in \mathscr{B}$에 대해 $A \times B = \bigsqcup_{i=1}^{\infty} (A_i \times B_i)$이고, $\mathbf{1}_A(x)\mathbf{1}_B(y) = \sum_{i=1}^{\infty} \mathbf{1}_{A_i}(x)\mathbf{1}_{B_i}(y)$에서 MCT로부터 $\mathbf{1}_A(x)\nu(B) = \int_Y \mathbf{1}_A(x)\mathbf{1}_B(y) \, d\nu(y) = \sum_{i=1}^{\infty} \int_Y \mathbf{1}_{A_i}(x)\mathbf{1}_{B_i}(y) \, d\nu(y) = \sum_{i=1}^{\infty} \mathbf{1}_{A_i}(x)\nu(B_i)$이고, 다시

$$\begin{aligned}
\rho(A \times B) &= \mu(A)\nu(B) \\
&= \int_X \mathbf{1}_A(x)\nu(B) \, d\mu(x) \\
&= \sum_{i=1}^{\infty} \int_X \mathbf{1}_{A_i}(x)\nu(B_i) \, d\mu(x) \\
&= \sum_{i=1}^{\infty} \mu(A_i)\nu(B_i) \\
&= \sum_{i=1}^{\infty} \rho(A_i \times B_i)
\end{aligned}$$

이다. 한편, $\rho(\emptyset) = 0$임은 자명하므로 ρ는 premeasure이다.

16 임의의 열린집합 $U \subseteq \mathbb{R}^n$를 택하여 $\mathbb{R}^n$의 rational box 중에서 U에 포함되는 것만을 뽑아 집합족 $\mathscr{U}$를 구성하면 이는 명백히 가산이고, 곧 $\{U_i\}$와 같이 집합열의 형태로 쓸 수 있다. (Rational box의 의미에 대해서는 각주 ? ? 를 참조하기 바란다.) 그렇다면, 각 $i \in \mathbb{N}$에 대해 적당한 $x_i, y_i \in \mathbb{Q}^n$가 존재하여 $U_i = (x_i, y_i)$이고, 여기서 $V_i = \prod_{j=1}^{l} (x_i^j, y_i^j), W_i = \prod_{j=l+1}^{m} (x_i^j, y_i^j)$라 하면 $\{V_i\}, \{W_i\}$는 각각 $\mathbb{R}^l, \mathbb{R}^m$에 속하는 열린 집합열로서 $U = \bigcup_{i=1}^{\infty} U_i = \bigcup_{i=1}^{\infty} (V_i \times W_i)$이다.

17 이는 각주 ? ? 와 거의 비슷하게 보일 수 있다. 집합 $\mathbb{R}^n$에서 모든 꼭짓점이 유리점인 semi-open box를 rational semi-box라 하면, $\mathbb{Q}$가 가산이므로 $\mathbb{Q}^n$도 가산이고, 곧 모든 rational semi-box의 모임 $\mathscr{S}$도 가산이다. 여기서 임의의 $B, B' \in \mathscr{S}$에 대해 $B \setminus B'$가 적당한 서로소인 $B_1, \cdots, B_k \in \mathscr{S}$에 대해 $B \setminus B' = \bigsqcup_{i=1}^k B_i$로 표현된다는 점을 주목하자. 이제 임의의 열린집합 $U \subseteq \mathbb{R}^n$에 대해 U에 포함되는 rational semi-box만을 모두 택하여 집합족을 구성하면 이가 명백히 가산이므로 WLOG, 필요하다면 공집합을 추가하여 이를 집합열 $\{B_i\}$로 쓸 수 있다. 한편, 각 $i \in \mathbb{N}$에 대해 $B_i' = B_i \setminus \bigcup_{j=1}^{i-1} B_j$를 생각하면 이는 적당한 서로소인 rational semi-box $B_{i1}, \cdots, B_{ik_i}$에 대해 $B_i' = \bigsqcup_{j=1}^{k_i} B_{ij}$이므로 rational semi-box의 열 $\{B_{ij}\}_{ij}$는 $\mathscr{S}$에 속하는 서로소인 집합열이고 $U = \bigcup_{i=1}^\infty B_i = \bigsqcup_{i=1}^\infty B_i' = \bigsqcup_{i=1}^\infty \bigsqcup_{j=1}^{k_i} B_{ij}$ 임을 쉽게 보일 수 있다.

18 부등식 $\varphi_{\mathscr{P}_i} \le f\mathbf{1}_A \le \psi_{\mathscr{P}_i}$가 성립하는 않는 점은 모두 $\bigcup_{P \in \mathscr{P}_i} (P \setminus \widetilde{P})$에 속하는데, 여기서 각 $P \setminus \widetilde{P}$는 위의 보조정리에서와 같은 초평면 n개의 합집합에 속하는 집합이므로 0의 측도를 가지게 되어 이 부등식은 거의 어디서나 성립한다.

19 이 함수는 집합 $\mathbb{Q} \cap [0,1]$이 영집합이라는 점에서 거의 어디서나 0이고, 따라서 그 Lebesgue 적분값도 0이다. 여기서 함수 $\mathbf{1}_{\mathbb{Q}}$는 **Dirichlet 함수(− function)**라는 이름을 가지고 있다. 이는 앞서 보았듯이 임의의 비어있지 않은 구간 $I \subseteq \mathbb{R}$에서 Riemann 적분은 불가능하지만 Lebesgue 적분은 가능하며 그 적분값은 항상 0이다. 때로 Dirichlet 함수는 $x \mapsto \lim_{m\to\infty} \lim_{n\to\infty} \cos^{2n}(m!\pi x)$와 같이 연속함수의 극한으로 표현되기도 한다.

20 이는 각주 ? ? 와 거의 비슷하게 보일 수 있다. 각 $i \in \mathbb{N}$에 대해 $2^{-i}\mathbb{Z}^n$에 속하는 점들을 꼭짓점으로 하여 만들 수 있는 모서리의 길이가 2^{-i}인 모든 semi-open box의 모임을 $\mathscr{C}_i$라 하면 이는 가산이고 모든 원소가 서로소이다. 잠시 이에 속하는 semi-open box를 ith dyadic semi-box라 하자. 이제 1st dyadic semi-box 중에서 U에 속하는 것만을 모아 집합족 $\mathscr{D}_1$을 구성한다. 비슷하게, 2nd dyadic semi-box 중에서 $U \setminus \bigcup \mathscr{D}_1$에 속하는 것만을 모아 집합족 $\mathscr{D}_2$를 구성하고 이를 반복하여 집합족의 열 $\{\mathscr{D}_i\}$를 구성한다. 그렇다면 각 $i \in \mathbb{N}$에 대해 $\mathscr{D}_i \subseteq \mathscr{C}_i$에서 $\mathscr{D}_i$는 가산이므로 곧 $\mathscr{D} = \bigcup_{i=1}^\infty \mathscr{D}_i$도 가산임이 분명하고, 구성으로부터 $\mathscr{D}$의 모든 원소들은 서로소이다. 따라서 $\mathscr{D}$의 원소들을 나열하여 WLOG, 필요하다면 공집합을 추가하여 이를 집합열의 형태로 $\{B_i\}$와 같이 쓸 수 있고, 곧 이가 우리가 찾던 집합열임을 쉽게 보일 수 있다.

21 여기서 Φ가 $\mathscr{C}^1$급 함수이므로 그 편미분이 모두 연속함수로서 가측이고, 곧 이들의 곱과 합으로 표현되는 행렬식 $\det \mathbf{D}\Phi$도 가측이다. 따라서 $f \circ \Phi$가 가측이면 $(f \circ \Phi)|\det \mathbf{D}\Phi|$가 가측임은 자명하다.
한편, 가측공간 $(X, \mathscr{A}), (Z, \mathscr{B})$와 $Y \in \mathscr{A}$에 대해 함수 $f : Y \to Z$가 가측이라 함은 f가 $\mathscr{A}|_Y/\mathscr{B}$−가측임을 의미한다.

22 보조정리 ? ? 로부터 임의의 닫힌 box에서 그 경계를 이루는 면 중 일부를 빼어 얻는 집합은 원래의 box와 같은 측도를 가짐을 알 수 있다. 만약 $c > 0$이라면 T^{-1}는 단순히 B를 scaling시킨 것에 불과하여 $T^{-1}(B)$가 다시 semi-open box가 되므로 위의 식이 자명하다. 그러나 $c < 0$인 경우에는 $T^{-1}(B)$가 semi-open box를 그 중심을 축으로 하여 점대칭시킨 모양이 되는데, 방금 지적한 바로부터 $T^{-1}(B)$에서 원래 속한 면은 빼고, 원래 빠진 면은 더하여 이를 다시 semi-open box로 만들도 그 측도는 변함이 없다. 따라서 이 경우에도 위의 식이 성립한다.

23 함수 Ψ_x와 Φ가 모두 $\mathscr{C}^1$급 미분동형사상이므로 이들은 열린집합을 열린집합으로, 닫힌집합을 닫힌집합으로 보내고, 곧 집합 $\Psi_x(B(x,(1-\varepsilon)r)), \Phi(B(x,r))$는 열린집합이고 집합 $\Phi(\overline{B(x,r)})$는 닫힌집합이다. 또한, 같은 이유로 Ψ_x와 Φ는 연결집합을 연결집합으로 보내며, 이러한 사실들은 이 증명의 전반에 걸쳐 별다른 언급 없이 사용되었다.

24 각주 ? ? 를 참조하면 $\mathscr{S}_n$에 속하는 적당한 서로소인 집합열 $\{B_i\}$가 존재하여 $W = \bigsqcup_{i=1}^\infty B_i$이다. 한편, 보조정리 ? ? 로부터 $\bigcup_{i=1}^\infty \partial B_i$가 영집합임을 알 수 있다. 이제 각 $i \in \mathbb{N}$에 대해 $B(x_i, r_i) = B_i^\circ$라 하고

WLOG, 필요하다면 B_i를 가산개의 semi-open box로 분할하여 $r_i < \delta$라 하면 $\{B(x_i, r_i)\}$가 서로소인 집합열임은 자명하고 $\lambda_n(W \setminus \bigsqcup_{i=1}^{\infty} B(x_i, r_i)) \le \lambda_n(\bigcup_{i=1}^{\infty} \partial B_i) = 0$이다.

25 각주 ? ? 를 참조하여 각주 ? ? 에서와 똑같은 방식으로 하되, 각주 ? ? 에서 집합열을 구성할 때 $\overline{B_i} \subseteq U$를 추가적인 조건으로 하여 $\{B_i\}$를 구성하면 된다.

26 집합 $\Phi(F)$가 유계이므로 적당한 유계인 열린집합 $V' \subseteq V$가 존재하여 $\Phi(F) \subseteq V'$이다. 이제 $W = \Phi^{-1}(V')$라 하면 Φ가 $\mathscr{C}^1$급 미분동형사상이라는 점에서 이가 열려있으며 본문의 조건을 모두 만족한다.

27 Gram-Schmidt 정규직교화를 통해 $\beta = \{\beta_1, \cdots, \beta_n\}$가 $\mathbb{R}^n$의 정규직교기저가 되는 동시에 $\gamma = \beta \setminus \{\beta_n\}$가 V의 정규직교기저가 되도록 $\beta_1, \cdots, \beta_n \in \mathbb{R}^n$을 택할 수 있다. 이제 $\mathbb{R}^n$의 표준기저 $\mathscr{E} = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$에 대해 선형사상 $T : \mathbb{R}^n \to \mathbb{R}^n$를 $[T]_{\beta}^{\mathscr{E}} = I$이도록 잡으면 $T(\mathsf{V}) = T(\operatorname{span}\gamma) = \operatorname{span} T(\gamma) = \operatorname{span}(\mathscr{E} \setminus \{\mathbf{e}_n\})$에서 $T(\mathsf{V})$는 $\mathbf{e}_n$과 직교하는 초평면이 되고, $|\det T| = 1$이다.

28 임의의 $x_0 \in \mathbb{R}$에 대해 함수 $f : \mathbb{R} \to \mathbb{R}$를 $f : x \mapsto e^{x-x_0}\mathbf{1}_{(-\infty, x_0]}$로 두면 $F : x \mapsto e^{x-x_0}\mathbf{1}_{(-\infty, x_0]} + \mathbf{1}_{(x_0, \infty)}$이므로 F는 x_0에서 연속이지만 미분불가능하다.

29 변수변환 공식으로부터 임의의 $\mathscr{C}^1$급 미분동형사상 $\Phi : U \to V$와(여기서 $U, V \subseteq \mathbb{R}^n$는 열린집합이다.) 임의의 가측인 $A \in \mathbb{R}^n$를 택하면 $\Phi^{-1}(A)$와 $\Phi(A)$가 모두 가측이므로 $\mathscr{M}_n = \Phi_* \mathscr{M}_n$이다. 한편, 다시 변수변환 공식으로부터

$$
\begin{aligned}
(\Phi_* \lambda_n)(A) &= \lambda_n(\Phi^{-1}(A)) \\
&= \int_U \mathbf{1}_{\Phi^{-1}(A)}\, d\lambda_n \\
&= \int_U \mathbf{1}_A \circ \Phi\, d\lambda_n \\
&= \int_V \mathbf{1}_A / |\det \mathbf{D}\Phi \circ \Phi^{-1}|\, d\lambda_n \\
&= \int_A d(\lambda_n)_{|\det \mathbf{D}\Phi \circ \Phi^{-1}|^{-1}} \\
&= (\lambda_n)_{|\det \mathbf{D}\Phi \circ \Phi^{-1}|^{-1}}(A)
\end{aligned}
$$

이므로 $\Phi_* \lambda_n = (\lambda_n)_{|\det \mathbf{D}\Phi \circ \Phi^{-1}|^{-1}}$이다. 이로부터 임의의 음이 아닌 가측함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$에 대해 $\widetilde{f} = f|\det \mathbf{D}\Phi \circ \Phi^{-1}|$라 하면 $\int_{\Phi^{-1}(V)} \widetilde{f} \circ \Phi\, d\lambda_n = \int_U (f \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n$이고 $\int_V \widetilde{f}\, d\Phi_* \lambda_n = \int_V \widetilde{f}\, d(\lambda_n)_{|\det \mathbf{D}\Phi \circ \Phi^{-1}|^{-1}} = \int_V f\, d\lambda_n$이므로 위의 변수변환 정리는 앞서 본 변수변환 정리의 일반화라 할 수 있다.

30 여기서 $\{\varphi(x) \in A\}$는 조건제시법으로 주어진 집합 $\{x \in X : f(x) \in A\}$의 약식 표현이다. 측도론에서는 혼동의 우려가 없다면 어떤 조건 $P(x)$에 대해 집합 $\{x \in X : P(x)\}$를 $\{P(x)\}$로 간단히 표기하는 관례가 있다.

31 [측도론] 함수 F를 Cantor 함수 $c : [0, 1] \to [0, 1]$로 두면 이는 연속이고 거의 어디서나 미분가능하므로 곧 거의 어디서나 symmetric differentiable하며, $F^{(s)} = 0 =: f$ (ae.)이지만 명백히 $F(x) \neq 0 = \int_{-\infty}^{x} f\, d\lambda_1$이다.

32 사실, 이렇게 다소 아쉬움이 남는 결론에 이른 것은 이후 확률론에서의 사용을 목적으로 일반적인 $n \in \mathbb{N}$에 대해 논의를 전개한 탓이 크다. 만약 우리가 일반성을 포기하고 $n = 1$인 경우에 대해서만 생각했다면 훨씬 더 편하게 더 깔끔한 결론을 얻을 수 있었을 것이다. 예컨대 i의 경우 미분계수의 정의인 극한식을 직접 다룸으로써 보다 쉽게 F가 거의 대부분의 점에서 미분가능하다는(symmetric differentiability가 아닌 통상의 미분가능성이다.) 더 강한 결론을 얻을 수 있고, ii의 경우 유계변동이라는 개념을 통해 F가 증가함수여야 한다는 조건을 획기적으로 완화시킬 수 있다. 이에 관심이 있는 독자들은 [3]을 참조하기 바란다. 하지만 이러한 방식들이 일반적인 n에 대하여 자명하게 일반화되지 않는 까닭에 i은

Hardy-Littlewood의 정리를 이용하여 우회적으로 접근할 수 밖에 없었고, ii의 경우 끝까지 F에게 증가함수여야 한다는 조건을 요구해야만 했다. 게다가 절대연속의 개념은 앞서 설명한 바와 같이 정리 ? ? 가 성립하는 것이 핵심인데, 이를 유지하면서 일반적인 n에 대해 절대연속의 개념을 '예쁘게' 정의하는 것은 꽤나 골치아픈 문제이다. 우리는 정리 ? ? 가 성립하도록 다소 무식하게 절대연속의 개념을 정의했는데, 그 결과 절대연속이 연속성을 함의하지 못하는 '못생긴' 절대연속이 되고 말았다. 이에 대한 보다 자세한 논의는 [7]과 [8]을 참조하기 바란다.

33 합 $\sum_{x\in X} f(x)$가 well-define된다고 함은 합의 과정에서 $\infty - \infty$와 같은 부정형이 등장하지 않고, X가 무한집합인 경우에는 이가 절대수렴하여 합의 순서에 무관하게 그 값이 주어짐을 뜻한다.

References

[1] Jerrold E. Marsden and Michael J. Hoffman, *Elementary Classical Analysis 2nd Edition*, W. H. Freeman, 1993.

[2] Walter Rudin, *Real and Complex Analysis 3rd Edition*, McGraw-Hill, 1987.

[3] Patrick Billingsley, *Probability and Measure 3rd Edition*, Wiley, 1995.

[4] Donald L. Cohn, *Measure Theory 2nd Edition*, Springer, 2013.

[5] Piermarco Cannarsa and Teresa D'Aprile, *Introduction to Measure Theory and Functional Analysis*, Springer, 2015.

[6] 이인석,『선형대수와 군』 개정판, 서울대학교출판문화원, 2015.

[7] Jan Maly, "Absolutely Continuous Functions of Several Variables", *Journal of Mathematical Analysis and Applications*, vol. 231, no. 2, 1999.

[8] Jiri Sremr, "A Note on Absolutely Continuous Functions of Two Variables in The Sense of Carathéodory", *Electronic Journal of Differential Equations*, no. 154, 2010.

[9] Alexander Grigoryan, *Measure Theory and Probability*, University of Bielefeld, 2008.

[10] Dietmar A. Salamon, *Measure and Integration*, ETH Zurich, 2016.

[11] Juha Kinnunen, *Real Analysis*, Aalto University, 2016.

[12] Unknown, *Functions of Real Variables*, Tel Aviv University, 2015.

[13] Tom Lindstrom, *Mathematical Analysis*, University of Oslo, 2013.

[14] Marco Gualtieri, *Geometry and Topology*, University of Toronto, 2009.

[15] Mathematics Stack Exchange, Available: https://math.stackexchange.com.

[16] Wolfram MathWorld, Available: http://mathworld.wolfram.com.

[17] Wikipedia, Available: https://en.wikipedia.org.

[18] 나무위키, Available: https://namu.wiki.

Chapter 2

Probability Theory

Abstract 확률론은 20세기 들어 급격하게 발전한 분야이다. 애초에 확률이라는 개념이 수학에 편입된 것이 그리 오래되지 않았다. 이는 Descart의 연역주의의 영향이 진하게 남아있던 근대 유럽의 수학에서 불확실성을 다루기를 꺼려했기 때문이다. 오죽했으면 "거의 확실한 것은 거의 확실히 거짓이다."라고까지 했을까. 하지만 도박 문제(de Méré's problem)와 같이 불확실성을 계량하여 다루어야 할 필요성은 조금씩 늘어갔고, 이러한 현실적 요구에 확률은 Pascal, Fermat, Lagrange 등의 기라성같은 수학자들에 의해 조금씩 건드려지기 시작했다. 이때까지만 하더라도 확률이 무엇인지에 대한 수학자들의 생각은 '어떤 사건이 발생할 가능성' 정도였다. 이러한 확률의 의미가 직관적으로 분명하였기에 이에 의문을 제기하는 사람도 없었고, 그럴 필요도 느끼지 못했다. 그러나 미적분학에서 극한의 개념이 그러하였듯, 확률에 대한 연구가 계속될수록 미묘한 잡음이 발생하기 시작했고, 이는 확률의 개념에 대한 엄밀한 수학적 접근이 필요함을 암시했다. 결국 '확률은 무엇인가?' 라는 질문의 답을 찾기 위한 긴 여정이 시작되었고, Laplace가 확률에 해석학을 끼얹은 것을 시작으로 Kolomogorov가 그의 명저 *Grundbegriffe der Wahrscheinlichkeitsrechnung*(영어: *Foundations of the Theory of Probability*)에서 측도론으로 확률을 정의하면서 그 여정은 일단락되게 된다. 본 장에서는 그 여정의 끝에서 수학자들이 꿰뚫어본 확률의 본질에 대해 살펴보도록 하자.

2.1 Probability Spaces

단도직입적으로 말하면, 확률은 측도의 특별한 한 종류에 불과하다. 곧 확률은 일종의 넓이나 부피와 같은 개념으로 무언가를 재는 역할을 한다. 현실적인 의미를 생각하면 '가능성'을 잰다고도 할 수 있겠다.

Definition 2.1 측도공간 $(\Omega, \mathscr{F}, \mathbb{P})$에 대해 $\mathbb{P}(\Omega) = 1$이면 이때의 유한 측도 $\mathbb{P}$를 **확률측도(probability measure)**라 하고, 유한 측도공간 $(\Omega, \mathscr{F}, \mathbb{P})$를 **확률공간(probability space)**이라 한다. 나아가 집합 Ω를 **표본공간(sample space)**이라 하고, σ-대수 $\mathscr{F}$에 속하는 임의의 집합 E를 **사건(event)**이라 하여 $\mathbb{P}(E)$의 값을 사건 E의 **확률(probability)**이라 한다.

확률의 본질이 측도라는 위의 정의는 나름 설득력이 있다. 그렇다면 이걸로 다 된 것일까? 아쉽게도 이제부터 해야 할 일이 태산이다. 일단 위의 정의를 받아들이기로 했다면, 지금까지 우리가 배웠던 확률의 대한 모든 내용들을 측도론의 언어로 다시 써야 한다. 곧 확률론을 다루는 본 장의 내용은 기본적으로 '번역 작업'으로, 다행히 대부분의 경우 이 번역 작업은 크게 어렵지 않을 것이다. 이는 측도론이 확률의 내용들을 형식화하기에 좋은 이론이라서이기도 하지만, 앞서 우리가 측도론을 배우며 이 순간을 위해 조금씩 준비해 둔 것들이 꽤 많기 때문이다.

본격적으로 시작하기에 앞서, 맥락상 다소 뜬금없기는 하지만, 우리의 확률에 대한 인식의 근간을 이루는 equally likely outcome model을 한 번은 언급하고 지나가는 것이 좋을 것 같다. 고등학교에서 경우의 수를 세는 문제로 흔히 접하는 equally likely outcome model은 표본공간으로 항상 유한집합 Ω를 가지고, 이의 모든 부분집합은 사건으로 간주된다. 나아가 한원소 집합인 사건은 특별히 **근원사건(elementary event)**이라 불리며 각 근원사건의 확률은 정확히 $1/|\Omega|$로 주어진다. (이렇게 각 근원사건의 확률이 같으므로 'equally likely'이다. 고등학교에서는 흔히 '같은 정도로 확실하다'로 번역한다.) 이러한 setting에서 우리는 임의의 사건 E의 확률을 $|E|/|\Omega|$로 정하고, 여기서의 $|E|$를 구하기 위해 여태껏 경우의 수를 열심히 계산하여 왔다. 이상의 내용을 측도론의 언어로 담백하게 번역하면 equally likely outcome model은 '표본공간 Ω에서의 셈측도 #에 대해 $(\Omega, \mathscr{P}(\Omega), \#/|\Omega|)$로 주어진 확률공간'이 된다. 곧 측도론으로 바라보면 equally likely outcome model도 측도공간의 특별한 한 종류에 불과하다.

보통 초급 확률론 교재에서는 고등학교에서와 마찬가지로 이 equally likely outcome model에 집중하여 경우의 수를 계산하는 화려한 기교를 소개하는 데 많은 분량을 할애하곤 하지만, 여기서는 이에 대한 논의는 Sheldon Ross의 A First Course in Probability를 참고문헌으로 실어두는 것으로 대신한다. 우리는 일반적인 이론의 전개를 목표로 하기에 확률공간의 한 예시에 불과한 equally likely outcome model에 대해서는 그다지 관심이 없고, 곧 이 책에서 equally likely outcome model이 다시 등장하는 일은 (아마) 없을 것이다. 또한, 일반적으로 표본공간 위의 σ-대수 $\mathscr{F}$가 모든 한원소 집합을 포함한다는 보장이 없으므로 근원사건이라는 개념도 딱히 쓸모가 없을 것이다.

다시 원래의 이야기로 돌아와, 본격적으로 번역 작업을 시작해보자. 우선 확률의 기본적인 성질 정도는 측도의 성질로부터 거의 자명하게 얻어진다.

Theorem 2.2 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건 E, F에 대해 다음이 성립한다.

i. $\mathbb{P}(\emptyset) = 0$.
ii. (σ−가법성) 서로소인 사건열 $\{E_i\}$에 대해 $\mathbb{P}(\bigsqcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$이다.
iii. $0 \leq \mathbb{P}(E) \leq 1$.
iv. $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$.
v. (단조성) 만약 $E \subseteq F$이면 $\mathbb{P}(E) \leq \mathbb{P}(F)$이다.

Proof 이는 정의와 측도의 기본적인 성질로부터 자명하다. □

Theorem 2.3 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에 대해 다음이 성립한다.

i. (포함배제의 원리) 사건 $E_1, \cdots, E_l$에 대해

$$\mathbb{P}\left(\bigcup_{i=1}^{l} E_i\right) = \sum_{i=1}^{l} (-1)^{i-1} \sum_{1 \leq j_1 < \cdots < j_i \leq l} \mathbb{P}\left(\bigcap_{k=1}^{i} E_{j_k}\right)$$

이다.
ii. (σ−반가법성) 사건열 $\{E_i\}$에 대해 $\mathbb{P}(\bigcup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} \mathbb{P}(E_i)$이다.

Proof i는 측도론의 포함배제의 원리를 $(\Omega, \mathscr{F}, \mathbb{P})$에 적용한 결과이고, ii는 측도의 σ−가법성이 σ−반가법성을 함의한다는 점에서 자명하다. □

Theorem 2.4 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건열 $\{E_i\}$에 대해

$$\mathbb{P}(\liminf_{i\to\infty} E_i) \leq \liminf_{i\to\infty} \mathbb{P}(E_i) \leq \limsup_{i\to\infty} \mathbb{P}(E_i) \leq \mathbb{P}(\limsup_{i\to\infty} E_i)$$

가 성립한다. 특별히, $E_i \to E$이면 $\mathbb{P}(E_i) \to \mathbb{P}(E)$이다.

Proof 이는 정리 ? ? 로부터 자명하다. □

위의 정리의 후단을 흔히 **확률측도의 연속성(continuity of probability measure)**이라 한다. 한편, 확률론에서는 사건열 $\{E_i\}$에 대해 이의 상극한 $\limsup_{i\to\infty} E_i$를 간단히 E_i io. 라 쓰기도 한다. 여기서 io.는 infinitely often의 줄임말로 곧 $\omega \in E_i$ io.는 $\omega \in \Omega$가 사건 $E_1, E_2, \cdots$에 무한히 많이 속한다는 것인데, 이는 $\limsup_{i\to\infty} E_i = \bigcap_{j=1}^{\infty} \bigcup_{i=j}^{\infty} E_i$임을 생각해 보면 나름 make sense하는 표기법이다.

Definition 2.5 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에 대해 영집합인 사건을 **영사건(null event)**이라 한다.

확률측도도 측도이기에 '($\mathbb{P}$−)거의 어디서나'라는 개념이 자주 쓰이는데, 확률론에서는 이를 **($\mathbb{P}$−)거의 확실하게(($\mathbb{P}$−)almost surely)**라 하기도 한다. 이는 영사건이 확률이 0인 사건이므로 어떤 성질이 $\mathbb{P}$−거의 어디서나 성립하면 곧 1의 확률로 성립하게 되기 때문에 생겨난 관례이다.

이와 관련하여 '확률이 0인 사건'과 '불가능한 사건'은 서로 다르다는 것에 주의할 필요가 있다. 확률인 0인 사건은 표본공간 위의 σ−대수 $\mathscr{F}$에 속하는 사건이지만 그 확률이 0일 뿐이고, 불가능한 사건은 애초에 $\mathscr{F}$에 속하지 않아 그 이름과는 달리 엄밀히는 사건이 아니다. 간단한 예시를 위해 $[0,1]$에서 임의로 점 하나를 택하는 상황을 생각해보자. 이를 확률공간으로 형식화한다면, 표본공간은 $\Omega = [0,1]$이고 $\mathscr{F} = \mathscr{B}_1|_\Omega$, 확률측도는 $\mathbb{P} = (\mu_1)_{\mathscr{F}}$ 정도로 둘 수 있을 것이다. 그렇다면 $\mathbb{P}\{0.5\} = 0$이므로 정확히 0.5를 뽑는 사건은 확률이 0인 사건이지만, 그렇다고 이가 일어나는 것 자체가 불가능한 것은 아니다. 한편, 2를 뽑는 사건은 애초에 발생이 불가능한 사건으로 이 경우에 $\{2\} \notin \mathscr{F}$이므로 이는 엄밀하게는 사건이 아니어서 확률의 부여가 불가능하다. 이와 비슷하게, '확률이 1인 사건'과 '항상 발생하는 사건'도 서로 다르다.

이어서, 조건부확률을 도입하고 그 성질을 측도론으로 보이자. 다만, 이후에 조건부기댓값과 조건부분포를 엄밀히 도입하기 위해서는 조건부확률의 개념을 격변에 준할 정도로 일반화시켜야 하는데, 이에 하나의 절을 오롯이 할애해야 할 정도의 논의가 필요하므로 여기에서는 아주 간단하게만 다루도록 한다.

Definition 2.6 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 영사건이 아닌 사건 E에 대해 **사건 E에 대한 조건부확률(conditional probability under event E)**을 $\mathbb{P}(\cdot|E) : \mathscr{F} \to \mathbb{R}_0^+$로 쓰고 $\mathbb{P}(\cdot|E) : F \mapsto \mathbb{P}(F \cap E)/\mathbb{P}(E)$로 정의한다.

Proposition 2.7 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 영사건이 아닌 사건 E에 대해 조건부확률 $\mathbb{P}(\cdot|E)$는 $(\Omega, \mathscr{F})$ 위의 확률측도이다. 따라서 $(\Omega, \mathscr{F}, \mathbb{P}(\cdot|E))$는 확률공간을 이룬다.

Proof 우선 $\mathbb{P}(\emptyset|E) = \mathbb{P}(\emptyset)/\mathbb{P}(E) = 0$, $\mathbb{P}(\Omega|E) = \mathbb{P}(E)/\mathbb{P}(E) = 1$임은 분명하고, 임의의 서로소인 사건열 $\{E_i\}$에 대해

$$\begin{aligned}
\mathbb{P}\left(\bigsqcup_{i=1}^{\infty} E_i | E\right) &= \frac{\mathbb{P}(\bigsqcup_{i=1}^{\infty} E_i \cap E)}{\mathbb{P}(E)} \\
&= \frac{\mathbb{P}(\bigsqcup_{i=1}^{\infty} (E_i \cap E))}{\mathbb{P}(E)} \\
&= \sum_{i=1}^{\infty} \frac{\mathbb{P}(E_i \cap E)}{\mathbb{P}(E)} \\
&= \sum_{i=1}^{\infty} \mathbb{P}(E_i | E)
\end{aligned}$$

이므로 $\mathbb{P}(\cdot|E)$가 확률측도임을 안다. □

Theorem 2.8 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에 대해 다음이 성립한다.

i. 사건 E와 영사건이 아닌 사건 F에 대해 $\mathbb{P}(E \cap F) = \mathbb{P}(E|F)\mathbb{P}(F)$이다.

ii. (전확률 공식) 서로소인 가산개의 사건 $E_1, E_2, \cdots$에 대해 각 E_i가 영사건이 아니고 $\bigsqcup_{i=1}^k E_i = \Omega$이면 임의의 사건 E에 대해 $\mathbb{P}(E) = \sum_{i=1}^k \mathbb{P}(E|E_i)\mathbb{P}(E_i)$이다. (여기서 k는 유한할 수도 있고, ∞일 수도 있다.)

PROOF i. 이는 조건부확률의 정의로부터 자명하다.

ii. i로부터 $\mathbb{P}(E) = \mathbb{P}(E \cap \bigsqcup_{i=1}^k E_i) = \mathbb{P}(\bigsqcup_{i=1}^k (E \cap E_i)) = \sum_{i=1}^k \mathbb{P}(E \cap E_i) = \sum_{i=1}^k \mathbb{P}(E|E_i)\mathbb{P}(E_i)$ 이다. □

Theorem 2.9 (Bayes) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 서로소인 가산개의 사건 $E_1, E_2, \cdots$에 대해 각 E_i가 영사건이 아니고 $\bigsqcup_{i=1}^k E_i = \Omega$이면 임의의 사건 E에 대해

$$\mathbb{P}(E_1|E) = \frac{\mathbb{P}(E|E_1)\mathbb{P}(E_1)}{\sum_{i=1}^k \mathbb{P}(E|E_i)\mathbb{P}(E_i)}$$

이다. (여기서 k는 유한할 수도 있고, ∞일 수도 있다.)

PROOF 전확률 공식으로부터 $\mathbb{P}(E_1|E) = \mathbb{P}(E \cap E_1)/\mathbb{P}(E) = \mathbb{P}(E|E_1)\mathbb{P}(E_1)/\sum_{i=1}^k \mathbb{P}(E|E_i)\mathbb{P}(E_i)$가 자명하다. □

비록 증명은 간단하지만 Bayes의 정리는 실험적으로 구하는 것이 불가능한 조건부확률을 구하게 해준다는 엄청난 실용성과 함의를 지닌다. 쉽고 즐거운 예시를 위해 유리가 예나에게 휴대전화로 0x2661(UTF-16 ♡)과 0x2665(UTF-16 ♥) 중 하나의 정보를 임의로 전송하였는데, 송수신의 과정에서 정보의 일부가 손상되어 결과적으로 예나는 0x2663(UTF-16 ♣)을 수신한 상황을 생각해보자. 이를 복원하기 위해 예나는 E를 0x2663을 수신하는 사건, F, G를 각각 0x2661과 0x2665를 전송하는 사건이라 두고 두 조건부확률 $\mathbb{P}(F|E)$와 $\mathbb{P}(G|E)$를 비교하여 전자가 더 크다면 0x2661로 복원하고 후자가 더 크다면 0x2665로 복원하며, 만약 같다면 재전송을 요청하기로 하였다. 이는 훌륭한 통계적 사고방식이지만 $\mathbb{P}(F|E)$와 $\mathbb{P}(G|E)$는 일의 선후가 뒤바뀐 확률이라 실험적으로 구할 수가 없다는 치명적인 문제가 있다. 예나가 자신의 휴대전화로 자신에게 0x2661과 0x2665를 수 회 전송하는 실험을 통해 근사하게나마 구할 수 있는 확률은 $\mathbb{P}(E|F)$와 $\mathbb{P}(E|G)$ 뿐이다. 여기서 Bayes의 정리는 $\mathbb{P}(F|E)$와 $\mathbb{P}(G|E)$를 $\mathbb{P}(E|F)$와 $\mathbb{P}(E|G)$의 조합으로써 계산할 수 있도록 하여 문제를 해결하는 결정적인 역할을 한다. 따라서 예나가 실험적으로 구한 $\mathbb{P}(E|F)$와 $\mathbb{P}(E|G)$의 근사치가 만약 $1/200, 1/300$이었다면 Bayes의 정리로부터 $\mathbb{P}(F|E) \approx 3/5, \mathbb{P}(G|E) \approx 2/5$가 되어 ♣를 ♡로 복원할 수 있다. 요컨대, Bayes의 정리는 선후관계나 인과관계를 역전시켜주는 마법의 공식이다.

이러한 Bayes 정리는 이후 통계학에 큰 지각변동을 일으켜 이를 기초로 하는 Bayesian이라는 독자적인 학파가 구성되기에 이르렀고, 오늘날 기계학습과 같은 분야에서 요긴하게 쓰이는 모양이다. (이와 구분하여 기존의 통계학 학파를 빈도주의라 한다.) 물론, 이런 학

파는 어디까지나 확률의 해석에 대한 차이로 구분되는 것이지 확률의 측도론적 정의나 접근방식으로 구분되는 것은 아니기에 빈도주의와 Bayesian의 구분이 본 장에서는 필요하지 않지만, 이 책에서는 특별한 언급이 없는 이상 빈도주의의 관점에서 확률을 바라본다. 여기에는 통계학 교양 정도를 들은 수준에서는 빈도주의의 관점이 Bayesian의 관점보다 조금 더 친숙하다는 것을 빼면 다른 그럴싸한 이유는 없다. 만약 자신이 Bayesian이라면 그들의 방식대로 해석하면 그만이고, 당연히 Bayesian에게도 측도론적인 엄밀한 확률론은 훌륭한 이론의 토대가 될 것이다.

이번 절에서 마지막으로 다룰 것은 바로 독립에 관한 내용인데, 앞서 확률의 기본적인 성질과 기본적인 조건부확률을 큰 어려움 없이 도입할 수 있었던 것과는 달리, 독립성을 측도론의 언어로 번역해 내는 것은 살짝 어렵다. 이는 독립이라는 개념이 측도론에서는 그 느낌조차 찾아보기 힘든, 완전히 새로운 개념이기 때문이다. 우리는 측도론에서 적분을 정의할 때와 비슷하게 독립을 그 정의를 점차 일반화시키는 방법으로 도입할 것이다. 우선 가장 기본적인 독립의 정의로 시작한다.

Definition 2.10 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건 $E_1, \cdots, E_k$를 생각하자. 만약 각 $l \leq k$과 임의의 서로다른 $i_1, \cdots, i_l \leq k$에 대해 $\mathbb{P}(\bigcap_{j=1}^{l} E_{i_j}) = \prod_{j=1}^{l} \mathbb{P}(E_{i_j})$가 성립하면 이때의 사건 $E_1, \cdots, E_k$를 **(서로) 독립((mutually) independent)**이라 한다. 한편, 만약 위의 성질이 $l = 2$에 대해서만 만족되면, 즉 임의의 서로다른 $i, j \leq k$에 대해서 $\mathbb{P}(E_i \cap E_j) = \mathbb{P}(E_i)\mathbb{P}(E_j)$가 성립하는 것에 그치면 이때의 사건 $E_1, \cdots, E_k$를 **pairwise 독립(− independent)**이라 한다.

서로 독립인 것과 pairwise 독립이 다르다는 사실은 잘 알려진 사실이다. 흔해빠진 주사위와 동전 예시를 피하기 위해 해석개론을 수강신청한 수지와 이제훈이 같이 밤새 과제를 하게 된 것을 계기로 서로 어느정도 이성으로서 호감을 가지게 된, 이른바 '썸 탄다' 불리우는 상황을 생각하자. 수업을 듣던 어느날, 수지와 이제훈의 교재가 실수로 서로 바뀌어 다음 수업시간 전에 만나 책을 다시 바꾸기로 하였는데, 둘은 이를 앞두고 책 사이에 고백편지를 살짝 끼워넣을까 고민하고 있다고 하자. 여기서 사건 E, F를 각각 수지와 이제훈이 고백편지를 끼워넣는 사건이라 하고, 수지나 이제훈이 고백편지를 넣을 확률은 $1/2$로 같으며 이는 서로 독립이라 하자. 이제 사건 G를 어느 한쪽만 고백편지를 받는 사건이라 하면 E, F, G는 pairwise 독립이지만 서로 독립은 아니다. (직접 계산해보자.)

한편, 유한개의 사건 사이의 독립을 무한개의 사건 사이의 독립으로 확장하는 것은 어렵지 않다.

Definition 2.11 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건의 모임 $\mathscr{C}$를 생각하자. 만약 임의의 사건 $E_1, \cdots, E_k \in \mathscr{C}$가 서로 독립이면 이때의 집합족 $\mathscr{C}$를 **독립(independent)**이라 한다.

다음으로, 유한개의 사건의 모임 사이의 독립을 정의한다.

Definition 2.12 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건의 모임 $\mathscr{C}_1, \cdots, \mathscr{C}_k$를 생각하자. 만약 각 $i \le k$에 대해 임의로 택한 사건 $E_i \in \mathscr{C}_i$가 서로 독립이면 이때의 집합족 $\mathscr{C}_1, \cdots, \mathscr{C}_k$를 **독립 (independent)**이라 한다.

마지막으로 이를 무한개의 사건의 모임 사이의 독립으로까지 확장하면 독립의 가장 일반적인 정의를 얻는다.

Definition 2.13 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건의 모임의 모임 Γ를 생각하자. 만약 임의의 사건의 모임 $\mathscr{C}_1, \cdots, \mathscr{C}_k \in \Gamma$가 서로 독립이면 이때의 집합족 Γ를 **독립(independent)**이라 한다.

이렇게까지 일반적인 형태의 독립성을 고려하는 이유는 다음 정리 때문이다.

Theorem 2.14 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 독립인 사건의 모임의 모임 $\{\mathscr{C}_\alpha\}$에 대해 각 $\mathscr{C}_\alpha$가 $\pi-$system이라 하면 $\{\sigma(\mathscr{C}_\alpha)\}$는 독립이다.

Proof 집합족 $\{\mathscr{C}_\alpha\}$에서 임의로 유한개의 원소를 택하여 이를 $\mathscr{C}_1, \cdots, \mathscr{C}_k$라 하고 이들이 생성하는 $\sigma-$대수 $\sigma(\mathscr{C}_1), \cdots, \sigma(\mathscr{C}_k)$가 서로 독립임을 보이면 증명은 충분하다. 이를 위해 사건 $E_2 \in \mathscr{C}_2, \cdots, E_k \in \mathscr{C}_k$를 임의로 택하고 집합족 $\mathscr{L} = \{F \in \mathscr{F} : F, E_2, \cdots, E_k$가 서로 독립$\}$을 생각하면 주어진 조건으로부터 $\mathscr{C}_1 \subseteq \mathscr{L}$임은 분명하다. 나아가 $\Omega \in \mathscr{L}$ 또한 분명하고, 임의의 $F \in \mathscr{L}$에 대해 $F, E_2, \cdots, E_k$가 서로 독립이면 $F^c, E_2, \cdots, E_k$도 서로 독립임을 쉽게 보일 수 있으므로 $F^c \in \mathscr{L}$이다. 비슷한 방법으로 $\mathscr{L}$에 속하는 임의의 서로소인 사건열 $\{F_i\}$에 대해 $\bigsqcup_{i=1}^{\infty} F_i, E_2, \cdots, E_k$가 서로 독립임을 보일 수 있으므로 $\bigsqcup_{i=1}^{\infty} F_i \in \mathscr{L}$도 성립한다. 이로부터 $\mathscr{L}$은 $\lambda-$system이 되어 $\mathscr{C}_1$이 $\pi-$system이라는 사실과 Dynkin의 $\pi-\lambda$ 정리로부터 $\sigma(\mathscr{C}_1) = \lambda(\mathscr{C}_1) \subseteq \mathscr{L}$이고, 곧 $\sigma(\mathscr{C}_1), \mathscr{C}_2, \cdots, \mathscr{C}_k$는 서로 독립이다. 이제 이를 $k-1$번 반복하면 $\sigma(\mathscr{C}_1), \cdots, \sigma(\mathscr{C}_k)$가 서로 독립임을 보일 수 있고, 증명이 끝난다. □

위의 정리는 서로 독립인 사건의 모임들이 생성하는 $\sigma-$대수도 서로 독립임을 함의하는데, 잠시 이 결과의 의미에 대해 생각해보자. 독립성은 기본적으로 '정보'에 대한 이야기이다. 가장 기본적인 형태로 두 사건 E, F가 서로 독립인 경우를 살펴보면, 이는 사건 E의 발생여부에 대한 정보가 주어지더라도 사건 F의 발생여부에 대한 정보는 일체 추론해 낼 수 없음을 의미한다. 이는 2개 이상의 사건, 나아가 무한개의 사건의 독립에 대해서도 마찬가지이다. 예컨대 앞서 든 수지와 이제훈의 예시에서 E, F, G는 서로 독립이 아니었는데, 이는 수지가 고백편지를 넣는 사건 E와 이제훈이 고백편지를 넣는 사건 F의 발생여부에 대한 정보가 사건 G의 발생여부에 대한 정보를 100% 함의하기 때문이다.

사건의 모임 사이의 독립도 이와 비슷하게 이해할 수 있다. 앞서 사건들 사이의 독립에서 각 사건은 그 사건의 발생여부에 대한 이진 정보를 의미했다. 그렇다면 사건의 모임은 그 모임에 속한 각 사건들의 발생여부에 대한 이진 정보의 조합으로 이루어진 보다 풍성한

정보로 이해하는 것이 자연스러울 것이다. 따라서 사건의 모임 사이의 독립은 각 사건의 모임이 함의하는 정보가 서로 독립적이라는, 즉 어느 하나를 알더라도 다른 하나를 일체 추론해내지 못한다는 것을 의미한다. 이러한 관점에서 보면, 위의 정리의 결과는 어떤 사건의 모임들이 함의하는 정보가 서로 추론이 불가능하다면, 각 모임을 그가 생성하는 σ−대수로 확장하여 훨씬 더 많은 정보로 얻어도, 여전히 서로 추론이 불가능함을 의미한다. 뭔가 자명한 듯 자명하지 않은 이 결과를 보다 효율적으로 쓰기 위해 따름정리 하나를 소개한다.

Corollary 2.15 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 가산개의 무한한 행과 열을 가지는 사건의 배열

$$\begin{matrix} E_{11} & E_{12} & \cdots \\ E_{21} & E_{22} & \cdots \\ \vdots & \vdots & \ddots \end{matrix}$$

에 대해 $\{E_{ij}\}_{ij}$가 독립이라 하고, 각 $i \in \mathbb{N}$에 대해 집합족 $\mathscr{G}_i$를 $\{E_{ij}\}_j$가 생성하는 σ−대수라 하면 $\{\mathscr{G}_i\}$는 독립이다. 한편, 행이나 열의 개수가 유한한 경우에도 같은 결과가 성립하며, 나아가 각 행의 열의 개수가 달라도 가산개이기만 하면 여전히 같은 결과가 성립한다.

Proof 각 $i \in \mathbb{N}$에 대해 $\{E_{ij}\}_j$의 모든 유한 교집합의 모임 $\mathscr{P}_i$를 생각하면 $\mathscr{P}_i$는 π−system이고 $\sigma(\mathscr{P}_i) = \mathscr{G}_i$임이 거의 분명하다. 따라서 $\{\mathscr{P}_i\}$가 서로 독립이라는 사실만 보이면 앞선 정리로부터 $\{\mathscr{G}_i\}$가 독립이 되어 증명이 끝난다. 이를 위해 $\{\mathscr{P}_i\}$에서 임의로 유한개의 원소를 택하여 이를 $\mathscr{P}_1, \cdots, \mathscr{P}_k$라 하고 다시 임의로 사건 $F_1 \in \mathscr{P}_1, \cdots, F_k \in \mathscr{P}_k$를 택하면 각 $i \le k$에 대해 $\mathscr{P}_i$의 구성으로부터 적당한 사건 $E_{i1}, \cdots, E_{il_i}$가 존재하여 $F_i = \bigcap_{j=1}^{l_i} E_{ij}$이다. 그렇다면 $\{E_{ij}\}_{ij}$가 독립이라는 사실로부터 $\mathbb{P}(\bigcap_{i=1}^k F_i) = \mathbb{P}(\bigcap_{i=1}^k \bigcap_{j=1}^{l_i} E_{ij}) = \prod_{i=1}^k \prod_{j=1}^{l_i} \mathbb{P}(E_{ij}) = \prod_{i=1}^k \mathbb{P}(\bigcap_{j=1}^{l_i} E_{ij}) = \prod_{i=1}^k \mathbb{P}(F_i)$가 되어 $\{\mathscr{P}_i\}$가 서로 독립임을 알고, 증명은 이로써 충분하다. 한편, 행이나 열의 개수가 유한하거나 각 행의 열의 수가 다른 경우에도 이와 비슷하게 하면 된다. □

이 따름정리를 적당히 응용하면 독립성에 대한 진부한 연습문제들, 예컨대 사건 E, F, G, H가 서로 독립이라면 $E \cap F$와 $G \setminus H$가 독립임을 보이라는 식의 문제들을 아주 깔끔하게 해결할 수 있다. 예시로 든 문제의 경우 배열 $E\,F//G\,H$를 생각하면 그만이다. 한편, 앞서 사건의 모임을 그에 포함된 각 사건의 발생여부로 구성된 정보의 집합으로 해석하였는데, 이는 독립의 경우에만 한정되는 해석이 아니어서 특히 그 모임이 σ−대수 $\mathscr{G}$인 경우 모든 사건의 집합 $\mathscr{F}$를 '전체 정보'로, $\mathscr{G}$는 이의 '부분 정보'로 해석하는 것이 유용한 경우가 많다.

나중을 위해 측도론에서 잠시 등장했던 Borel-Cantelli의 정리를 조금 보강하는 것으로 이번 절을 마친다.

Theorem 2.16 (Borel-Cantelli) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건열 $\{E_i\}$에 대해 다음이 성립한다.

i. 만약 $\sum_{i=1}^{\infty}\mathbb{P}(E_i) < \infty$이면 $\mathbb{P}(E_i \text{ io.}) = 0$이다.
ii. 만약 $\sum_{i=1}^{\infty}\mathbb{P}(E_i) = \infty$이고 $\{E_i\}$가 독립이면 $\mathbb{P}(E_i \text{ io.}) = 1$이다.

PROOF i. 이는 측도론에서 배운 Borel-Cantelli의 정리로부터 자명하다.

ii. 우선 $\{E_i\}$가 독립이므로 따름정리 ? ? 로부터 $\{E_i^c\}$도 독립이다. 이제 임의의 $\varepsilon > 0$과 임의의 $j \in \mathbb{N}$를 택하면 $\sum_{i=j}^{\infty}\mathbb{P}(E_i) = \infty$이므로 $\lim_{k\to\infty}\exp(-\sum_{i=j}^{k}\mathbb{P}(E_i)) = 0$이다. 이로부터 적당한 $k_0 \in \mathbb{N}$가 존재하여 $k_0 \geq j$이고 $\exp(-\sum_{i=j}^{k_0}\mathbb{P}(E_i)) < \varepsilon$이므로 $\mathbb{P}(\bigcap_{i=j}^{k_0}E_i^c) = \prod_{i=j}^{k_0}\mathbb{P}(E_i^c) = \prod_{i=j}^{k_0}[1-\mathbb{P}(E_i)] \leq \prod_{i=j}^{k_0}\exp(-\mathbb{P}(E_i)) = \exp(-\sum_{i=j}^{k_0}\mathbb{P}(E_i)) < \varepsilon$이다. 이는 곧 $\mathbb{P}(\bigcap_{i=j}^{\infty}E_i^c) \leq \mathbb{P}(\bigcap_{i=j}^{k_0}E_i^c) < \varepsilon$임을 뜻하므로 $\mathbb{P}(\bigcap_{i=j}^{\infty}E_i^c) = 0$임을 알고, 이는 다시 $\mathbb{P}(E_i \text{ io.}) = \mathbb{P}((\liminf_{i\to\infty}E_i^c)^c) = 1-\mathbb{P}(\liminf_{i\to\infty}E_i^c) \geq 1-\sum_{j=1}^{\infty}\mathbb{P}(\bigcap_{i=j}^{\infty}E_i^c) = 1$에서 $\mathbb{P}(E_i \text{ io.}) = 1$임을 뜻한다. □

2.2 Random Variables and Random Vectors

앞선 절에서 사건 그 자체에 관련된 확률의 내용들을 측도론의 틀에 맞추어 열심히 옮겨 놓았으니, 이번 절에서는 확률벡터라는 개념을 추가하여 내용을 보다 풍성하게 만들어보도록 하자. 확률벡터를 통해 우리는 일일히 사건을 정의하지 않고도 확률의 여러 개념들을 편리하게 사용할 수 있다.

Definition 2.17 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에 대해 가측함수 $X : \Omega \to \mathbb{R}^n$를 **($n$차원) 확률벡터(($n$ dimensional) random vector)**라 하고, 특별히 $n = 1$이면 **확률변수(random variable)**라 한다.

확률벡터는 본질적으로 가측함수이기에 확률벡터의 기본적인 성질들이 가측함수의 성질들로부터 자명하게 성립하는 것이 전혀 이상하지 않다.

Proposition 2.18 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 함수 $X : \Omega \to \mathbb{R}^n$에 대해 X가 rv.일 필요충분조건은 $X_1, \cdots, X_n$이 모두 rv.인 것이다.

PROOF 이는 벡터함수가 가측일 필요충분조건은 그 성분이 모두 가측인 것이라는 점에서 자명하다. □

Theorem 2.19 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^m$와 Borel 함수 $g : \mathbb{R}^m \to \mathbb{R}^n$에 대해 합성 $g \circ X : \Omega \to \mathbb{R}^n$도 rv.이다.

PROOF 이는 Borel 함수와 가측함수의 합성은 가측이라는 점에서 자명하다. □

정의상 확률벡터는 엄연히 함수이지만, 그 이름에서도 잘 드러나듯이 확률론에서는 이를 마치 변수처럼 생각하고 사용하는 경우가 많다. 이는 이론적인 이유에서라기보다 실생활의 응용에서 이렇게 생각하는 편이 조금 더 직관적으로 편하기 때문이다. 이러한 우리의 인식은 확률벡터와 관련된 여러 표기상의 관례에 잘 나타나는데, 위의 정리에서의 합성 $g\circ X$를 마치 g에 X라는 변수를 대입한 것으로 생각하여 $g(X)$로 쓰는 관례가 대표적인 예시이다. 그러나 이런 표기법은 어디까지나 관례일 뿐, 확률벡터가 변수가 아닌 함수라는 사실은 항상 염두에 두고 있어야 한다.

이어서, 측도론에서 FTC를 일반화하는 과정에서 잠시 스쳐 지나갔던 pushfowarding이 다시 등장한다. 비록 측도론에서의 pushfowarding은 도구 역할에 그쳤지만, 확률론에서의 pushfowarding은 빼놓을 수 없는 핵심적인 개념이다.

Definition 2.20 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}^n$에 대해 pushfoward 측도 $X_*\mathbb{P}$를 rv. X의 **분포(distribution)**라 하고 $\mathbf{P}_X$로 쓴다. 특별히, $n\geq 2$인 경우 $\mathbf{P}_X$를 rv. $X_1,\cdots,X_n$의 **결합분포(joint distribution)**라 하기도 한다.

교양 통계학에서 배운 PDF나 CDF를 비롯하여 확률론에는 방금 정의한 분포와 쉽게 혼동될 법한 개념들이 많고, 실제로 각자의 정의도 서로 긴밀히 연결되어 있다. 여기에 한술 더 떠서 문헌마다 조금씩 용어를 다르게 쓰는 바람에 혼란이 가중되는 부분이 없지 않지만, 대부분 논의의 맥락으로 적당히 구별할 수 있다. 아무튼 구태여 혼란을 초래할 필요는 없기에, 이 책에서는 용어를 최대한 잘 구별하여 사용하였다.

나중에 하나씩 보겠지만, 확률벡터가 가지는 성질과 특성 중에는 그 확률벡터 자체가 아닌, 그 분포에 의존하는 성질들이 있다. 곧, 이러한 성질들은 서로다른 확률공간에서 정의된 서로다른 확률벡터라도 분포만 같다면 서로 공유하게 된다. 이를 뒤집어 생각하면 때로는 하나의 확률공간에 정의된 확률벡터를 그 분포에 의존하는 성질들을 그대로 유지하면서 다른 확률공간으로 옮길 수도 있다. 이러한 이유로 분포가 같은 확률벡터를 다루게 되는 경우가 많고, 표기의 편의를 위해 확률공간 $(\Omega,\mathscr{F},\mathbb{P}),(\Omega',\mathscr{F}',\mathbb{P}')$에서 각각 정의된 확률벡터 $X:\Omega\to\mathbb{R}^n, Y:\Omega'\to\mathbb{R}^n$에 대해 $\mathbf{P}_X=\mathbf{P}_Y$이면 이를 간단히 $X\equiv Y$로 쓴다. 이때 $X\equiv Y$이지만 $X\neq Y$일 수 있다는 사실에 주의해야 한다.

Proposition 2.21 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}^n$에 대해 분포 $\mathbf{P}_X$는 $\mathscr{B}_n$ 위의 측도이다. 따라서 $(\mathbb{R}^n,\mathscr{B}_n,\mathbf{P}_X)$는 확률공간을 이룬다.

Proof 먼저 $\mathscr{B}_n\subseteq X_*\mathscr{F}$임을 보이기 위해 임의의 $A\in\mathscr{B}_n$를 택하면 $X^{-1}(A)\in\mathscr{F}$이므로 $A\in X_*\mathscr{A}$에서 $\mathscr{B}_n\subseteq X_*\mathscr{F}$이다. 따라서 $\mathbf{P}_X$가 확률측도임을 보이면 충분한데, 이는 $\mathbf{P}_X(\mathbb{R}^n)=\mathbb{P}(X^{-1}(\mathbb{R}^n))=\mathbb{P}(\Omega)=1$에서 쉽게 알 수 있고, 곧 증명이 끝난다. □

위의 명제는 서로다른 확률공간에 정의된 확률측도를 각각 다루는 대신 이들을 pushfowarding하여 $\mathscr{B}_n$에서 정의된 확률측도인 분포로 일관되게 다룰 수 있음을 함의한다. 그리고 생각해

보면, 이가 곧 사건을 직접 정의하고 사용하는 대신 확률벡터를 도입하여 사용하는 이유이다. 무엇이 될 지 모르는 확률공간 대신 우리가 잘 알고있는 실수공간에서의 확률측도를 다루는 것이 훨씬 편하다. 한편, 위의 정리의 역이 성립한다는 것도 꽤나 흥미로운 사실이다.

Theorem 2.22 Borel $\sigma-$대수 $\mathscr{B}_n$ 위의 확률측도 μ에 대해 적당한 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$가 존재하여 $\mu = \mathbf{P}_X$이다. 특별히, 이때 $(\Omega, \mathscr{F}) = (\mathbb{R}^n, \mathscr{B}_n)$이도록 잡을 수 있다.

Proof 거의 자명하다. 함수 $X : \mathbb{R}^n \to \mathbb{R}^n$를 항등함수로 두면 이는 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mu)$에서 정의된 rv.이고, 임의의 사건 E에 대해 $\mathbf{P}_X(E) = \mu(X^{-1}(E)) = \mu(E)$에서 $\mu = \mathbf{P}_X$이다. □

Theorem 2.23 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^m$와 Borel 함수 $g : \mathbb{R}^m \to \mathbb{R}^n$에 대해 $\mathbf{P}_{g(X)} = \mathbf{P}_X \circ g^{-1}$이다.

Proof 임의의 $A \in \mathscr{B}_n$에 대해 $\mathbf{P}_{g(X)}(A) = \mathbb{P}((g \circ X)^{-1}(A)) = \mathbb{P}(X^{-1}(g^{-1}(A))) = \mathbf{P}_X(g^{-1}(A)) = (\mathbf{P}_X \circ g^{-1})(A)$이므로 $\mathbf{P}_{g(X)} = \mathbf{P}_X \circ g^{-1}$이다. □

자연스러운 다음 순서는 고등학교 시절부터 들어와 이름만은 익숙한 이산확률변수와 연속확률변수를 엄밀하게 측도론의 언어로 정의하는 것이다. 물론, 고등학교나 교양 통계학에서 각각을 정의하지 않는 것은 아니지만, 그 정의가 뭔가 어색하고 작위적이라는 느낌을 지우기 힘든데, 아래의 측도론적인 정의는 더할 나위 없이 깔끔하고 명쾌하다.

Definition 2.24 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 $\mathbf{P}_X$가 이산측도이면 이때의 rv. X를 **이산확률벡터(discrete rv.)**라 한다. 또한, 만약 $\mathbf{P}_X$가 μ_n에 대해 절대연속이거나 특이연속이면 이때의 rv. X를 각각 **연속확률벡터(continuous rv.)** 혹은 **특이확률벡터(singular rv.)**라 한다. 특별히, $n = 1$인 경우 이산확률벡터, 연속확률벡터, 특이확률벡터를 각각 **이산확률변수**, **연속확률변수**, **특이확률변수**라 한다.

Proposition 2.25 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 이는 이산확률벡터, 연속확률벡터, 특이확률벡터의 정의 중에서 두 개의 이상을 동시에 만족시킬 수 없다.

Proof 모순을 유도하기 위해 X가 이산확률벡터인 동시에 연속확률벡터라고 하자. 그렇다면 $\mathbf{P}_X$는 가산 지지집합 $A \in \mathscr{B}_n$를 가지는데, $\mu_n(A) = 0$에서 $\mathbf{P}_X(A) = 0$의 모순이 발생한다. 이번에는 X가 이산확률벡터인 동시에 특이확률벡터라 하자. 그렇다면 이건과 같이 $\mathbf{P}_X$는 가산 지지집합 $A \in \mathscr{B}_n$를 가지는데, 이의 가산개의 원소를 $x_1, x_2, \cdots$와 같이 나열하면 $\mathbf{P}_X(A) = \sum_{i=1}^{k} \mathbf{P}_X\{x_i\} = 0$에서 모순이 발생한다. (여기서 k는 유한할 수도 있고, ∞일 수도 있다.) 마지막으로 X가 연속확률벡터인 동시에 특이확률벡터라 하면 $\mathbf{P}_X \ll \mu_n$이고 $\mathbf{P}_X \perp \mu_n$이므로 명제 ? ? 의 iv로부터 $\mathbf{P}_X = 0$의 모순이 발생하고, 증명은 이로써 충분하다. □

일반적으로, 어떤 확률벡터가 이산인지, 연속인지, singular인지는 확률측도 $\mathbb{P}$와 확률벡터 X 모두에 의해 결정되는 것이지, 이 중 어느 하나에 의해 일방적으로 결정되는 것이 아니다. 즉, 둘 중 어느 하나만 보고서 X가 이산인지, 연속인지, singular인지는 알 수 없다. 이로 말미암아 우리가 확률벡터에 대해 당연하게 생각하던 사실들에 미묘한 혼란이 생겨나게 된다.

우선 X가 이산확률벡터이지만 그 치역은 가산이 아닐 수 있다. 물론, 대부분의 응용에서는 이산확률변수의 치역도 가산으로 주어지지만, 이는 우연의 일치 그 이상도 이하도 아니다. 극단적인 예시로 가측공간 $(\mathbb{R}, \mathscr{B}_1)$ 위의 측도 $\mathbb{P}$를

$$\mathbb{P} : A \mapsto \begin{cases} 1 & 0 \in A\text{인 경우} \\ 0 & \text{ow.} \end{cases}$$

로 잡아 확률공간 $(\mathbb{R}, \mathscr{B}_1, \mathbb{P})$를 구성하고 확률변수 $X : \mathbb{R} \to \mathbb{R}$를 항등함수로 두면 명백히 X는 모든 실수를 그 함숫값으로 가지고 심지어 연속이지만, 분포 $\mathbf{P}_X$가 한원소 집합 $\{0\}$을 지지집합으로 가지므로 X는 이산확률벡터이다. 다만, 정의로부터 분포가 가산 지지집합을 가지므로 X가 그 가산개의 값을 제외한 나머지 값을 가질 확률이 0이 되어 '사실상' 치역이 가산이라고 생각할 수는 있다. 하지만 영집합과 공집합이 비슷하지만 완전히 같지는 않은 것처럼 이 경우에도 '사실상' 치역이 가산인 것과 치역이 정말 가산인 것은 구분해야 할 것이다.

비슷하게, X가 연속확률벡터이지만 함수로서 연속이 아닐 수 있다. 애초에 표본공간 Ω에 위상구조가 존재한다는 보장이 없으므로 연속성을 논할 수조차 없다. 그렇다고 표본공간에 위상구조가 적당히 정의되어 있고, 이에 대해 X가 연속이라고 해서 이가 연속확률변수냐 하면 이것 또한 아니다. 앞서 든 예시를 생각해보면 표본공간이 $\mathbb{R}$이고 이 위에 표준위상을 잡더라도 X가 연속이지만 연속확률변수가 아닐 수 있다. 다만, 정의로부터 연속확률벡터와 특이확률벡터는 point mass를 가질 수 없으므로 확률이 표본공간의 어느 한 점에 집중되어 있지 않고 전체에 고르게 퍼져 있으니, 이런 의미에서 '연속'이라 생각할 수는 있다.

한편, 위의 정의에서 고등학교나 교양 통계학에서는 들어보지 못한 특이확률벡터라는 새로운 종류의 확률벡터가 등장했다. 다른 두 종류의 확률벡터는 고등학교 수준에서도 접할 수 있는 반면, 이제서야 특이확률벡터를 도입하는 것에는 그럴만한 이유가 있다. 우선 특이확률벡터는 다분히 이론적인 필요에 의한 확률벡터로 실생활의 응용에서는 거의 쓸모가 없다. 또한, 이산확률변수나 연속확률변수의 경우 기댓값이나 분산과 같은 개념의 도입과 계산이 쉬운 반면, 특이확률변수의 경우 이에 상당한 이론적 뒷받침이 필요하다. 이런 이유에서 이 책에서도 특이확률변수가 구체적인 예시로 주어지는 것은 이후에 배울 Cantor 분포 하나 뿐이다. 그렇다면 이런 단점에도 불구하고 특이확률벡터를 도입해야 할 이론적

인 필요가 대체 무엇인가? 다음 정리는 이 질문에 대한 답이자 측도론의 에필로그에서 예고한 이번 절의 클라이막스이다.

Theorem 2.26 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$ 위의 rv. $X : \Omega \to \mathbb{R}^n$에 대해 다음의 조건

i. Rv. $X_{\mathrm{ac}} : \mathbb{R}^n \to \mathbb{R}^n$는 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\mathrm{ac}})$에서 정의된 연속확률벡터이다.
ii. Rv. $X_{\mathrm{pp}} : \mathbb{R}^n \to \mathbb{R}^n$는 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\mathrm{pp}})$에서 정의된 이산확률벡터이다.
iii. Rv. $X_{\mathrm{sc}} : \mathbb{R}^n \to \mathbb{R}^n$는 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\mathrm{sc}})$에서 정의된 특이확률벡터이다.

를 만족하는 적당한 $\mathscr{B}_n$ 위의 확률측도 $\mathbb{P}_{\mathrm{ac}}, \mathbb{P}_{\mathrm{pp}}, \mathbb{P}_{\mathrm{sc}}$와 rv. $X_{\mathrm{ac}}, X_{\mathrm{pp}}, X_{\mathrm{sc}}$가 존재하여 $\alpha + \beta + \gamma = 1$인 적당한 $\alpha, \beta, \gamma \geq 0$에 대해 $\mathbf{P}_X = \alpha \mathbf{P}_{X_{\mathrm{ac}}} + \beta \mathbf{P}_{X_{\mathrm{pp}}} + \gamma \mathbf{P}_{X_{\mathrm{sc}}}$이다.

Proof 정리 ? ? 로부터 $(\mathbb{R}^n, \mathscr{B}_n, \mathbf{P}_X)$가 확률공간이므로 Lebesgue의 분해정리로부터 $\mathbf{P}_X$는 절대연속성분 $(\mathbf{P}_X)_{\mathrm{ac}}$, 순수 점 성분 $(\mathbf{P}_X)_{\mathrm{pp}}$, 특이연속성분 $(\mathbf{P}_X)_{\mathrm{sc}}$에 대해 $\mathbf{P}_X = (\mathbf{P}_X)_{\mathrm{ac}} + (\mathbf{P}_X)_{\mathrm{pp}} + (\mathbf{P}_X)_{\mathrm{cs}}$와 같이 분해된다. 또한, $\mathbf{P}_X(\mathbb{R}^n) = 1$이므로 $\alpha = (\mathbf{P}_X)_{\mathrm{ac}}(\mathbb{R}^n), \beta = (\mathbf{P}_X)_{\mathrm{pp}}(\mathbb{R}^n), \gamma = (\mathbf{P}_X)_{\mathrm{cs}}(\mathbb{R}^n)$라 하면 α, β, γ는 모두 유한하고 $\alpha + \beta + \gamma = 1$이다. 이제 α, β, γ가 모두 0이 아닌 특별한 경우를 생각해보자. 그렇다면 $\mathbb{P}_1 := (\mathbf{P}_X)_{\mathrm{ac}}/\alpha, \mathbb{P}_2 := (\mathbf{P}_X)_{\mathrm{pp}}/\beta, \mathbb{P}_3 := (\mathbf{P}_X)_{\mathrm{cs}}/\gamma$가 모두 $\mathscr{B}_n$ 위의 확률측도이므로 $(\mathbf{P}_X)_{\mathrm{ac}}, (\mathbf{P}_X)_{\mathrm{pp}}, (\mathbf{P}_X)_{\mathrm{cs}}$의 성질과 정리 ? ? 로부터 적당한 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\mathrm{ac}})$에서 정의된 연속확률벡터 $X_{\mathrm{ac}} : \mathbb{R}^n \to \mathbb{R}^n$, 적당한 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\mathrm{pp}})$에서 정의된 이산확률벡터 $X_{\mathrm{pp}} : \mathbb{R}^n \to \mathbb{R}^n$, 적당한 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\mathrm{sc}})$에서 정의된 특이확률벡터 $X_{\mathrm{sc}} : \mathbb{R}^n \to \mathbb{R}^n$가 존재하여 $\mathbb{P}_1 = \mathbf{P}_{X_{\mathrm{ac}}}, \mathbb{P}_2 = \mathbf{P}_{X_{\mathrm{pp}}}, \mathbb{P}_3 = \mathbf{P}_{X_{\mathrm{sc}}}$이고, 곧 $\mathbf{P}_X = (\mathbf{P}_X)_{\mathrm{ac}} + (\mathbf{P}_X)_{\mathrm{pp}} + (\mathbf{P}_X)_{\mathrm{cs}} = \alpha\mathbb{P}_1 + \beta\mathbb{P}_2 + \gamma\mathbb{P}_3 = \alpha \mathbf{P}_{X_{\mathrm{ac}}} + \beta \mathbf{P}_{X_{\mathrm{pp}}} + \gamma \mathbf{P}_{X_{\mathrm{sc}}}$이다. 한편, α, β, γ 중 일부가 0인 경우에 대해서도 이와 비슷하게 하면 된다. □

연속확률벡터, 이산확률벡터, 특이확률벡터의 세 가지 분류가 서로 배타적인 관계인 것은 맞지만 그렇다고 임의의 확률벡터가 반드시 이 셋 중 하나에 속하는 것은 아니다. 즉, 확률벡터 중에는 연속도, 이산도, singular도 아닌 골치아픈 것들이 존재한다. (이런 확률벡터를 흔히 **mixed type**이라 부르며 특이확률벡터에 비할 바는 아니지만 그 실용성은 많이 떨어지는 편이다.) 이런 상황에서 위의 정리는 임의의 확률벡터에 대해 비록 이가 mixed type이더라도 그 분포는 적당한 연속확률벡터, 이산확률벡터, 특이확률벡터의 분포의 합으로 분해할 수 있다는 놀라운 결과를 함의한다. 곧 연속확률벡터, 이산확률벡터, 특이확률벡터의 세 가지 분류는 확률벡터의 공간의 기저와 비슷한 역할을 하며, 이 세 가지 확률벡터를 정의한 순간 사실상 모든 확률벡터의 분류를 끝마친 것과 다름없다.

따라서 이론 전개에 있어서는 mixed type rv.를 고려할 필요 없이 연속확률벡터, 이산확률벡터, 특이확률벡터의 세 가지 확률벡터만 생각하면 되고, mixed type 확률벡터는 이 세 종류의 확률벡터의 성질들을 적당히 섞어 가질 뿐이다. 이러니 고등학교 시절부터 연속확률벡터, 이산확률벡터의 두 가지 종류에만 지대한 관심을 가진 것은 너무나 당연하다. 이

론적으로 다루기 힘든 특이확률벡터를 제외하면 이 둘을 다룸으로써 우리는 고등학교때부터 우리도 모르는 사이에 사실상 온갖 종류의 확률벡터를 모두 다루고 있던 셈이다!

이제 클라이막스의 여운을 뒤로 하고, CDF를 살펴볼 순서이다. 흔히 교양 통계학에서는 PDF를 배운 뒤 CDF를 배우므로 PDF의 개념을 도입하지도 않고 CDF를 정의하는 것이 의아할 수 있다. 하지만, 적어도 이론적으로는 CDF가 PDF보다 더 기본적인 개념이기에 이를 먼저 도입한다.

Definition 2.27 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 rv. X의 **(누적)분포함수((cumulative) distribution function)**를 $F_X : \mathbb{R}^n \to \mathbb{R}$로 쓰고 $F_X : x \mapsto \mathbf{P}_X(\prod_{i=1}^n(-\infty, x_i])$로 정의한다. 특별히, $n \geq 2$인 경우 F_X를 rv. $X_1, \cdots, X_n$의 **결합(누적)분포함수(joint (cumulative) distribution function)**라 하기도 한다.

CDF의 기본적인 성질은 정리 ? ? 로부터 대부분 자명하게 유도된다.

Theorem 2.28 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 다음이 성립한다.

i. $0 \leq F_X \leq 1$.
ii. 임의의 유계인 semi-open box $B \subseteq \mathbb{R}^n$에 대해 $\Delta_B F_X = \mathbf{P}_X(B) \geq 0$이다.
iii. CDF F_X는 각 변수에 대해 증가한다.
iv. CDF F_X는 오른쪽 연속이다.
v. 각 $i \leq n$에 대해 $\lim_{x_i \to -\infty} F_X(x) = 0$이다.[1]
vi. $\lim_{x_1, \cdots, x_n \to \infty} F_X(x) = 1$.[2]
vii. 임의의 $x \in \mathbb{R}^n$에 대해 $B_x = (-\infty, x]$라 하면 $F_X(x-) = \mathbf{P}_X(B_x^\circ)$이고 $F_X(x) - F_X(x-) = \mathbf{P}_X(\partial B_x)$이다.

Proof i − vi. 이는 CDF의 정의와 정리 ? ? 로부터 자명하다.

vii. 임의의 $x \in \mathbb{R}^n$를 택하여 집합열 $\{B_j\}$를 $B_j := B_{x-\mathbf{1}/j}$로 두면 이는 $\mathscr{S}_n$에 속하는 증가하는 집합열로서 $B_j \uparrow (-\infty, x) = B_x^\circ$이다. 따라서 $F_X(x - \mathbf{1}/j) = \mathbf{P}_X(B_j) \uparrow \mathbf{P}_X(B_x^\circ)$이므로 임의의 $\varepsilon > 0$을 택하면 적당한 $j_0 \in \mathbb{N}$가 존재하여 $\mathbf{P}_X(B_x^\circ) - \mathbf{P}_X(B_{j_0}) < \varepsilon$이다. 이제 $\delta = 1/j_0$라 하면 $||x - y|| < \delta$이고 $x > y$인 모든 $y \in \mathbb{R}^n$에 대해 $B_{j_0} \subseteq (-\infty, y) \subseteq B_x^\circ$에서 $\mathbf{P}_X(B_x^\circ) - \varepsilon < \mathbf{P}_X(B_{j_0}) \leq F_X(y) = \mathbf{P}_X((-\infty, y]) \leq \mathbf{P}_X(B_x^\circ)$이므로 $|F_X(y) - \mathbf{P}_X(B_x^\circ)| < \varepsilon$가 되어 $F_X(x-) = \mathbf{P}_X(B_x^\circ)$임을 안다. 이제 $F_X(x) - F_X(x-) = \mathbf{P}_X(B_x) - \mathbf{P}_X(B_x^\circ) = \mathbf{P}_X(\partial B_x)$임은 자명하다. □

위의 정리에서 $n = 1$인 경우 vii는 $F_X(x) - F_X(x-) = \mathbf{P}_X\{x\}$가 되어 이로써 CDF를 분포의 point mass를 구하는 용도로 사용할 수 있다. 한편, 위의 정리의 역 비슷한 정리도 성립한다.

Theorem 2.29 함수 $F : \mathbb{R}^n \to \mathbb{R}$에 대해 이가

i. 함수 F는 오른쪽 연속이고 각 변수에 대해 증가한다.
ii. 유계인 semi-open box $B \subseteq \mathbb{R}^n$에 대해 $\Delta_B F \geq 0$이다.
iii. 각 $i \leq n$에 대해 $\lim_{x_i \to -\infty} F(x) = 0$이고 $\lim_{x_1, \cdots, x_n \to \infty} F_X(x) = 1$이다.

를 만족하면 적당한 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$가 존재하여 $F = F_X$이다. 특별히, 이때 $(\Omega, \mathscr{F}) = (\mathbb{R}^n, \mathscr{B}_n)$이도록 잡을 수 있다.

Proof 정리 ? ? 로부터 적당한 $\mathscr{B}_n$ 위의 측도 μ가 존재하여 임의의 $x \in \mathbb{R}^n$에 대해 $F(x) = \mu((-\infty, x])$이다. 그런데 정리 ? ? 의 vii와 주어진 조건으로부터 $\mu(\mathbb{R}^n) = 1$이 되어 μ는 확률측도이고, 곧 정리 ? ? 로부터 적당한 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$ 위의 n차원 rv. X가 존재하여 $\mu = \mathbf{P}_X$이다. 이상으로부터 임의의 $x \in \mathbb{R}^n$에 대해 $F(x) = \mu((-\infty, x]) = \mathbf{P}_X((-\infty, x]) = F_X(x)$가 성립한다. 한편, 이때 $(\Omega, \mathscr{F}) = (\mathbb{R}^n, \mathscr{B}_n)$이도록 잡을 수 있음은 정리 ? ? 로부터 자명하다. □

곧 우리는 CDF가 될 수 있는 함수를 위의 정리의 조건 i – iii으로 완벽히 characterize할 수 있다. 한편, CDF의 연속성에 대한 다음 정리들도 꽤나 흥미롭다.

Theorem 2.30 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 한 점 $x_0 \in \mathbb{R}^n$ 대해 $B_{x_0} = (-\infty, x_0]$라 하면 TFAE.

i. CDF F_X가 x_0에서 연속이다.
ii. $F_X(x_0) = \mathbf{P}_X(B_{x_0}) = \mathbf{P}_X(B_{x_0}^\circ)$.
iii. $\mathbf{P}_X(\partial B_{x_0}) = 0$.

Proof i $\Rightarrow$ ii. 집합열 $\{B_j\}$를 $B_j = (-\infty, x_0 - \mathbf{1}/j]$로 두면 이는 $B_{x_0}^\circ$로 수렴하는 증가하는 집합열이므로 $F_X(x_0 - \mathbf{1}/j) = \mathbf{P}_X(B_j) \uparrow \mathbf{P}_X(B_{x_0}^\circ)$이다. 한편, 가정으로부터 F_X가 x_0에서 연속이므로 $\mathbf{P}_X(B_j) = F_X(x_0 - \mathbf{1}/j) \uparrow F_X(x_0) = \mathbf{P}_X(B_{x_0})$가 되어 $F_X(x_0) = \mathbf{P}_X(B_{x_0}) = \mathbf{P}_X(B_{x_0}^\circ)$임을 안다.

ii $\Rightarrow$ iii. 이는 $\mathbf{P}_X(\partial B_{x_0}) = \mathbf{P}_X(B_{x_0}) - \mathbf{P}_X(B_{x_0}^\circ) = 0$에서 자명하다.

iii $\Rightarrow$ i. 0으로 수렴하는 $\mathbb{R}^n$에 속하는 임의의 수열 $\{h_j\}$를 택하여 각 $j \in \mathbb{N}$에 대해 $a_j = \min_{i=1}^n h_j^i, b_j = \max_{i=1}^n h_j^i$라 하면 $F_X(x_0 + a_j\mathbf{1}) \leq F_X(x_0 + h_j) \leq F_X(x_0 + b_j\mathbf{1})$이고 $a_j, b_j \to 0$이다. 따라서 함수 $f : \mathbb{R} \to \mathbb{R}$를 $f : x \mapsto F_X(x_0 + x\mathbf{1})$로 두고 이가 0에서 연속임을 보이는 것으로 증명은 충분한데, CDF의 성질로부터 $\lim_{x \downarrow 0} f(x) = f(0)$임은 이미 알고 있으므로 $\lim_{x \uparrow 0} f(x) = f(0)$이라는 사실만 보이면 된다. 이를 위해 $x_j \uparrow 0$인 임의의 실수열 $\{x_j\}$를 택하여 집합열 $\{B_j\}$를 $B_j := (-\infty, x_0 + x_j]$로 두면 이는 $B_{x_0}^\circ$로 수렴하는 증가하는 집합열이다. 그렇다면 가정으로부터 $f(x_j) = F_X(x_0 + x_j\mathbf{1}) = \mathbf{P}_X(B_j) \uparrow \mathbf{P}_X(B_{x_0}^\circ) = \mathbf{P}_X(B_{x_0}) = F_X(x_0) = f(0)$이고, 증명이 끝난다. □

Lemma 2.31 단조함수 $f : \mathbb{R} \to \mathbb{R}$는 가산개의 불연속점만을 가진다.

Proof 우선 f가 증가함수라 하고 집합 $A = \{x \in \mathbb{R} : f$가 x에서 불연속$\}$를 생각하자. 그렇다면 임의의 $x \in A$에 대해 $f(x-) < f(x+)$이므로 $f(x-) < p_x < f(x+)$인 적당한 $p_x \in \mathbb{Q}$를 택할 수 있고, 이로써 함수 $g : A \to \mathbb{Q}$를 $g : x \mapsto p_x$로 두자. 한편, 임의의 $x, y \in A$에 대해 $x < y$인데 $f(y-) < f(x+)$이면 $z = (x+y)/2$에 대해 $f(z) \le f(y-) < f(x+) \le f(z)$의 모순이 발생하므로 $(f(x-), f(x+))$와 $(f(y-), f(y+))$는 서로소가 되어 g는 단사이고, 곧 A는 가산이다. 이제 f가 감소함수인 경우에는 $-f$를 대신 생각하면 된다. □

Theorem 2.32 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 F_X의 연속점의 집합은 조밀하다. 나아가, $n = 1$인 경우 F_X는 가산개의 불연속점만을 가진다.

Proof 임의의 $x_0 \in \mathbb{R}^n$를 택하여 함수 $f : \mathbb{R} \to [0, 1]$를 $f : h \mapsto F_X(x_0 + h\mathbf{1})$로 두자. 그렇다면 f는 증가함수가 되어 위의 보조정리로부터 가산개의 불연속점만을 가지고, 곧 0으로 수렴하면서 각 점에서 f가 연속인 실수열 $\{h_i\}$를 적당히 택할 수 있다. 이는 각 $i \in \mathbb{N}$에 대해 $\mathbf{P}_X((-\infty, x_0 + h)) = F_X((x_0 + h_i\mathbf{1})-) = f(h_i-) = f(h_i) = F_X(x_0 + h_i\mathbf{1})$임을 함의하므로 정리 ? ? 로부터 F_X는 $x_0 + h_i\mathbf{1}$에서 연속이고, $x_0 + h_i\mathbf{1} \to x_0$임은 자명하므로 F_X의 연속점의 집합이 조밀함을 안다. 한편, $n = 1$인 경우에는 F_X가 증가함수이므로 다시 위의 보조정리로부터 가산개의 불연속점만을 가짐이 자명하다. □

이제 PDF를 도입하는 것으로 이번 절을 마무리하자. 측도론으로 PDF를 도입하는데 핵심적인 역할을 하는 것은 다름아닌 Radon-Nikodým 도함수의 개념이다.

Definition 2.33 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 적당한 $\mathscr{B}_n$ 위의 $\sigma-$유한 측도 μ가 존재하여 $\mathbf{P}_X \ll \mu$라 하자. 이때 Radon-Nikodým 도함수 $d\mathbf{P}_X/d\mu$를 X의 μ에 대한 **(확률)밀도함수((probability) density function)**라 하고 f_X로 쓴다. 특별히, $n \ge 2$인 경우 f_X를 rv. $X_1, \cdots, X_n$의 **결합(확률)밀도함수(joint (probability) density function)**라 하기도 한다.

이렇게 Radon-Nikodým 도함수로 PDF를 정의하면서 이번에도 우리가 PDF에 대해 당연하게 생각하던 사실들에 미묘한 혼란이 생겨나게 된다. 우선, Radon-Nikodým 도함수가 유일하기는 하지만 그 유일성이 '거의 어디서나 같은 함수를 하나로 볼 때' 성립하는 것이므로 엄밀히는 유일하지 않다. 즉, 영집합에서 조금씩 다른 무한히 많은 함수들이 모두 하나의 확률벡터의 PDF가 될 수 있으므로 일반적으로 특정 점에서의 PDF의 값을 물어보는 것은 의미가 없다. 그러나 Radon-Nikodým 도함수를 추상적인 도구로만 사용했던 측도론에서와 달리 확률론에서는 PDF를 구체적인 함수로 다루어야 할 필요가 있으므로 혼란을 피하기 위해 PDF가 될 수 있는 무한히 많은, 거의 어디서나 같은 함수 중 특정한 하나를 PDF의 **version**이라 한다. 물론, 표기상의 편의를 이유로 논의의 대상이 PDF인지, PDF의 version인지를 명시적으로 밝히지 않는 경우가 대부분이지만, L^p 공간에서 동치류로서의 f와 구

체적인 함수 f를 논의의 맥락으로 큰 혼란 없이 구분하였듯이 이 둘 또한 구분에 큰 어려움은 없을 것이다.

다음으로, 지금까지는 연속확률벡터와 이산확률벡터에 대해서만 PDF를 생각했지만 위의 정의에서 볼 수 있듯이 임의의 확률벡터 X에 대해서도 $\mathbf{P}_X \ll \mu$인 $\sigma-$유한 측도 μ만 잘 잡아주면 이의 PDF를 생각할 수 있다. 한편, 연속확률벡터, 이산확률벡터, 특이확률벡터의 세 가지 확률벡터만 생각해도 충분하다는 사실에 놀라워했던 것이 불과 몇 페이지 전인데 굳이 이렇게 지나칠 정도로 일반적인 PDF를 도입하는 것에 의아할 수 있다. 그러나 이는 mixed-type 확률벡터의 PDF를 직접 다루기 위함이 아니라 연속확률벡터와 이산확률벡터의 PDF를 동시에 다루기 위함이다. 고등학교나 교양 통계학에서 연속확률변수와 이산확률변수의 기댓값이나 분산 같은 성질들을 논할 때, 각각 경우를 나누어 전자는 적분으로, 후자는 합으로 접근했던 것을 기억할 것이다. 우리는 위에서 일반적으로 정의된 PDF와 측도론의 에필로그에서 소개한 셈측도를 이용해 드디어 이런 번거로움에서 벗어날 수 있다.

한편, 정의에서 볼 수 있듯이 $\mathbf{P}_X \ll \mu$인 $\sigma-$유한 측도 μ를 무엇으로 택하는지에 따라 그 PDF가 달라지므로 μ를 반드시 명시해 주어야 하는데, 우리가 주로 다룰 연속확률벡터와 이산확률벡터의 경우 이에 대한 관례가 있다. 먼저 연속확률벡터 $X : \Omega \to \mathbb{R}^n$의 경우 그 정의상 $\mathbf{P}_X \ll \mu_n$이 항상 성립하므로 특별한 언급이 없는 이상 연속확률벡터의 PDF는 μ_n에 대한 PDF로 생각한다. 이산확률벡터 $X : \Omega \to \mathbb{R}^n$의 경우는 조금 복잡한데, 우선 정의상 $\mathbf{P}_X$는 가산 지지집합 $A \in \mathscr{B}_n$를 가진다. 물론, 이때의 가산 지지집합은 유일하지 않지만 임의의 $x \in A$에 대해 $\mathbf{P}_X\{x\} > 0$이라는 조건을 추가하면 유일하게 주어진다. 그렇다면 $(\mathbf{P}_X, \mathscr{B}_n, \#)$에 대해 $\mathbf{P}_X \ll \#_A$임이 자명하고, 이때 $\#_A$가 $\sigma-$유한임도 자명하므로 특별한 언급이 없는 이상 이산확률벡터의 PDF는 $\#_A$에 대한 PDF로 생각한다. 나아가, 이산확률벡터의 경우 특별한 언급이 없는 이상 PDF의 version으로 $\mathbb{R}^n \setminus A$에서 0인 version을 택하는 관례도 있다. 이런 관례를 따르면 이때의 PDF가 우리가 기존에 알던 PMF가 된다는 것을 잠시 후에 알게 될 것이다.

이제 PDF의 기본적인 성질을 보자. 이는 대부분 적분의 성질로부터 거의 자명하게 얻어진다.

Theorem 2.34 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 적당한 $\mathscr{B}_n$ 위의 $\sigma-$유한 측도 μ가 존재하여 $\mathbf{P}_X \ll \mu$라 하자. 그렇다면 μ에 대한 X의 PDF f_X에 대해 다음이 성립한다.

i. $f_X \geq 0$ ($\mu-$ae.).
ii. $\int_{\mathbb{R}^n} f_X \, d\mu = 1$.
iii. 임의의 $A \in \mathscr{B}_n$에 대해 $\int_A f_X \, d\mu = \mathbf{P}_X(A) = \mathbb{P}\{X \in A\}$이다.

Proof iii. 이는 $\int_A f_X \, d\mu = \int_A (d\mathbf{P}_X / d\mu) \, d\mu = \int_A d\mathbf{P}_X = \mathbf{P}_X(A)$에서 자명하다.

i. 만약 $f_X \geq 0$ (μ−ae.)가 아니라면 집합 $A = \{x \in \mathbb{R}^n : -f_X(x) > 0\}$이 양의 측도를 가지므로 iii과 따름정리 ? ? 의 iii으로부터 $\mathbf{P}_X(A) = \int_A f_X \, d\mu < 0$의 모순이 발생한다. 따라서 $f_X \geq 0$ (μ−ae.)이어야 한다.

ii. iii으로부터 $\int_{\mathbb{R}^n} f_X \, d\mu = \mathbf{P}_X(\mathbb{R}^n) = 1$이므로 자명하다. □

특별히, 이산확률벡터의 경우 위의 정리는 다음 따름정리와 같이 합의 형태로 표현된다.

Corollary 2.35 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 이산확률벡터 $X : \Omega \to \mathbb{R}^n$에 대해 $A \in \mathscr{B}_n$를 $\mathbf{P}_X$의 가산 지지집합이라 하고 임의의 $x \in A$에 대해 $\mathbf{P}_X\{x\} > 0$이라 하면 다음이 성립한다.

i. 임의의 $x \in \mathbb{R}^n$에 대해 $f_X(x) = \mathbf{P}_X\{x\}$이고, 따라서 $0 \leq f_X \leq 1$이다.
ii. $\sum_{x \in \mathbb{R}^n} f_X(x) = 1$.
iii. 임의의 $B \in \mathscr{B}_n$에 대해 $\sum_{x \in B} f_X(x) = \mathbf{P}_X(B) = \mathbb{P}\{X \in B\}$이다.

Proof 표기의 편의를 위해 $\mathscr{A} = \mathscr{B}_n|_A$라 하면 $\#_A|_{\mathscr{A}} = \#|_{\mathscr{A}}$는 가측공간 $(A, \mathscr{A})$ 위의 셈측도이다.

i. 만약 $x \in A$이면 정리 ? ? 의 iii으로부터 $\mathbf{P}_X\{x\} = \int_{\{x\}} f_X \, d\#_A = \int_{\{x\}} f_X|_A \, d\#|_{\mathscr{A}} = f_X(x)$이고 $x \notin A$이면 f_X의 version을 택하는 관례로부터 $\mathbf{P}_X\{x\} = 0 = f_X(x)$이다.

ii. 위의 정리와 셈측도의 성질 그리고 f_X의 version을 택하는 관례와 정리 ? ? 의 iii으로부터 $1 = \mathbf{P}_X(A) = \int_A f_X \, d\#_A = \int_A f_X|_A \, d\#|_{\mathscr{A}} = \sum_{x \in A} f_X(x) = \sum_{x \in \mathbb{R}^n} f_X(x)$이다.

iii. 위의 정리와 셈측도의 성질 그리고 f_X의 version을 택하는 관례와 정리 ? ? 의 iii으로부터 $\mathbf{P}_X(B) = \mathbf{P}_X(A \cap B) = \int_{A \cap B} f_X \, d\#_A = \int_{A \cap B} f_X|_A \, d\#|_{\mathscr{A}} = \sum_{x \in A \cap B} f_X(x) = \sum_{x \in B} f_X(x)$이다. □

다음 정리는 연속확률벡터의 CDF와 PDF의 관계에 대한 정리로서 연속확률벡터의 CDF와 PDF를 계산하는 데 유용하게 사용된다.

Theorem 2.36 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 연속확률벡터 $X : \Omega \to \mathbb{R}^n$에 대해 다음이 성립한다.

i. $f_X(x) = F_X^{(s)}(x)$. [3]
ii. $F_X(x) = \int_{(-\infty, x]} f_X \, d\mu_n$.

특별히, $x_0 \in \mathbb{R}^n$의 근방에서 $\partial^{\mathbf{1}} F_X$가 존재하고 이가 x_0에서 연속이면 $f_X(x_0) = \partial^{\mathbf{1}} F_X(x_0)$이다.

Proof 이는 따름정리 ? ? 와 정리 ? ? 의 iii으로부터 자명하다 □

비슷하게, 다음 정리는 어떤 확률벡터가 연속인지를 판단할 때 유용하다.

Theorem 2.37 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 TFAE.

i. Rv. X는 연속확률벡터이다.
ii. CDF F_X가 절대연속이다.
iii. 적분가능한 Borel 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$가 존재하여 $F_X(x) = \int_{(-\infty,x]} f\,d\mu_n$이다.

PROOF i $\Leftrightarrow$ ii. 이는 정리 ? ? 로부터 자명하다.

i $\Rightarrow$ iii. 이는 $f = f_X$로 두고 정리 ? ? 의 ii를 생각하면 자명하다.

iii $\Rightarrow$ ii. 함수 f가 적분가능하므로 거의 어디서나 유한하고, 따라서 WLOG, 필요다하면 영집합에서의 f의 값을 0으로 바꾸어 $|f| < \infty$라 하자. 이제 집합열 $\{A_j\}$를 $A_j = |f|^{-1}((j,\infty))$로 두면 이는 $\emptyset$으로 수렴하는 감소하는 집합열이 되어 $\int_{A_j} |f|\,d\lambda_n \to \int_{\emptyset} |f|\,d\lambda_n = 0$이다. 이로부터 임의의 $\varepsilon > 0$을 택하면 적당한 $j_0 \in \mathbb{N}$가 존재하여 $\int_{A_{j_0}} |f|\,d\lambda_n < \varepsilon/2$이고, $\mu_n(A) < \varepsilon/2j_0 =: \delta$인 임의의 $A \in \mathscr{B}_n$에 대해 $\int_A |f|\,d\mu_n \le \int_{A\setminus A_{j_0}} |f|\,d\mu_n + \int_{A_{j_0}} |f|\,d\mu_n < \mu_n(A \setminus A_{j_0})j_0 + \varepsilon/2 < \varepsilon$이다. 따라서 임의의 유계이고 서로소인 $B_1, \cdots, B_l \in \mathscr{S}_n$에 대해 $\sum_{k=1}^l \mu_n(B_k) = \mu_n(\bigsqcup_{k=1}^l B_k) < \delta$이면 $\sum_{k=1}^l \Delta_{B_k} F_X = \sum_{k=1}^l \int_{B_k} f\,d\mu_n = \int_{\bigsqcup_{k=1}^l B_k} f\,d\mu < \varepsilon$이 되어 F_X가 절대연속임을 안다. □

2.3 Expectation

이번 절에서는 확률변수의 대표적인 통계량 중 하나인 기댓값과 분산을 엄밀하게 도입하고, 이와 관련된 부등식을 증명한다. 통계학의 큰 연구주제 중 하나는 '정보의 단순화'라 할 것이다. 당연히, 이때 단순화의 정도와 이에 따른 정보의 손실은 서로 trade-off의 관계에 있어서 단순화를 많이 하면 할수록 정보의 전달이 용이하지만 그만큼 손실되는 정보도 많아지고, 역으로 손실되는 정보의 양을 줄이면 줄일수록 단순화의 정도가 감소하여 전달이 어려워진다. 이런 상황에서 우리는 정보의 손실을 최소한으로 유지하면서 정보를 단순하게 전달할 수 있는 좋은 방법을 찾고자 하는데, 이번 절에서 알아볼 기댓값은 어떤 분포의 정보를 단 하나의 통계량으로 단순화하여 전달할 수 있는 좋은 방법 중 하나이다. (이와 같이 분포의 정보를 요약하는 통계량을 그 분포의 대푯값이라 하며 기댓값 외에도 중앙값, 최빈값 등이 있다. 각 대푯값은 나름의 장단점이 있기에 어떤 상황에서 무엇을 써야 하는지를 잘 판단하여 사용해야 한다.)

기댓값은 확률변수가 가질 것으로 가장 기대할 수 있는 값이라 해석할 수 있다. 기댓값은 분포의 여러 대푯값 중에서 가장 널리 이용되는 통계량으로 추정이나 관련 가설의 검정 등에 있어 많은 장점을 지니지만, outlier와 같은 극단치에 민감하여 때로는 분포에 대해 다소 오인의 가능성이 있는 정보를 전달할 수 있다는 단점을 가진다. 이러한 기댓값은 기본적으로 확률변수의 적분으로 정의된다.

Definition 2.38 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 만약 X가 적분가능하다면 $\int_\Omega X \, d\mathbb{P}$를 X의 **기댓값(expectation)** 혹은 **평균(mean)**이라 하고 $\mathbf{E}(X)$, μ_X 혹은 간단히 μ로 쓴다.

기댓값이 정의되기 위해서는 확률공간과 그 위에서 정의된 rv.만 있으면 충분하다는 점에 주목하기 바란다. 고등학교나 교양 통계학에서는 기댓값이 정의되려면 확률변수가 반드시 연속이거나 이산이어서 PDF나 PMF가 있어야 했다. 그러나 위의 정의는 일반적인 mixed-type 확률변수에 대해서도 기댓값을 정의할 수 있도록 해준다. 한편, 정의에서 볼 수 있듯이 기댓값은 확률변수의 적분에 불과하므로 이에 관한 기본적인 성질은 대부분 적분의 성질로부터 자명하게 유도된다.

Proposition 2.39 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 $\mathbf{E}(X)$가 존재할 필요충분조건은 $\mathbf{E}(|X|)$가 존재하는 것이다.

Proof 이는 X가 적분가능할 필요충분조건이 $|X|$가 적분가능할 것이라는 점에서 자명하다. $\square$

Theorem 2.40 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X, Y : \Omega \to \mathbb{R}$에 대해 $\mathbf{E}(X), \mathbf{E}(Y)$가 모두 존재한다면 다음이 성립한다.

i. (선형성) 임의의 $a, b \in \mathbb{R}$에 대해 $\mathbf{E}(aX + bY)$가 존재하고 $\mathbf{E}(aX + bY) = a\mathbf{E}(X) + b\mathbf{E}(Y)$이다.
ii. 만약 $X = Y$ (as.)라면 $\mathbf{E}(X) = \mathbf{E}(Y)$이다.
iii. 만약 $X \le Y$ (as.)라면 $\mathbf{E}(X) \le \mathbf{E}(Y)$이다.
iv. $\mathbf{E}(1) = 1$.
v. $\inf_{\omega \in \Omega} X(\omega) \le \mathbf{E}(X) \le \sup_{\omega \in \Omega} X(\omega)$.
vi. $|\mathbf{E}(X)| \le \mathbf{E}(|X|)$.

Proof i − iii, vi. 이는 적분의 성질로부터 자명하다.
iv. 이는 $\mathbf{E}(1) = \int_\Omega d\mathbb{P} = \mathbb{P}(\Omega) = 1$에서 분명하다.
v. 이는 iii과 iv로부터 자명하다. $\square$

기댓값의 선형성으로부터 $X - \mu_X$의 기댓값은 항상 0이다. 이렇게 원래의 확률변수에서 그 기댓값을 빼는 것을 **centering**이라 하며, 기댓값이 서로다른 분포를 비교할 때에 자주 사용되는 방법이다. 한편, 다양한 수렴정리들도 동일하게 성립한다.

Theorem 2.41 (Monotone convergence theorem) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 음이 아닌 rv. $X_i : \Omega \to \mathbb{R}_0^+$의 열 $\{X_i\}$와 음이 아닌 rv. $X : \Omega \to \mathbb{R}_0^+$에 대해 $X_i \uparrow X$ (as.)이고 $\mathbf{E}(X)$가 존재한다고 하면 모든 $i \in \mathbb{N}$에 대해 $\mathbf{E}(X_i)$도 존재하고 $\mathbf{E}(X_i) \uparrow \mathbf{E}(X)$이다.

Proof 모든 $i \in \mathbb{N}$에 대해 $0 \le X_i \le X$이므로 $\mathbf{E}(X_i)$가 존재한다. 이제 나머지는 측도론에서 배운 MCT에서 자명하다. □

Theorem 2.42 (Lebesgue's dominated convergence theorem) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X_i : \Omega \to \mathbb{R}$의 열 $\{X_i\}$와 rv. $X : \Omega \to \mathbb{R}$에 대해 $X_i \to X$ (as.)라 하자. 또한 어떤 rv. $Y : \Omega \to \mathbb{R}_0^+$가 존재하여 $\mathbf{E}(Y)$가 존재하고 모든 $i \in \mathbb{N}$에 대해 $|X_i| \le Y$이면 $\mathbf{E}(X_i)$와 $\mathbf{E}(X)$가 모두 존재하고 $\mathbf{E}(X_i) \to \mathbf{E}(X)$이다.

Proof 이는 측도론에서 배운 DCT에서 자명하다. □

Corollary 2.43 (Bounded convergence theorem) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X_i : \Omega \to \mathbb{R}$의 열 $\{X_i\}$와 rv. $X : \Omega \to \mathbb{R}$에 대해 $X_i \to X$ (as.)라 하자. 또한 $\{X_i\}$가 균등하게 유계라 하자. 즉, 어떤 $M > 0$이 존재하여 모든 $i \in \mathbb{N}$에 대해 $|X_i| \le M$이라 하자. 그렇다면 $\mathbf{E}(X_i)$와 $\mathbf{E}(X)$가 모두 존재하고 $\mathbf{E}(X_i) \to \mathbf{E}(X)$이다.

Proof 이는 측도론에서 배운 유계수렴정리로부터 자명하다. □

비록 기댓값의 정의는 충분히 일반적이지만, 이가 기댓값을 계산하는 구체적인 방법을 알려주지는 않는다. 결국 기댓값을 계산하기 위해서는 우리가 여태껏 해왔던 것처럼 PDF의 도움을 받아야 한다.

Theorem 2.44 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 Borel 함수 $g : \mathbb{R}^n \to \mathbb{R}$에 대해 $\mathbf{E}(g(X))$가 존재한다면(혹은 g가 $\mathbf{P}_X$−적분가능하다면) $\mathbf{E}(g(X)) = \int_{\mathbb{R}^n} g\, d\mathbf{P}_X$이고, 이때 g는 $\mathbf{P}_X$−적분가능하다(혹은 $\mathbf{E}(g(X))$가 존재한다).

Proof 기댓값 $\mathbf{E}(g(X))$가 존재한다면 $\mathbf{E}(g(X)) = \int_\Omega g \circ X\, d\mathbb{P} = \int_{\mathbb{R}^n} g\, d\mathbf{P}_X < \infty$이므로 이는 자명하다. 한편, $\mathbf{P}_X$−적분가능한 g에 대해서도 비슷하게 하면 된다. □

Theorem 2.45 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 Borel 함수 $g : \mathbb{R}^n \to \mathbb{R}$에 대해 적당한 $\mathscr{B}_n$ 위의 σ−유한 측도 μ가 존재하여 $\mathbf{P}_X \ll \mu$라 하자. 만약 $\mathbf{E}(g(X))$가 존재한다면(혹은 $f_X g$는 μ−적분가능하다면) $\mathbf{E}(g(X)) = \int_{\mathbb{R}^n} f_X g\, d\mu$이고, 이때 $f_X g$는 μ−적분가능하다(혹은 $\mathbf{E}(g(X))$가 존재한다).

Proof 기댓값 $\mathbf{E}(g(X))$가 존재한다면 $\mathbf{E}(g(X)) = \int_{\mathbb{R}^n} g\, d\mathbf{P}_X = \int_{\mathbb{R}^n} f_X g\, d\mu < \infty$이므로 이는 자명하다. 한편, $f_X g$가 μ−적분가능한 경우에 대해서도 비슷하게 하면 된다. □

특별히, 이산확률벡터의 경우 위의 정리는 다음 따름정리와 같이 합의 형태로 표현된다.

Corollary 2.46 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 이산확률벡터 $X : \Omega \to \mathbb{R}^n$와 Borel 함수 $g : \mathbb{R}^n \to \mathbb{R}$에 대해 $\mathbf{E}(g(X))$가 존재한다면 $\mathbf{E}(g(X)) = \sum_{x \in \mathbb{R}^n} f_X(x) g(x)$이다.

Proof 가정으로부터 $\mathbf{P}_X$의 가산 지지집합 $A \in \mathscr{B}_n$가 존재하고, WLOG, 필요하다면 몇몇 원소를 제거하여 임의의 $x \in A$에 대해 $\mathbf{P}_X\{x\} > 0$이라 해도 된다. 표기의 편의를 위해 $\mathscr{A} = \mathscr{B}_n|_A$라 하면 $\#_A|_{\mathscr{A}} = \#|_{\mathscr{A}}$는 가측공간 $(A, \mathscr{A})$ 위의 셈측도이므로 위의 정리와 f_X의 version을 택하는 관례 그리고 정리 ? ? 의 iii으로부터 $\mathbf{E}(g(X)) = \int_{\mathbb{R}^n} f_X g \, d\#_A = \int_A f_X g \, d\#_A = \int_A (f_X g)|_A \, d\#|_{\mathscr{A}} = \sum_{x\in A} f_X(x)g(x) = \sum_{x\in\mathbb{R}^n} f_X(x)g(x)$이다. □

만약 확률변수가 음이 아니라면 CDF를 사용하여 기댓값을 계산할 수도 있다.

Theorem 2.47 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 음이 아닌 rv. $X : \Omega \to \mathbb{R}_0^+$에 대해 $\mathbf{E}(X)$가 존재하면 $1 - F_X$는 $\mathbb{R}_0^+$에서 μ_1–적분가능하고 $\mathbf{E}(X) = \int_{\mathbb{R}_0^+} 1 - F_X \, d\mu_1$이다.

Proof 함수 $\mathbf{1}_{\{0\le x<y\}}(x, y)$가 음이 아닌 Borel이므로 Fubini의 정리로부터

$$\begin{aligned}
\mathbf{E}(X) &= \int_{\mathbb{R}} x \, d\mathbf{P}_X(x) \\
&= \int_{\mathbb{R}_0^+} x \, d\mathbf{P}_X(x) \\
&= \int_{\mathbb{R}_0^+} \left(\int_{[0,x]} d\mu_1 \right) d\mathbf{P}_X(x) \\
&= \int_{\mathbb{R}_0^+} \left(\int_{\mathbb{R}_0^+} \mathbf{1}_{[0,x)}(y) \, d\mu_1(y) \right) d\mathbf{P}_X(x) \\
&= \int_{\mathbb{R}_0^+} \left(\int_{\mathbb{R}_0^+} \mathbf{1}_{[0,x)}(y) \, d\mathbf{P}_X(x) \right) d\mu_1(y) \\
&= \int_{\mathbb{R}_0^+} \left(\int_{\mathbb{R}_0^+} \mathbf{1}_{(y,\infty)}(x) \, d\mathbf{P}_X(x) \right) d\mu_1(y) \\
&= \int_{\mathbb{R}_0^+} \mathbf{P}_X((y,\infty)) \, d\mu_1(y) \\
&= \int_{\mathbb{R}_0^+} 1 - F_X(y) \, d\mu_1(y)
\end{aligned}$$

이고 이때 $1 - F_X$는 $\mathbb{R}_0^+$에서 μ_1–적분가능하다. □

다음으로 넘어가기 전에, 기댓값에 대해 한 가지 유념해야 할 점이 있다. 기댓값의 정의에는 이가 마치 어떤 확률변수의 특성인 것처럼 쓰여 있지만 사실 기댓값은 확률변수의 특성이 아닌, 분포의 특성이어서 서로 다른 확률공간에서 정의된 서로다른 확률변수라도 그 분포가 같다면 기댓값도 같다. 실제로 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 $(\Omega', \mathscr{F}', \mathbb{P}')$에서 각각 정의된 확률변수 $X : \Omega \to \mathbb{R}, Y : \Omega' \to \mathbb{R}$에 대해 만약 $X \equiv Y$라면 정리 ? ? 로부터 $\mathbf{E}(X) = \int_{\mathbb{R}^n} x \, d\mathbf{P}_X(x) = \int_{\mathbb{R}^n} y \, d\mathbf{P}_Y(y) = \mathbf{E}(Y)$이다. 이러한 사실은 기댓값이 분포의 정보를 요약하여 전달하는 통계량이라는 점에서 어떻게 보면 당연한 것이다.

기댓값에 이어 살펴볼 통계량은 분산이다. 기댓값이 분포의 정보를 요약하여 전달하는 좋은 통계량인 것은 사실이지만, 이가 분포의 정보를 완벽하게 전달하지는 못한다. 이때 발생하는 정보의 손실이 크게 중요하지 않은 경우도 있지만, 때로는 기댓값만으로는 부족한 경우도 있다. 이에 기댓값만을 전달할 때 손실되는 정보를 보충할 목적으로 보조적인 통계량을 제공할 수 있는데, 보통 분포가 기댓값으로부터 퍼져있는 정도를 의미하는 분산이나 표준편차를 제공한다. (이와 같이 분포가 퍼져있는 정도를 의미하는 통계량을 그 분포의 산포도라 하며 분산과 표준편차 이외에도 평균절대편차, 사분위수범위 등이 있다.) 이러한 분산은 기본적으로 기댓값으로 정의된다.

Definition 2.48 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 만약 $\mathbf{E}((X-\mu_X)^2)$이 존재한다면 이를 X의 **분산(variance)**이라 하고 $\mathbf{Var}(X), \sigma_X^2$ 혹은 간단히 σ^2으로 쓴다. 나아가, $\mathbf{Var}(X)$가 존재한다면 $\sqrt{\mathbf{Var}(X)}$를 X의 **표준편차(standard deviation)**라 하고 $\mathbf{Sd}(X), \sigma_X$ 혹은 간단히 σ로 쓴다.

Proposition 2.49 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 $\mathbf{Var}(X)$가 존재할 필요충분조건은 $\mathbf{E}(X^2)$이 존재하는 것이다.

Proof 만약 $\mathbf{Var}(X)$가 존재한다면 정의로부터 $\mathbf{E}(X)$와 $\mathbf{E}((X-\mu_X)^2)$이 존재하고, 곧 $X^2 = (X-\mu_X)^2 + 2\mu_X X - \mu_X^2$에서 $\mathbf{E}(X^2)$이 존재함을 안다. 역으로 $\mathbf{E}(X^2)$이 존재한다면 $|X| \le X^2 \mathbf{1}_{\{|X|\ge 1\}} + \mathbf{1}_{\{|X|<1\}} \le X^2 + 1$에서 $\mathbf{E}(X)$가 존재함을 알고, 곧 $(X-\mu_X)^2 = X^2 - 2\mu_X X + \mu_X^2$에서 $\mathbf{Var}(X) = \mathbf{E}((X-\mu_X)^2)$이 존재한다. □

분산의 기본적인 성질은 모두 기댓값의 성질에 기인하다. 특히 다음 정리의 ii는 간단한 공식이지만 분산을 실제로 계산할 때 유용하게 사용되는 공식이다.

Theorem 2.50 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 $\mathbf{Var}(X)$가 존재한다면 다음이 성립한다.

i. $\mathbf{Var}(X) \ge 0$.
ii. $\mathbf{Var}(X) = \mathbf{E}(X^2) - \mu_X^2$.
iii. 임의의 $a, b \in \mathbb{R}$에 대해 $\mathbf{Var}(aX+b)$가 존재하고 $\mathbf{Var}(aX+b) = a^2\mathbf{Var}(X)$이다.

Proof i. 이는 $(X-\mu_X)^2 \ge 0$에서 자명하다.
ii. 이는

$$\begin{aligned}\mathbf{Var}(X) &= \mathbf{E}((X-\mu_X)^2) \\ &= \mathbf{E}(X^2 - 2\mu_X X + \mu_X^2) \\ &= \mathbf{E}(X^2) - 2\mu_X \mathbf{E}(X) + \mu_X^2\end{aligned}$$

$$= \mathbf{E}(X^2) - \mu_X^2$$

에서 자명하다.

ii. 우선 $\mathbf{Var}(X)$가 존재하므로 $\mathbf{E}(X)$와 $\mathbf{E}(X^2)$이 존재하여 $\mathbf{E}((aX+b)^2) = \mathbf{E}(a^2X^2 + 2abX + b^2)$가 존재하여 $\mathbf{Var}(aX+b)$도 존재한다. 이제

$$\begin{aligned}\mathbf{Var}(aX+b) &= \mathbf{E}([aX+b-\mathbf{E}(aX+b)]^2)\\ &= \mathbf{E}(a^2[X-\mathbf{E}(X)]^2)\\ &= a^2\mathbf{E}((X-\mu_X)^2)\\ &= a^2\mathbf{Var}(X)\end{aligned}$$

에서 정리는 자명하다. □

기댓값과 분산의 성질로부터 $\mathbf{Var}(X)$가 존재하고 0이 아니라면 $(X-\mu_X)/\sigma_X$의 기댓값과 분산은 항상 각각 $0, 1$이다. 이렇게 원래의 확률변수에서 그 기댓값을 빼고, 그 표준편차로 나누는 것을 **표준화(standardization)**라 하며, 기댓값과 분산이 서로다른 분포를 비교할 때 자주 사용되는 방법이다. 한편, 기댓값과 분산은 각각 다음과 같이 $\mathbf{E}((X-x)^2)$가 최솟값을 갖도록 하는 값과 그 최솟값으로 정의할 수도 있다.

Theorem 2.51 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 $\mathbf{Var}(X)$가 존재하면 함수 $x \mapsto \mathbf{E}((X-x)^2)$는 well-define되고 $\mathbf{E}(X)$에서 최솟값 $\mathbf{Var}(X)$를 가진다.

Proof 가정으로부터 $\mathbf{E}(X)$와 $\mathbf{E}(X^2)$이 존재하고 곧 임의의 $x \in \mathbb{R}$에 대해 $(X-x)^2 = X^2 - 2xX + x^2$에서 $\mathbf{E}((X-x)^2)$이 존재하므로 함수 $x \mapsto \mathbf{E}((X-x)^2)$가 well-define된다. 한편, 임의의 $x \in \mathbb{R}$에 대해 $(X-x)^2 = (X-\mu_X)^2 + 2(X-\mu_X)(\mu_X - x) + (\mu_X - x)^2$이므로 $\mathbf{E}((X-x)^2) = \mathbf{E}((X-\mu_X)^2) + 2(\mu_X - x)\mathbf{E}(X-\mu_X) + (\mu_X - x)^2 = \mathbf{Var}(X) + (\mu_X - x)^2$에서 정리가 성립한다. □

이제 기댓값에 관한 다양한 부등식을 알아보는 것으로 이번 절을 마무리하자. 이러한 부등식은 이후 확률론의 다양한 정리를 증명할 때에 약방의 감초와 같이 사용될 것이다.

Theorem 2.52 (Jensen's inequality) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X, Y : \Omega \to \mathbb{R}$와 함수 $g : \mathbb{R} \to \mathbb{R}$에 대해 적당한 구간 $I \subseteq \mathbb{R}$가 존재하여 g가 I에서 볼록하고(혹은 오목하고) $X \in I$ (as.)라 하자. 만약 $\mathbf{E}(X)$와 $\mathbf{E}(g(X))$가 존재한다면 $g(\mathbf{E}(X)) \le \mathbf{E}(g(X))$이다(혹은 $g(\mathbf{E}(X)) \ge \mathbf{E}(g(X))$이다).

Proof 먼저 I가 열린구간인 경우를 생각하여 $I = (a, b)$로 두면 가정과 따름정리 ? ? 의 iii으로부터 $\mathbf{E}(X) \in I$이다. 한편, g가 I에서 볼록하므로 $(\mathbf{E}(X), g(\mathbf{E}(X)))$를 지나는 supporting line $l : y = mx+n$이 존재하여 $mX+n \le g(X)$ (as.)이고, 따라서 $g(\mathbf{E}(X)) = m\mathbf{E}(X)+n = \mathbf{E}(mX+n) \le \mathbf{E}(g(X))$이다.

이제 I가 닫힌구간인 경우를 생각하여 $I = [a,b]$로 두면 가정으로부터 $\mathbf{E}(X) \in I$인데, 만약 $\mathbf{E}(X) \in I^\circ$라면 이전과 같이 하여 $g(\mathbf{E}(X)) \le \mathbf{E}(g(X))$임을 보일 수 있으므로 $\mathbf{E}(X)$가 I의 양 끝점인 경우만 고려하면 된다. 간결한 논의를 위해, $\mathbf{E}(X) = b$라 하자. (반대로 $\mathbf{E}(X) = a$인 경우에도 이와 비슷하게 하면 된다.) 그렇다면 정리 ? ? 의 ii로부터 $X = b$ (as.)이므로 $g(X) = g(b)$ (as.)에서 $g(\mathbf{E}(X)) = g(b) = \mathbf{E}(g(X))$이고, 증명이 끝난다.

한편, 구간 I가 한쪽 끝점만 포함하는 경우에도 I가 닫힌구간인 경우와 비슷하게 하면 같은 결론을 얻고, g가 I에서 오목한 경우에는 $-g$를 생각하면 되므로 증명은 이로써 충분하다. □

Theorem 2.53 (Hölder's inequality) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X, Y : \Omega \to \mathbb{R}$와 $1/p + 1/q = 1$인 $p, q > 1$에 대해 $\mathbf{E}(|X|^p)$와 $\mathbf{E}(|Y|^q)$가 존재한다면 $\mathbf{E}(|XY|)$도 존재하고 $\mathbf{E}(|XY|) \le [\mathbf{E}(|X|^p)]^{1/p}[\mathbf{E}(|Y|^q)]^{1/q}$이다.

Proof 이는 측도론에서 배운 Hölder의 부등식으로부터 자명하다. □

Theorem 2.54 (Minkowski's inequality) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X, Y : \Omega \to \mathbb{R}$와 $p \ge 1$에 대해 $\mathbf{E}(|X|^p)$와 $\mathbf{E}(|Y|^p)$가 존재한다면 $\mathbf{E}(|X+Y|^p)$도 존재하고 $[\mathbf{E}(|X+Y|^p)]^{1/p} \le [\mathbf{E}(|X|^p)]^{1/p} + [\mathbf{E}(|Y|^p)]^{1/p}$이다.

Proof 이는 측도론에서 배운 Minkowski의 부등식으로부터 자명하다. □

Corollary 2.55 (Cauchy-Schwarz's inequality) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X, Y : \Omega \to \mathbb{R}$에 대해 $\mathbf{E}(X^2)$과 $\mathbf{E}(Y^2)$이 존재한다면 $\mathbf{E}(|XY|)$도 존재하고 $\mathbf{E}(|XY|) \le \sqrt{\mathbf{E}(X^2)\mathbf{E}(Y^2)}$이다.

Proof 이는 Hölder의 부등식에서 $p = q = 2$인 특수한 경우이다. □

Theorem 2.56 (Liapounov's inequality) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$와 $0 < p < q$인 $p, q \in \mathbb{R}$에 대해 $\mathbf{E}(|X|^q)$가 존재한다면 $\mathbf{E}(|X|^p)$도 존재하고 $[\mathbf{E}(|X|^p)]^{1/p} \le [\mathbf{E}(|X|^q)]^{1/q}$이다.

Proof 우선 $|X|^p \le |X|^q \mathbf{1}_{\{|X| \ge 1\}} + \mathbf{1}_{\{|X|<1\}} \le |X|^q + 1$에서 $\mathbf{E}(|X|^p)$가 존재함을 안다. 이제 함수 $g : \mathbb{R} \to \mathbb{R}$를 $g : x \mapsto x^{q/p}\mathbf{1}_{\mathbb{R}_0^+}(x)$로 두면 이는 볼록함수이므로 Jensen의 부등식으로부터 $[\mathbf{E}(|X|^p)]^{q/p} = g(\mathbf{E}(|X|^p)) \le \mathbf{E}(g(|X|^p)) = \mathbf{E}(|X|^q)$이고, 곧 증명이 끝난다. □

아래의 두 부등식은 증명의 도구로 쓰일 뿐만 아니라 어떤 확률분포가 특정 값을 얼마 이상 벗어날 확률에 대한 상계의 역할을 한다. 이를 두고 아래의 두 부등식이 'tail probability를 bounding한다' 고 한다.

Theorem 2.57 (Markov's inequality) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$와 $p > 0$에 대해 $\mathbf{E}(|X|^p)$가 존재한다면 임의의 $x > 0$에 대해 $\mathbb{P}\{|X| \geq x\} \leq \mathbf{E}(|X|^p)/x^p$이다.

Proof 이는 측도론에서 배운 Markov의 부등식으로부터 자명하다. □

Corollary 2.58 (Chebyshev's inequality) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 $\mathbf{Var}(X)$가 존재한다면 임의의 $x > 0$에 대해 $\mathbb{P}\{|X - \mu_X| \geq x\} \leq \mathbf{Var}(X)/x^2$이다.

Proof 이는 rv. $X - \mu_X$에 $p = 2$인 경우의 Markov의 부등식을 적용한 결과이다. □

다음은 Chebyshev의 부등식의 기초적인 응용으로, 분산이 0이라는 것의 의미를 잘 나태내준다. 이 결과는 이후에 이따금씩 쓰일 것이다.

Proposition 2.59 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 $\mathbf{Var}(X) = 0$이면 $\mathbf{E}(X)$가 존재하고 $X = \mathbf{E}(X)$ (as.)이다.

Proof 우선 $\mathbf{E}(X)$가 존재함은 자명하다. 이제 사건열 $\{E_i\}$를 $E_i = \{|X - \mu_X| > 1/i\}$로 두면 이는 $\{X \neq \mu_X\}$로 수렴하는 사건열로 $\mathbb{P}\{X = \mu_X\} = 1 - \mathbb{P}\{X \neq \mu_X\} = 1 - \lim_{i\to\infty} \mathbb{P}(E_i)$이다. 한편, 각 $i \in \mathbb{N}$에 대해 Chebyshev의 부등식으로부터 $\mathbb{P}(E_i) \leq \mathbf{Var}(X)/i^2 = 0$이므로 $\mathbb{P}(E_i) \to 0$에서 $\mathbb{P}\{X = \mu_X\} = 1$임을 안다. □

2.4 Moments

이번 절에서는 적률이라는 통계량에 대해 알아본다. 앞서 어떤 분포에 대한 정보를 간단히 전달하기 위해 기댓값이나 분산을 사용하였듯이 적률 또한 어떤 분포의 정보를 전달하기 위해 사용한다. 당장에 기댓값과 분산도 적률의 일종이다. 형식적으로 k차 적률은 확률변수 X의 k제곱의 기댓값으로 정의되는데, 기댓값이나 분산이 그 자체로 나름의 의미를 가지는 것과 달리 k차 적률 그 자체의 의미를 해석하는 것은 쉽지 않다. 그럼에도 불구하고, 하나의 절을 할애하여 적률에 대해 논하는 까닭은 많은 경우에 적률은 분포의 정보를 '손실 없이' 전달할 수 있기 때문이다. 즉, 모든 차수의 적률을 수열의 형태로 전달하면 그것만으로 분포를 완벽하게 복원해낼 수 있다. 그렇다면 적률의 정의부터 시작하자.

Definition 2.60 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$와 $k \in \mathbb{N}_0$에 대해 다음을 정의한다.

i. 만약 $\mathbf{E}(X^k)$가 존재한다면 이를 X의 k**차 적률**(**kth moment**)이라 하고 $\mu_{k,X}$ 혹은 간단히 μ_k로 쓴다.

ii. 만약 $\mathbf{E}((X-\mu_X)^k)$가 존재한다면 이를 X의 k**차 중심화된 적률**(**nth centeral moment**)이라 하고 $\tau_{k,X}$ 혹은 간단히 τ_k로 쓴다.
iii. 만약 $\mathbf{E}([(X-\mu_X)/\sigma_X]^k)$가 존재한다면 이를 X의 k**차 표준화된 적률**(**kth standardized moment**)이라 하고 $\kappa_{k,X}$ 혹은 간단히 κ_k로 쓴다.
iv. 만약 $\mathbf{E}(X^{\underline{k}})$가 존재한다면 이를 X의 k**차 (하향)계승적률**(**kth (falling) factorial moment**)이라 하고 $\mu_{\underline{k},X}$ 혹은 간단히 $\mu_{\underline{k}}$로 쓴다.
v. 만약 $\mathbf{E}(X^{\overline{k}})$가 존재한다면 이를 X의 k**차 (상향)계승적률**(**nth (rising) factorial moment**)이라 하고 $\mu_{\overline{k},X}$ 혹은 간단히 $\mu_{\overline{k}}$로 쓴다.

위에서 정의한 적률을 제외하고도 많은 종류의 적률이 존재하는데, 이는 계산상의 편의를 위한 목적이 크다. 나아가 중심화된 적률과 표준화된 적률은 그 이름에서도 알 수 있듯이 각각 원래 확률변수의 평행이동과 Affine 변환에 대해 불변이라는 유용한 특성을 가지기도 한다. 한편, 적률의 존재성을 따지는 데에는 다음 두 명제가 유용하다.

Proposition 2.61 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}$에 대해 만약 X의 k차 적률(혹은 중심화된 적률, 표준화된 적률, 하향계승적률, 상향계승적률)이 존재하면 X의 $1,\cdots,k$차 적률(혹은 중심화된 적률, 표준화된 적률, 하향계승적률, 상향계승적률)이 존재한다.

Proof 만약 $k\le 1$이면 명제가 자명하므로 $k>1$이라 하자. 먼저 적률, 중심화된 적률, 표준화된 적률의 경우는 Liapounov의 부등식으로부터 자명하다. 한편, 하향계승적률의 경우, 만약 X의 k차 하향계승적률이 존재한다면 $X^{\underline{k}}=X^{\underline{k-1}}(X-k+1)$에서

$$\begin{aligned}
|X^{\underline{k-1}}| &= \left|\frac{X^{\underline{k}}}{X-k+1}\right|\mathbf{1}_{\{|X-k+1|\ge 1\}}+|X^{\underline{k-1}}|\mathbf{1}_{\{|X-k+1|<1\}}\\
&\le |X^{\underline{k}}|\mathbf{1}_{\{|X-k+1|\ge 1\}}+|X^{\underline{k-1}}|\mathbf{1}_{\{|X-k+1|<1\}}\\
&< |X^{\underline{k}}|+k!
\end{aligned}$$

이다. 여기서 마지막 부등호는 함수 $x\mapsto|x^{\underline{k-1}}|$이 $(k-2,\infty)$에서 증가하므로 성립한다. 이로부터 X의 $k-1$차 하향계승적률이 존재함을 알고, 이를 $k-2$번 반복하면 X의 $1,\cdots,k$차 하향계승적률이 존재함을 안다. 마지막으로 상향계승적률의 경우에는 하향계승적률과 비슷하게 하면 된다. □

Proposition 2.62 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}$에 대해 TFAE.

i. Rv. X의 k차 적률이 존재한다.
ii. Rv. X의 k차 중심화된 적률이 존재한다.
iii. Rv. X의 k차 표준화된 적률이 존재한다.

iv. Rv. X의 k차 하향계승적률이 존재한다.
v. Rv. X의 k차 상향계승적률이 존재한다.

단, ii의 경우 $\mathbf{E}(X)$가 존재함을 전제로 하고, iii의 경우 $\mathbf{Var}(X)$가 존재하며 0이 아님을 전제로 한다.

PROOF 만약 $k = 0$이면 명제가 자명하므로 $k > 0$이라 하자.

i ⇔ ii. 만약 X의 k차 적률이 존재하면 명제 ? ? 로부터 X의 $1, \cdots, k$차 적률이 존재한다. 이로부터

$$(X - \mu_X)^k = \sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} X^l \mu_X^{k-l}$$

에서 X의 k차 중심화된 적률이 존재한다. 역으로, $\mathbf{E}(X)$가 존재하고 X의 k차 중심화된 적률이 존재하면 다시 명제 ? ? 로부터 X의 $1, \cdots, k$차 중심화된 적률이 존재한다. 그렇다면

$$\begin{aligned} X^k &= (X - \mu_X + \mu_X)^k \\ &= \sum_{l=0}^{k} \binom{k}{l} (X - \mu_X)^l \mu_X^{k-l} \end{aligned}$$

에서 X의 k차 적률이 존재한다.

ii ⇔ iii. 이는 정의로부터 자명하다.

i ⇒ iv. 명제 ? ? 로부터 X의 $1, \cdots, k$차 적률이 존재하는데, 적당한 $a_0, \cdots, a_k \in \mathbb{R}$에 대해 $X^{\underline{k}} = \prod_{l=0}^{k-1}(X - l) = \sum_{l=0}^{k} a_l X^l$이므로 곧 X의 k차 하향계승적률이 존재한다.

iv ⇒ i. 이번에도 명제 ? ? 로부터 X의 $1, \cdots, k$차 하향계승적률이 존재한다. 한편, 적당한 $a_0, \cdots, a_k \in \mathbb{R}$에 대해 $X^{\underline{k}} = \prod_{l=0}^{k-1}(X - l) = \sum_{l=0}^{k} a_l X^l$이고 이때 $a_k = 1$임이 명백하므로 $X^k = X^{\underline{k}} - \sum_{l=0}^{k-1} a_l X^l$이다. 이를 $k - 1$번 반복하면 다시 적당한 $b_0, \cdots, b_k \in \mathbb{R}$에 대해 $X^k = \sum_{l=0}^{k} b_l X^{\underline{l}}$이고, 곧 X의 k차 적률이 존재한다.

i ⇔ v. 이는 i ⇔ iv의 증명과 비슷하게 하면 된다. □

앞서 적률 그 자체에 의미를 해석하는 것은 쉽지 않다고 했는데, 특별히 3차와 4차 표준화된 적률에 대해서는 어느정도 널리 받아들여지는 해석이 있어 잠시 소개한다.

Definition 2.63 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 만약 X의 3차 표준화된 적률이 존재한다면 이를 특별히 X의 **왜도(skewness)**라 하고 $\mathbf{Skew}(X)$, τ_X 혹은 간단히 τ로 쓴다. 비슷하게, X의 4차 표준화된 적률이 존재한다면 $\kappa_{4,X} - 3$을 특별히 X의 **(excess) 첨도(− kurtosis)**라 하고 $\mathbf{Kurt}(X)$, κ_X 혹은 간단히 κ로 쓴다.

왜도는 분포가 얼마나 치우쳤는가, 즉 얼마나 비대칭인가에 대한 척도이다. 정규분포와 같이 분포가 기댓값을 기준으로 정확히 대칭을 이루면 왜도는 0이 되고, 기댓값이 분포

의 중앙값보다 작아서 분포가 왼쪽으로 치우쳐 있으면 왜도는 양수가 되며 이때 그 분포는 **positively skewed**되었다고 한다. 반대로 기댓값이 분포의 중앙값보다 커서 분포가 오른쪽으로 치우쳐 있으면 왜도는 음수가 되며 이때 그 분포는 **negatively skewed**되었다고 한다.

첨도는 분포가 얼마나 뾰족한가, 즉 분포의 꼬리가 얼마나 두꺼운가에 대한 척도이다. 왜도와는 달리 4차 표준화된 적률에서 3을 뺀 것으로 정의되어 있는데, 이는 표준정규분포의 4차 표준화된 적률이 3인 바, 비교의 기준이 되는 표준정규분포의 첨도를 0으로 만들기 위함이다. 표준정규분포와 같이 첨도가 0이면 이때 그 분포는 **mesokurtic**하다고 하고, 표준정규분포보다 얇은 꼬리를 가지면 첨도는 양수가 되며 이때 그 분포는 **leptokurtic**하다고 한다. 반대로 표준정규분포보다 두꺼운 꼬리를 가지면 첨도는 음수가 되며 이때 그 분포는 **platykurtic**하다고 한다. 일반적으로 꼬리의 두께는 outlier와 같은 극단적 사건의 발생확률을 의미하므로 leptokurtic한 분포와 platykurtic한 분포는 표준정규분포에 비해 극단적 사건의 확률이 각각 더 낮고 높다.

한편, 왜도와 첨도는 모두 Affine 변환에 대해 불변이다. 다만, 왜도의 경우 부호 정도의 영향은 받으며, 이는 그 정의를 생각해보면 당연하다.

Theorem 2.64 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$와 임의의 $a, b \in \mathbb{R}$에 대해 $a \neq 0$이면 다음이 성립한다.

i. 만약 $\mathbf{Skew}(X)$가 존재하면 $\mathbf{Skew}(aX+b)$도 존재하고 $\mathbf{Skew}(aX+b) = \mathrm{sgn}(a)\mathbf{Skew}(X)$이다.
ii. 만약 $\mathbf{Kurt}(X)$가 존재하면 $\mathbf{Kurt}(aX+b)$도 존재하고 $\mathbf{Kurt}(aX+b) = \mathbf{Kurt}(X)$이다.

Proof 먼저 $a \neq 0$이므로 $\mathbf{Var}(aX+b) = a^2\mathbf{Var}(X) \neq 0$이다. 또한, $\mathbf{Skew}(X)$가 존재하면 X의 3차 적률이 존재하여 곧 $\mathbf{E}(X), \mathbf{E}(X^2), \mathbf{E}(X^3)$이 모두 존재하며, 이로부터 $\mathbf{E}((aX+b)^3) = a^3\mathbf{E}(X^3) + 3a^2b\mathbf{E}(X^2) + 3ab^2\mathbf{E}(X) + b^3$이 존재하므로 $\mathbf{Skew}(aX+b)$도 존재한다. 한편,

$$\begin{aligned}
\mathbf{Skew}(aX+b) &= \mathbf{E}\left(\left(\frac{aX+b-\mu_{aX+b}}{\sigma_{aX+b}}\right)^3\right) \\
&= \mathbf{E}\left(\left[\frac{a(X-\mu_X)}{|a|\sigma_X}\right]^3\right) \\
&= \mathbf{E}\left(\mathrm{sgn}(a)\left(\frac{X-\mu_X}{\sigma_X}\right)^3\right) \\
&= \mathrm{sgn}(a)\mathbf{Skew}(X)
\end{aligned}$$

에서 i이 성립한다. 비슷하게 $\mathbf{Kurt}(X)$가 존재하면 $\mathbf{Kurt}(aX+b)$도 존재하고

$$\mathbf{Kurt}(aX+b) = \mathbf{E}\left(\left(\frac{aX+b-\mu_{aX+b}}{\sigma_{aX+b}}\right)^4\right) - 3$$

$$
\begin{aligned}
&= \mathbf{E}\left(\left[\frac{a(X-\mu_X)}{|a|\sigma_X}\right]^4\right) - 3 \\
&= \mathbf{E}\left(\left(\frac{X-\mu_X}{\sigma_X}\right)^4\right) - 3 \\
&= \mathbf{Kurt}(X)
\end{aligned}
$$

에서 ii가 성립한다. □

적률은 확률벡터에 대해 보다 일반적으로 정의할 수도 있다. 물론, 이렇게 확률벡터에 대해 정의된 적률에 대해서도 명제 ? ? , ? ? 는 비슷하게 성립한다.

Definition 2.65 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 $\alpha \in \mathbb{N}_0^n$에 대해 다음을 정의한다.

i. 만약 $\mathbf{E}(X^\alpha)$가 존재한다면 이를 X의 $|\alpha|$**차 (결합)적률(**$|\alpha|$**th (joint) moment)**이라 하고 $\mu_{\alpha,X}$ 혹은 간단히 μ_α로 쓴다.
ii. 만약 $\mathbf{E}((X_i - \mu_{X_i})^\alpha)$가 존재한다면 이를 X의 $|\alpha|$**차 중심화된 (결합)적률(**$|\alpha|$**th (joint) centeral moment)**이라 하고 $\tau_{\alpha,X}$ 혹은 간단히 τ_α로 쓴다.
iii. 만약 $\mathbf{E}(((X_i - \mu_{X_i})/\sigma_{X_i})^\alpha)$가 존재한다면 이를 X의 $|\alpha|$**차 표준화된 (결합)적률(**$|\alpha|$**th (joint) standardized moment)**이라 하고 $\kappa_{\alpha,X}$ 혹은 간단히 κ_α로 쓴다.
iv. 만약 $\mathbf{E}(X^{\underline{\alpha}})$가 존재한다면 이를 X의 $|\alpha|$**차 (하향결합)계승적률(**$|\alpha|$**th (joint falling) factorial moment)**이라 하고 $\mu_{\underline{\alpha},X}$ 혹은 간단히 $\mu_{\underline{\alpha}}$로 쓴다.
v. 만약 $\mathbf{E}(X^{\overline{\alpha}})$가 존재한다면 이를 X의 $|\alpha|$**차 (상향결합)계승적률(**$|\alpha|$**th (joint rising) factorial moment)**이라 하고 $\mu_{\overline{\alpha},X}$ 혹은 간단히 $\mu_{\overline{\alpha}}$로 쓴다.

Proposition 2.66 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 만약 X의 모든 k차 적률(혹은 중심화된 적률, 표준화된 적률, 하향계승적률, 상향계승적률)이 존재하면 X의 모든 $1, \cdots, k$차 적률(혹은 중심화된 적률, 표준화된 적률, 하향계승적률, 상향계승적률)이 존재한다.

Proof 만약 $k = 0$이면 명제가 자명하므로 $k > 0$이라 하고, 적률에 대해서는 수학적 귀납법을 사용하자. 우선 각 $i \le n$에 대해 $\mathbf{E}(X_i^k)$가 존재하므로 $\mathbf{E}(X_i)$도 존재하여 모든 1차 적률이 존재한다. 이제 귀납가정으로서 $l < k$에 대해 모든 $1, \cdots, l$차 적률이 존재한다고 하고 $|\alpha| = l + 1$인 임의의 $\alpha \in \mathbb{N}_0^n$를 택하여 WLOG, $\alpha_1 \ne 0$이라 하면

$$
\begin{aligned}
|X^\alpha| &\le |X_1|^{\alpha_1+k-l-1} \prod_{i=2}^{n} |X_i|^{\alpha_i} \mathbf{1}_{\{|X_1| \ge 1\}} + \prod_{i=2}^{n} |X_i|^{\alpha_i} \mathbf{1}_{\{|X_1| < 1\}} \\
&\le |X_1|^{\alpha_1+k-l-1} \prod_{i=2}^{n} |X_i|^{\alpha_i} + \prod_{i=2}^{n} |X_i|^{\alpha_i}
\end{aligned}
$$

에서 $\mathbf{E}(|X^{\alpha}|)$가 존재함을 알고, 곧 모든 $l+1$차 적률이 존재하여 모든 $1,\cdots,k$차 적률이 존재함을 안다. 한편, 중심화된 적률과 표준화된 적률에 대해서도 이와 비슷하게 하면 된다.

다음으로 X의 하향계승적률에 대해서도 수학적 귀납법을 사용하자. 우선 각 $i \le n$에 대해 $\mathbf{E}(X_i^k)$가 존재하므로 $\mathbf{E}(X_i)$도 존재하여 모든 1차 하향계승적률이 존재한다. 이제 귀납가정으로서 $l<k$에 대해 모든 $1,\cdots,l$차 하향계승적률이 존재한다고 하고 $|\alpha|=l+1$인 임의의 $\alpha \in \mathbb{N}_0^n$을 택하여 WLOG, $\alpha_1 \neq 0$이라 하자. 한편, 함수 $f,g:\mathbb{R}\to\mathbb{R}$를 각각 $f: x \mapsto |(x-\alpha_1)^{\underline{k-l-1}}|$, $g: x \mapsto |x^{\underline{\alpha_1}}|$이라 두면 적당한 $M>\alpha_1+k-l-2$이 존재하여 $(-\infty,-M)$와 (M,∞)에서 f는 각각 감소하고 증가하며 $f(-M), f(M) \ge 1$이고 $(-M,M)$에서 $g<g(M)=M^{\underline{\alpha_1}}$이다. 이상을 종합하면 $X^{\underline{\alpha_1+k-l-1}} = X^{\underline{\alpha_1}}(X_1-\alpha_1)^{\underline{k-l-1}}$에서

$$\begin{aligned}
|X^{\underline{\alpha}}| &= \frac{|X_1^{\underline{\alpha_1+k-l-1}}|}{|(X_1-\alpha_1)^{\underline{k-l-1}}|}\left(\prod_{i=2}^{n}|X_i^{\underline{\alpha_i}}|\right)\mathbf{1}_{\{|X_1|\ge M\}} + |X^{\underline{\alpha}}|\mathbf{1}_{\{|X_1|<M\}} \\
&\le |X_1^{\underline{\alpha_1+k-l-1}}|\left(\prod_{i=2}^{n}|X_i^{\underline{\alpha_i}}|\right)\mathbf{1}_{\{|X_1|\ge M\}} + M^{\underline{\alpha_1}}\left(\prod_{i=2}^{n}|X_i^{\underline{\alpha_i}}|\right)\mathbf{1}_{\{|X_1|<M\}} \\
&< |X_1^{\underline{\alpha_1+k-l-1}}|\prod_{i=2}^{n}|X_i^{\underline{\alpha_i}}| + M^{\underline{\alpha_1}}\prod_{i=2}^{n}|X_i^{\underline{\alpha_i}}|
\end{aligned}$$

이므로 $\mathbf{E}(|X^{\underline{\alpha}}|)$가 존재함을 알고, 곧 모든 $l+1$차 하향계승적률이 존재하여 모든 $1,\cdots,k$차 하향계승적률이 존재함을 안다. 마지막으로 상향계승적률의 경우에는 하향계승적률과 비슷하게 하면 된다. □

Proposition 2.67 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}^n$에 대해 TFAE.

i. Rv. X의 모든 k차 적률이 존재한다.
ii. Rv. X의 모든 k차 중심화된 적률이 존재한다.
iii. Rv. X의 모든 k차 표준화된 적률이 존재한다.
iv. Rv. X의 모든 k차 하향계승적률이 존재한다.
v. Rv. X의 모든 k차 상향계승적률이 존재한다.

단, ii의 경우 각 $i \le n$에 대해 $\mathbf{E}(X_i)$가 존재함을 전제로 하고, iii의 경우 각 $i \le n$에 대해 $\mathbf{Var}(X_i)$가 존재하며 0이 아님을 전제로 한다.

Proof 만약 $k=0$이면 명제가 자명하므로 $k>0$이라 하자.

i $\Leftrightarrow$ ii. 만약 X의 모든 k차 적률이 존재하면 명제 ? ? 로부터 X의 모든 $1,\cdots,k$차 적률이 존재한다. 이제 $|\alpha|=k$인 임의의 $\alpha\in\mathbb{N}_0^n$에 대해 $(X_i-\mu_{X_i})^{\alpha}$의 전개식에서 X^{α}가 최고차항임을 생각해보면 충분조건임이 자명하고, 필요조건임도 이와 비슷하게 보일 수 있다.

ii $\Leftrightarrow$ iii. 이는 정의로부터 자명하다.

i $\Rightarrow$ iv. 명제 ? ? 로부터 X의 모든 $1, \cdots, k$차 적률이 존재하는데, $|\alpha| = k$인 임의의 $\alpha \in \mathbb{N}_0^n$에 대해 $X^{\underline{\alpha}} = \prod_{i=1}^n \prod_{j=0}^{\alpha_i - 1}(X_i - j)$의 전개식에서 X^α가 최고차항임을 생각해보면 명제가 자명하다.

iv $\Rightarrow$ i. 수학적 귀납법을 사용하기 위해 증명하고자 하는 명제를 $\mathsf{P}(k)$라 하자. 우선 $\mathsf{P}(1)$은 명제 ? ? 로부터 자명하다. 이제 귀납가정으로서 $l < k$에 대해 $\mathsf{P}(1), \cdots, \mathsf{P}(l)$이 모두 성립한다고 하고 모든 $l+1$차 하향계승적률이 존재한다고 하여 $|\alpha| = l+1$인 임의의 $\alpha \in \mathbb{N}_0^n$을 택하자. 그렇다면 명제 ? ? 로부터 X의 모든 $1, \cdots, l+1$차 하향계승적률이 존재하고, $X^{\underline{\alpha}} = \prod_{i=1}^n \prod_{j=0}^{\alpha_i - 1}(X_i - j)$의 전개식에서 X^α가 최고차항이므로 X^α는 $X^{\underline{\alpha}}$와 $1, \cdots, l$차항 그리고 상수항의 선형결합으로 쓸 수 있어서 $\mathbf{E}(X^\alpha)$가 존재하여 $\mathsf{P}(l+1)$도 성립하고, 곧 임의의 $k \in \mathbb{N}$에 대해 $\mathsf{P}(k)$가 성립한다.

i $\Leftrightarrow$ v. 이는 i $\Leftrightarrow$ iv의 증명과 비슷하게 하면 된다. □

보통 적률에 대한 이론을 전개할 때에는 적률을 그 자체로 다루는 것보다 아래의 MGF를 통해 다루는 것이 더 편리하다.

Definition 2.68 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 적당한 0의 근방 $U \subseteq \mathbb{R}^n$가 존재하여 임의의 $t \in U$에 대해 $\mathbf{E}(e^{t^\top X})$가 존재하면 rv. X의 **적률생성함수(moment generating function)**를 $M_X : U \to \mathbb{R}$로 쓰고 $M_X : t \mapsto \mathbf{E}(e^{t^\top X})$로 정의한다. 특별히, $n \geq 2$인 경우 M_X를 rv. $X_1, \cdots, X_n$의 **결합적률생성함수(joint moment generating function)**라 하기도 한다.

Theorem 2.69 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 M_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재하면 다음이 성립한다.

i. $M_X(0) = 1$.
ii. $M_X > 0$.
iii. MGF M_X는 log−볼록하다.
iv. 임의의 $a \in \mathbb{R}$와 임의의 $b \in \mathbb{R}^n$에 대해 M_{aX+b}가 적당한 0의 근방 $V \subseteq \mathbb{R}^n$ 위에서 존재하고, $at \in U$인 임의의 $t \in V$에 대해 $M_{aX+b}(t) = e^{t^\top b} M_X(at)$이다.

Proof i. 이는 $M_X(0) = \mathbf{E}(1) = 1$에서 자명하다.

ii. 따름정리 ? ? 의 iii으로부터 임의의 $t \in \mathbb{R}^n$에 대해 $M_X(t) = \mathbf{E}(e^{t^\top X}) > \mathbf{E}(0) = 0$이므로 이는 자명하다.

iii. 임의의 $s, t \in U$와 $\lambda \in (0,1)$에 대해 Hölder의 부등식으로부터 $M_X(\lambda s + (1-\lambda)t) = \mathbf{E}(e^{[\lambda s + (1-\lambda)t]^\top X}) \leq [\mathbf{E}(e^{s^\top X})]^\lambda [\mathbf{E}(e^{t^\top X})]^{1-\lambda} = M_X(s)^\lambda M_X(t)^{1-\lambda}$이고, ii로부터 $M_X > 0$이므로 양변에 log를 취하면 $\log M_X(\lambda s + (1-\lambda)t) \leq \lambda \log M_X(s) + (1-\lambda) \log M_X(t)$이다. 이 부등식이 $\lambda = 0, 1$인 경우에도 성립함이 자명하므로 M_X는 log−볼록하다.

iv. 집합 $V = \{t \in \mathbb{R}^n : at \in U\}$를 생각하면 이는 명백히 0의 근방이고, 임의의 $t \in V$에 대해 $\mathbf{E}(e^{t^\top(aX+b)}) = e^{t^\top b}\mathbf{E}(e^{(at)^\top X}) = e^{t^\top b}M_X(at)$이므로 정리가 성립한다. □

MGF의 정의만 보아서는 이와 적률 사이의 관계를 눈치채기가 쉽지 않다. 이에 다음 정리는 MGF와 적률의 관계를 명시적으로 잘 보여준다.

Lemma 2.70 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 M_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재하면, 적당한 0의 근방 $V \subseteq U$가 존재하여 임의의 $t \in V$에 대해 $\mathbf{E}(e^{|t|^\top |X|})$가 존재한다. 특별히, $U = \mathbb{R}^n$이면 $V = \mathbb{R}^n$으로 잡을 수 있다.

Proof 먼저 $V := B(||t_0||) \subseteq U$인 0이 아닌 $t_0 \in U$를 택하고, 이어서 모든 성분이 음이 아닌 임의의 $t \in V$를 택한 뒤, $\mathbb{R}^n$을 2^n개의 사분공간으로 나누어 각각을 $A_1, \cdots, A_{2^n}$이라 하자. (이때 각 사분면의 경계도 적당히 나눈다.) 이제 임의의 A_j를 고정하고 임의의 $x \in A_j^\circ$에 대해 $v = (\mathrm{sgn}(x_i))$라 하면 이는 x의 선택과는 무관하게 well-define된다. 그렇다면 $s = (v_i t_i) \in V$임이 명백하므로

$$
\begin{aligned}
\int_\Omega e^{t^\top |X|}\mathbf{1}_{A_j}(X)\,d\mathbb{P} &= \int_{A_j} e^{t^\top |x|}\,d\mathbf{P}_X(x) \\
&= \int_{A_j} \exp\left(\sum_{i=1}^n v_i t_i x_i\right) d\mathbf{P}_X(x) \\
&= \int_{A_j} e^{s^\top x}\,d\mathbf{P}_X(x) \\
&\leq \int_{\mathbb{R}^n} e^{s^\top x}\,d\mathbf{P}_X(x) \\
&= \mathbf{E}(e^{s^\top X}) \\
&< \infty
\end{aligned}
$$

에서 $e^{t^\top |X|}\mathbf{1}_{A_j}(X)$가 적분가능하고, 곧 $e^{t^\top |X|} = \sum_{j=1}^{2^n} e^{t^\top |X|}\mathbf{1}_{A_j}(X)$도 적분가능하여 $\mathbf{E}(e^{t^\top |X|})$가 존재한다. 이는 임의의 $t \in V$에 대해 $\mathbf{E}(e^{|t|^\top |X|})$가 존재함을 함의하는 한편, $U = \mathbb{R}^n$인 경우에는 이상의 논의에서 $V = \mathbb{R}^n$으로 택할 수 있음이 분명하므로 증명이 끝난다. □

Theorem 2.71 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 M_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재하면 X의 모든 적률이 존재하고, 적당한 0의 근방 $V \subseteq U$가 존재하여 임의의 $t \in V$에 대해 $M_X(t) = \sum_{i=0}^\infty \sum_{|\alpha|=i} \mathbf{E}(X^\alpha)t^\alpha/\alpha! = \sum_{\alpha \in \mathbb{N}_0^n} \mathbf{E}(X^\alpha)t^\alpha/\alpha!$이다. 특별히, $U = \mathbb{R}^n$이면 $V = \mathbb{R}^n$으로 잡을 수 있으며, 여기서 급수는 절대수렴하여 well-define된다.

Proof 먼저 임의의 $\alpha \in \mathbb{N}_0^n$에 대해 $\mathbf{E}(X^\alpha)$가 존재함을 보이자. 이를 위해 각 성분이 양수인 $t \in U$를 택하면

$$\begin{aligned}
|X^\alpha| &= \frac{(t_i|X_i|)^\alpha}{t^\alpha} \\
&\le \frac{1}{t^\alpha}\left(\sum_{i=1}^n t_i|X_i|\right)^{|\alpha|} \bigg/ \binom{|\alpha|}{\alpha} \\
&= \frac{\alpha!}{t^\alpha} \cdot \frac{(t^\top|X|)^{|\alpha|}}{|\alpha|!} \\
&\le \frac{\alpha!}{t^\alpha} \sum_{i=0}^\infty \frac{(t^\top|X|)^i}{i!} \\
&= \frac{\alpha!}{t^\alpha} e^{t^\top|X|}
\end{aligned}$$

인데, 보조정리로부터 t를 충분히 작게 잡아주면 $\mathbf{E}(e^{t^\top|X|})$가 존재하므로 $\mathbf{E}(X^\alpha)$도 존재함을 알고, 곧 X의 모든 적률이 존재한다.

다음으로, 임의의 $k \in \mathbb{N}_0$와 $t \in U$에 대해

$$\begin{aligned}
\left|\sum_{i=0}^k \frac{(t^\top X)^i}{i!}\right| &\le \sum_{i=0}^k \frac{|t^\top X|^i}{i!} \\
&\le \sum_{i=0}^\infty \frac{|t^\top X|^i}{i!} \\
&= e^{|t^\top X|} \\
&\le e^{t^\top X}\mathbf{1}_{\{t^\top X \ge 0\}} + \mathbf{1}_{\{t^\top X < 0\}} \\
&\le e^{t^\top X} + 1
\end{aligned}$$

이므로 DCT에서

$$\begin{aligned}
M_X(t) &= \mathbf{E}\left(\sum_{i=0}^\infty \frac{(t^\top X)^i}{i!}\right) \\
&= \sum_{i=0}^\infty \frac{\mathbf{E}((t^\top X)^i)}{i!} \\
&= \sum_{i=0}^\infty \frac{1}{i!} \sum_{|\alpha|=i} \binom{i}{\alpha} \mathbf{E}((t_i X_i)^\alpha) \\
&= \sum_{i=0}^\infty \sum_{|\alpha|=i} \mathbf{E}(X^\alpha)\frac{t^\alpha}{\alpha!} \qquad (*)
\end{aligned}$$

이다. 이제 $V := B(||t_0||) \subseteq U$인 0이 아닌 $t_0 \in U$를 택하자. 그렇다면 $\sum_{i=0}^\infty \sum_{|\alpha|=i} |\mathbf{E}(X^\alpha)t^\alpha/\alpha!| \le \sum_{i=0}^\infty \sum_{|\alpha|=i} \mathbf{E}(|X|^\alpha)|t|^\alpha/\alpha!$이므로 보조정리로부터 t_0를 충분히 작게 잡아주면 $\mathbf{E}(e^{|t|^\top|X|})$가 존재하고, 위에서와 비슷하게 하여 $(*)$의 급수가 절대수렴함을 쉽게 보일 수 있다. 이로부

터 $(*)$의 급수에서의 합의 순서를 바꿀 수 있으므로 원하는 형태의 식을 얻고, $U = \mathbb{R}^n$인 경우에는 이상의 논의에서 $V = \mathbb{R}^n$으로 택할 수 있음이 분명하므로 증명이 끝난다. $\square$

한편, MGF는 적률을 계산하는 방법으로도 유용하게 쓰인다.

Theorem 2.72 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 M_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재한다고 하자. 그렇다면 임의의 $\alpha \in \mathbb{N}_0^n$에 대해 $\mathbf{E}(X^\alpha) = \partial^\alpha M_X(0)$이다.

Proof 정리 ? ? 로부터 적당한 0의 근방 $V \subseteq U$가 존재하여 임의의 $t \in V$에 대해 $M_X(t) = \sum_{\beta \in \mathbb{N}_0^n} \mathbf{E}(X^\beta) t^\beta / \beta!$가 절대수렴하는 멱급수이므로 $\partial^\alpha M_X(0) = \sum_{\beta \in \mathbb{N}_0^n} \mathbf{E}(X^\beta) \partial^\alpha t^\beta |_{t=0} / \beta! = \sum_{\beta \geq \alpha} \mathbf{E}(X^\beta) t^{\beta - \alpha} |_{t=0} / (\beta - \alpha)! = \mathbf{E}(X^\alpha)$이다. $\square$

이렇게 적률과 긴밀하게 연결되어 있는 MGF이지만 이는 그 존재성이 항상 보장되지 않는다는 큰 단점을 가진다. 이론 전개에 있어 이러한 단점은 많은 귀찮음을 유발하므로 자연스럽게 항상 존재하면서도 MGF의 역할을 해 줄수 있는 대체재를 생각해내기에 이르렀는데, 그것이 바로 확률벡터의 PDF의 Fourier 변환인 CF이다. 아래의 정의에서 기댓값에 복소함수가 들어갔는데, 측도론의 에필로그에서 논한 복소함수의 적분에 대한 내용을 그대로 적용하면 된다. 즉, 확률변수 X, Y에 대해 $\mathbf{E}(X)$와 $\mathbf{E}(Y)$가 존재한다면 $\mathbf{E}(X + \mathbf{i}Y) = \mathbf{E}(X) + \mathbf{i}\mathbf{E}(Y)$이다.

Definition 2.73 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 rv. X의 **특성함수 (characteristic function)**를 $\varphi_X : \mathbb{R}^n \to \mathbb{C}$로 쓰고 $\varphi_X : t \mapsto \mathbf{E}(e^{\mathbf{i}t^\top X})$로 정의한다. 특별히, $n \geq 2$인 경우 φ_X를 rv. $X_1, \cdots, X_n$의 **결합특성함수(joint characteristic function)**라 하기도 한다.

Proposition 2.74 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 임의의 $t \in \mathbb{R}^n$에 대해 $\mathbf{E}(e^{\mathbf{i}t^\top X})$가 항상 존재한다. 따라서 φ_X는 well-define된다.

Proof 함수 $e^{\mathbf{i}t^\top X}$의 실수부 $\cos t^\top X$와 허수부 $\sin t^\top X$가 모두 적분가능함을 보이면 되는데, 이는 $|\cos t^\top X|, |\sin t^\top X| \leq 1$에서 자명하다. $\square$

CF는 항상 존재한다는 장점과 더불어 함수로서도 훌륭한 성질을 가진다.

Theorem 2.75 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 다음이 성립한다.

i. $\varphi_X(0) = 1$.
ii. $|\varphi_X| \leq 1$.
iii. φ_X는 균등연속이다.
iv. 임의의 $a \in \mathbb{R}$와 임의의 $b, t \in \mathbb{R}^n$에 대해 $\varphi_{aX+b}(t) = e^{\mathbf{i}t^\top b} \varphi_X(at)$이다.

Proof i. 이는 $\varphi_X(0)=\mathbf{E}(1)=1$에서 자명하다.

ii. 임의의 $t\in\mathbb{R}^n$에 대해 $|\varphi_X(t)|=|\mathbf{E}(e^{\mathbf{i}t^\top X})|\le\mathbf{E}(|e^{\mathbf{i}t^\top X}|)=\mathbf{E}(1)=1$이므로 이는 자명하다.

iii. 함수 $f:\Omega\times\mathbb{R}^n$를 $f:(\omega,t)\mapsto|e^{\mathbf{i}t^\top X(\omega)}-1|$로 두면 임의의 $\omega\in\Omega$에 대해 section f_ω는 연속이다. 또한, 임의의 $(\omega,t)\in\Omega\times\mathbb{R}^n$에 대해 $|e^{\mathbf{i}t^\top X(\omega)}-1|\le|e^{\mathbf{i}t^\top X(\omega)}|+1=2$이므로 임의의 $t\in\mathbb{R}^n$에 대해 section f^t는 적분가능하고 $|f|\le 2$이다. 이상으로부터 f가 따름정리 ? ? 의 조건을 모두 만족시킴을 알 수 있으므로 함수 $t\mapsto\int_\Omega|e^{\mathbf{i}t^\top X}-1|\,d\mathbb{P}=\mathbf{E}(|e^{\mathbf{i}t^\top X}-1|)$는 연속이다. 특히 이가 0에서 연속이므로 $t\to 0$이면 $\mathbf{E}(|e^{\mathbf{i}t^\top X}-1|)\to 0$이고, 곧 임의의 $\varepsilon>0$을 택하면 적당한 $\delta>0$가 존재하여 $|t|<\delta$이면 $\mathbf{E}(|e^{\mathbf{i}t^\top X}-1|)<\varepsilon$이다. 이로부터 $\|s-t\|<\delta$인 임의의 $s,t\in\mathbb{R}^n$에 대해 $|\varphi_X(s)-\varphi_X(t)|=|\mathbf{E}(e^{\mathbf{i}s^\top X}-e^{\mathbf{i}t^\top X})|\le\mathbf{E}(|e^{\mathbf{i}t^\top X}(e^{\mathbf{i}(s-t)^\top X}-1)|)=\mathbf{E}(|e^{\mathbf{i}(s-t)^\top X}-1|)<\varepsilon$이 되어 φ_X가 균등연속임을 안다.

iv. 이는 $\mathbf{E}(e^{\mathbf{i}t^\top(aX+b)})=e^{\mathbf{i}t^\top b}\mathbf{E}(e^{\mathbf{i}(at)^\top X})=e^{\mathbf{i}t^\top b}\varphi_X(at)$에서 자명하다. □

나아가, MGF와 적률의 관계를 잘 보여주었던 정리 ? ? , ? ? 가 CF에 대해서도 거의 비슷하게 (사실은 더 잘) 성립한다.

Lemma 2.76 임의의 $x,y\in\mathbb{R}^n$와 임의의 $k\in\mathbb{N}_0$에 대해

$$\left|e^{\mathbf{i}x^\top y}-\sum_{i=0}^k\mathbf{i}^i\frac{(x^\top y)^i}{i!}\right|\le\min\left\{\frac{(|x|^\top|y|)^{k+1}}{(k+1)!},\frac{2(|x|^\top|y|)^k}{k!}\right\}$$

이다.

Proof 만약 $k=0$이면 임의의 $t\in\mathbb{R}$에 대해 $\cos t\ge 1-|t|$가 성립하여 곧

$$\begin{aligned}|e^{\mathbf{i}x^\top y}-1|&=|\cos x^\top y+\mathbf{i}\sin x^\top y-1|\\&\le|\cos x^\top y-1|+|\mathbf{i}\sin x^\top y|\\&\le|x^\top y|+1\end{aligned}$$

에서 보조정리가 자명하므로 $k>0$이라 하자. 그렇다면 Taylor의 정리로부터

$$\begin{aligned}e^{\mathbf{i}x^\top y}&=\sum_{i=0}^{k-1}\mathbf{i}^i\frac{(x^\top y)^i}{i!}+\mathbf{i}^k\frac{(x^\top y)^k}{(k-1)!}\int_0^1(1-t)^{k-1}e^{\mathbf{i}tx^\top y}\,dt\\&=\sum_{i=0}^{k}\mathbf{i}^i\frac{(x^\top y)^i}{i!}+\mathbf{i}^k\frac{(x^\top y)^k}{(k-1)!}\int_0^1(1-t)^{k-1}(e^{\mathbf{i}tx^\top y}-1)\,dt\end{aligned}$$

이다. 여기서 마지막 등호는 $\int_0^1(1-t)^{k-1}\,dt=1/k$에서 성립한다. 이로부터

$$\left|e^{\mathbf{i}x^\top y}-\sum_{i=0}^k\mathbf{i}^i\frac{(x^\top y)^i}{i!}\right|=\left|\mathbf{i}^{k+1}\frac{(x^\top y)^{k+1}}{k!}\int_0^1(1-t)^k e^{\mathbf{i}tx^\top y}\,dt\right|$$

$$
\begin{aligned}
&\le \frac{(|x|^\top|y|)^{k+1}}{k!}\int_0^1 (1-t)^k|e^{\mathbf{i}tx^\top y}|\,dt\\
&= \frac{(|x|^\top|y|)^{k+1}}{k!}\int_0^1 (1-t)^k\,dt\\
&= \frac{(|x|^\top|y|)^{k+1}}{(k+1)!}
\end{aligned}
$$

이고

$$
\begin{aligned}
\left|e^{\mathbf{i}x^\top y}-\sum_{i=0}^k \mathbf{i}^i\frac{(x^\top y)^i}{i!}\right| &= \left|\mathbf{i}^k\frac{(x^\top y)^k}{(k-1)!}\int_0^1 (1-t)^{k-1}(e^{\mathbf{i}tx^\top y}-1)\,dt\right|\\
&\le \frac{(|x|^\top|y|)^k}{(k-1)!}\int_0^1 (1-t)^{k-1}|e^{\mathbf{i}tx^\top y}-1|\,dt\\
&\le \frac{(|x|^\top|y|)^k}{(k-1)!}\int_0^1 (1-t)^{k-1}(|e^{\mathbf{i}tx^\top y}|+1)\,dt\\
&= \frac{(|x|^\top|y|)^k}{(k-1)!}\int_0^1 2(1-t)^{k-1}\,dt\\
&= \frac{2(|x|^\top|y|)^k}{k!}
\end{aligned}
$$

이므로

$$
\left|e^{\mathbf{i}x^\top y}-\sum_{i=0}^k \mathbf{i}^i\frac{(x^\top y)^i}{i!}\right| \le \min\left\{\frac{(|x|^\top|y|)^{k+1}}{(k+1)!}, \frac{2(|x|^\top|y|)^k}{k!}\right\}
$$

이다. □

Theorem 2.77 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}^n$에 대해 M_X가 $\mathbb{R}^n$ 위에서 존재하면 임의의 $t\in\mathbb{R}^n$에 대해 $\varphi_X(t)=\sum_{i=0}^\infty \mathbf{i}^i\sum_{|\alpha|=i}\mathbf{E}(X^\alpha)t^\alpha/\alpha! = \sum_{\alpha\in\mathbb{N}_0^n}\mathbf{i}^{|\alpha|}\mathbf{E}(X^\alpha)t^\alpha/\alpha!$이다. 여기서 급수는 절대수렴하여 well-define된다.

Proof MGF M_X가 $\mathbb{R}^n$ 위에서 존재하므로 보조정리 ? ? 로부터 임의의 $t\in\mathbb{R}^n$에 대해 $\mathbf{E}(e^{|t|^\top|X|})$가 존재한다. 또한, 임의의 $t\in\mathbb{R}^n$에 대해 $e^{t^\top|X|}\le e^{|t|^\top|X|}$이므로 이상으로부터 $M_{|X|}$가 $\mathbb{R}^n$ 위에서 존재함을 안다. 한편, 위의 보조정리로부터 임의의 $t\in\mathbb{R}^n$와 임의의 $k\in\mathbb{N}$를 택하면

$$
\begin{aligned}
&\left|\varphi_X(t)-\sum_{i=0}^k \mathbf{i}^i\sum_{|\alpha|=i}\mathbf{E}(X^\alpha)\frac{t^\alpha}{\alpha!}\right|\\
&\qquad= \left|\varphi_X(t)-\sum_{i=0}^k\frac{\mathbf{i}^i}{i!}\sum_{|\alpha|=i}\binom{i}{\alpha}\mathbf{E}((t_iX_i)^\alpha)\right|
\end{aligned}
$$

$$
\begin{aligned}
&= \left| \varphi_X(t) - \sum_{i=0}^{k} \mathbf{i}^i \frac{\mathbf{E}((t^\top X)^i)}{i!} \right| \\
&= \left| \mathbf{E}\left(e^{\mathbf{i}t^\top X} - \sum_{i=0}^{k} \mathbf{i}^i \frac{(t^\top X)^i}{i!} \right) \right| \\
&\le \mathbf{E}\left(\left| e^{\mathbf{i}t^\top X} - \sum_{i=0}^{k} \mathbf{i}^i \frac{(t^\top X)^i}{i!} \right| \right) \\
&\le \min\left\{ \frac{\mathbf{E}((|t|^\top |X|)^{i+1})}{(i+1)!}, \frac{2\mathbf{E}((|t|^\top |X|)^i)}{i!} \right\} \\
&= \min\left\{ \frac{1}{(i+1)!} \sum_{|\alpha|=i+1} \binom{i+1}{\alpha} \mathbf{E}(|X|^\alpha)|t|^\alpha, \frac{2}{i!} \sum_{|\alpha|=i} \binom{i}{\alpha} \mathbf{E}(|X|^\alpha)|t|^\alpha \right\}
\end{aligned}
$$

인데, 이미 $\mathbb{R}^n$ 위에서 $M_{|X|}$가 존재함을 알고 있으므로 정리 ? ? 로부터 $k \to \infty$이면 우변의 두 항이 모두 0으로 수렴하여 $\varphi_X(t) = \sum_{i=0}^{\infty} \mathbf{i}^i \sum_{|\alpha|=i} \mathbf{E}(X^\alpha)t^\alpha/\alpha!$이다. 한편, 다시 정리 ? ? 로부터 이 급수가 절대수렴함을 알고, 곧 합의 순서를 바꿀 수 있으므로 원하는 형태의 식을 얻는다. □

Theorem 2.78 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 $\alpha \in \mathbb{N}_0^n$에 대해 $\mathbf{E}(X^\alpha)$가 존재한다고 하자. 그렇다면 임의의 $t \in \mathbb{R}^n$에 대해 $\mathbf{E}(X^\alpha e^{\mathbf{i}t^\top X})$도 존재하고 $\partial^\alpha \varphi_X(t) = \mathbf{i}^{|\alpha|}\mathbf{E}(X^\alpha e^{\mathbf{i}t^\top X})$이다.

Proof 간결한 논의를 위해 $\alpha_n \neq 0$이라 하자. ($\alpha_n = 0$인 경우에도 이와 비슷하게 하면 된다.) 임의의 $t \in \mathbb{R}^n$를 고정하고 함수 $f : \Omega \times \mathbb{R}$를 $f : (\omega, h) \mapsto e^{\mathbf{i}(t+h\mathbf{e}_n)^\top X(\omega)}$로 두면 임의의 $(\omega, h) \in \Omega \times \mathbb{R}$에 대해 $|(\partial/\partial h)f(\omega, h)| = |\mathbf{i}X_n(\omega)e^{\mathbf{i}(t+h\mathbf{e}_n)^\top X(\omega)}| = |X_n(\omega)|$이고 $\mathbf{E}(X_n)$이 존재하므로 f가 Leibniz의 법칙의 모든 조건을 만족시킨다. 따라서 함수 $h \mapsto \int_\Omega e^{\mathbf{i}(t+h\mathbf{e}_n)^\top X}\, d\mathbb{P} = \varphi_X(t + h\mathbf{e}_n)$는 미분가능하고 $\mathbf{E}(X_n e^{\mathbf{i}t^\top X})$가 존재하며, $(\partial/\partial t_n)\varphi_X(t) = (d/dh)\varphi_X(t + h\mathbf{e}_n)|_{h=0} = \int_\Omega \mathbf{i}X_n e^{\mathbf{i}t^\top X}\, d\mathbb{P} = \mathbf{i}\mathbf{E}(X_n e^{\mathbf{i}t^\top X})$이다. 이제 이와 비슷하게 $|\alpha| - 1$번 반복하면 원하는 결과를 얻는다. □

Corollary 2.79 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 $\alpha \in \mathbb{N}_0^n$에 대해 $\mathbf{E}(X^\alpha)$가 존재한다고 하자. 그렇다면 $\mathbf{E}(X^\alpha) = \partial^\alpha \varphi_X(0)/\mathbf{i}^{|\alpha|}$이다.

Proof 이는 위의 정리로부터 자명하다. □

앞서 적률을 도입하며 이를 이용하면 분포의 정보를 손실 없이 전달할 수 있다고 하였는데, 지금부터는 이에 대해 조금 자세히 알아보도록 하자. 분포의 정보를 손실 없이 전달할 수 있다고 함은 곧 적률만 알면 그 분포를 다시 복원해 낼 수 있음을 의미하는데, 기껏해야 가산개인 적률로써 $\mathscr{B}_n$ 위의 측도인 분포를 유일하게 특정해 내는 것은 결코 쉬운 일이 아니며, 이와 같은 문제를 흔히 **problem of moments**라 한다. 다음의 정리는 사실상 Fourier 역변환 공식으로 이 문제를 공략하는 좋은 출발점이 된다.

Theorem 2.80 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 유계인 $B = (x, y] \in \mathscr{S}_n$에 대해 $\mathbf{P}_X(\partial B) = 0$이면

$$\mathbf{P}_X(B) = \lim_{M\to\infty} \frac{1}{(2\pi)^n} \int_{[-M,M]^n} \left(\prod_{i=1}^{n} \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} \right) \varphi_X(t)\, d\mu_n(t)$$

이다. [4]

Proof 임의의 $M > 0$을 택하여 $\mathscr{A} = \mathscr{B}_n|_{[-M,M]^n}$이라 하면 $\mathbf{P}_X \otimes \mu_n|_{\mathscr{A}}$가 유한하고, 모든 성분이 0이 아닌 임의의 $t \in \mathbb{R}^n$와 임의의 $u \in \mathbb{R}^n$에 대해

$$\begin{aligned}\left(\prod_{i=1}^{n} \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} \right) e^{\mathbf{i}t^\top u} &= \prod_{i=1}^{n} \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} e^{\mathbf{i}t_i u_i} \\ &= \prod_{i=1}^{n} \frac{e^{\mathbf{i}t_i(u_i - x_i)} - e^{\mathbf{i}t_i(u_i - y_i)}}{\mathbf{i}t_i}\end{aligned}$$

이므로 곧

$$\begin{aligned}\left| \left(\prod_{i=1}^{n} \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} \right) e^{\mathbf{i}t^\top u} \right| &= \prod_{i=1}^{n} \frac{|e^{\mathbf{i}t_i(u_i - x_i)} - e^{\mathbf{i}t_i(u_i - y_i)}|}{|\mathbf{i}t_i|} \\ &= \prod_{i=1}^{n} \frac{|e^{\mathbf{i}t_i(u_i - x_i)}||1 - e^{\mathbf{i}t_i(x_i - y_i)}|}{|t_i|} \\ &\le \prod_{i=1}^{n} \frac{|t_i(x_i - y_i)|}{|t_i|} \\ &= \prod_{i=1}^{n} (y_i - x_i) \\ &= \mu_n(B) \\ &< \infty\end{aligned}$$

가 되어 함수 $(u, t) \mapsto [\prod_{i=1}^{n}(e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i})/\mathbf{i}t_i] e^{\mathbf{i}t^\top u}$는 $\mathbf{P}_X \otimes \mu_n|_{\mathscr{A}}$-적분가능하다. 여기서 첫번째 부등호는 보조정리 ? ? 로부터 성립한다. 그렇다면 Fubini의 정리로부터

$$\begin{aligned}&\int_{[-M,M]^n} \left(\prod_{i=1}^{n} \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} \right) \varphi_X(t)\, d\mu_n(t) \\ &\qquad = \int_{([-M,M]\setminus\{0\})^n} \left(\prod_{i=1}^{n} \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} \right) \varphi_X(t)\, d\mu_n(t) \\ &\qquad = \int_{([-M,M]\setminus\{0\})^n} \left(\prod_{i=1}^{n} \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} \right) \varphi_X(t)\, d\mu_n|_{\mathscr{A}}(t) \\ &\qquad = \int_{\mathbb{R}^n \times ([-M,M]\setminus\{0\})^n} \left(\prod_{i=1}^{n} \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} \right) e^{\mathbf{i}t^\top u}\, d(\mathbf{P}_X \otimes \mu_n|_{\mathscr{A}})(u, t)\end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^n} \left(\int_{([-M,M]\setminus\{0\})^n} \prod_{i=1}^{n} \frac{e^{\mathbf{i}t_i(u_i-x_i)} - e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\, d\mu_n|_{\mathscr{A}}(t) \right) d\mathbf{P}_X(u) \\
&= \int_{\mathbb{R}^n} \left(\int_{([-M,M]\setminus\{0\})^n} \prod_{i=1}^{n} \frac{e^{\mathbf{i}t_i(u_i-x_i)} - e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\, d\mu_n(t) \right) d\mathbf{P}_X(u) \\
&= \int_{\mathbb{R}^n} \prod_{i=1}^{n} \left(\int_{[-M,M]\setminus\{0\}} \frac{e^{\mathbf{i}t_i(u_i-x_i)} - e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\, d\mu_1(t_i) \right) d\mathbf{P}_X(u)
\end{aligned}
$$

인데, 여기서 각 $i \le n$에 대해

$$
\begin{aligned}
&\int_{[-M,M]\setminus\{0\}} \frac{e^{\mathbf{i}t_i(u_i-x_i)} - e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\, d\mu_1(t_i) \\
&\qquad = \int_{[-M,M]\setminus\{0\}} \frac{\cos t(u_i-x_i) + \mathbf{i}\sin t(u_i-x_i) - \cos t(u_i-y_i) - \mathbf{i}\sin t(u_i-y_i)}{\mathbf{i}t}\, d\mu_1(t) \\
&\qquad = 2\int_{(0,M]} \frac{\sin t(u_i-x_i)}{t}\, d\mu_1(t) - 2\int_{(0,M]} \frac{\sin t(u_i-y_i)}{t}\, d\mu_1(t) \\
&\qquad = 2\operatorname{sgn}(u_i-x_i)\int_{(0,M|u_i-x_i|]} \frac{\sin t}{t}\, d\mu_1(t) - 2\operatorname{sgn}(u_i-y_i)\int_{(0,M|u_i-y_i|]} \frac{\sin t}{t}\, d\mu_1(t)
\end{aligned}
$$

이다. 표기의 편의를 위해 함수 $\mathrm{Si}: \mathbb{R}_0^+ \to \mathbb{R}$를 $\mathrm{Si}: x \mapsto \int_{(0,x]} \sin t/t\, d\mu_1(t)$로 두면 이는 FTC로부터 연속이고 $x \to \infty$일 때 $\pi/2$로 수렴하므로 곧 Si는 적당한 $N > 0$에 대해 유계이다. 이로부터 $M_j \to \infty$인 임의의 수열 $\{M_j\}$를 생각하여 $j \in \mathbb{N}$가 충분히 크다고 하면

$$
\begin{aligned}
&\left| \int_{[-M_j,M_j]} \frac{e^{\mathbf{i}t_i(u_i-x_i)} - e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\, d\mu_1(t_i) \right| \\
&\qquad\qquad = |2\operatorname{sgn}(u_i-x_i)\mathrm{Si}(M_j|u_i-x_i|) - 2\operatorname{sgn}(u_i-y_i)\mathrm{Si}(M_j|u_i-y_i|)| \\
&\qquad\qquad \le 2|\mathrm{Si}(M_j|u_i-x_i|)| + 2|\mathrm{Si}(M_j|u_i-y_i|)| \\
&\qquad\qquad \le 4N
\end{aligned}
$$

이 되어 DCT에서

$$
\begin{aligned}
&\lim_{j\to\infty} \int_{[-M_j,M_j]^n} \left(\prod_{i=1}^{n} \frac{e^{-\mathbf{i}t_ix_i} - e^{-\mathbf{i}t_iy_i}}{\mathbf{i}t_i} \right) \varphi_X(t)\, d\mu_n(t) \\
&\qquad = \int_{\mathbb{R}^n} \lim_{j\to\infty} \prod_{i=1}^{n} \left(\int_{[-M_j,M_j]} \frac{e^{\mathbf{i}t_i(u_i-x_i)} - e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\, d\mu_1(t_i) \right) d\mathbf{P}_X(u) \\
&\qquad = \int_{\mathbb{R}^n} \lim_{j\to\infty} \prod_{i=1}^{n} [2\operatorname{sgn}(u_i-x_i)\mathrm{Si}(M_j|u_i-x_i|) - 2\operatorname{sgn}(u_i-y_i)\mathrm{Si}(M_j|u_i-y_i|)]\, d\mathbf{P}_X(u) \\
&\qquad = \int_{\mathbb{R}^n} \prod_{i=1}^{n} \pi[\operatorname{sgn}(u_i-x_i) - \operatorname{sgn}(u_i-y_i)]\, d\mathbf{P}_X(u)
\end{aligned}
$$

$$= \int_{B^\circ} (2\pi)^n \, d\mathbf{P}_X(u) + \int_{\partial B} \prod_{i=1}^n \pi[\operatorname{sgn}(u_i - x_i) - \operatorname{sgn}(u_i - y_i)] \, d\mathbf{P}_X(u)$$
$$= (2\pi)^n \mathbf{P}_X(B)$$

이고, 증명이 끝난다. 여기서 마지막 등호는 $\mathbf{P}_X(\partial B) = 0$이므로 성립한다. □

위의 정리는 CF를 알면 이를 통해 그 분포를 (거의) 알 수 있음을 의미한다. 여기서 이론적인 기교를 조금 더 부리면 다음 따름정리를 얻는다.

Lemma 2.81 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 $\mathbb{R}^n$의 각 축에 수직인 초평면 중에 $\mathbf{P}_X$−영집합이 아닌 것은 가산개이다.

Proof 임의의 $i \le n$에 대해 CDF F_{X_i}가 $x_0 \in \mathbb{R}$에서 연속이라면 $\mathbf{P}_X(\mathbb{R}^{i-1} \times \{x_0\} \times \mathbb{R}^{n-i}) = \mathbb{P}\{X_i = x_0\} = \mathbf{P}_{X_i}\{x_0\} = F_{X_i}(x_0) - F_{X_i}(x_0-) = 0$이므로 초평면 $\pi : x_i = x_0$는 $\mathbf{P}_X$−영집합이다. 그런데 정리 ? ? 로부터 각 $i \le n$에 대해 F_{X_i}는 가산개의 불연속점만을 가지고, 곧 보조정리가 성립한다. □

Corollary 2.82 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$, $(\Omega', \mathscr{F}', \mathbb{P}')$에서 각각 정의된 rv. $X : \Omega \to \mathbb{R}^n$, $Y : \Omega' \to \mathbb{R}^n$에 대해 $X \equiv Y$일 필요충분조건은 $\varphi_X = \varphi_Y$인 것이다.

Proof 만약 $X \equiv Y$이면 $\mathbf{P}_X = \mathbf{P}_Y$에서 $\varphi_X = \varphi_Y$임이 자명하므로 역만 성립함을 보이면 된다. 이를 위해 $\varphi_X = \varphi_Y$라 하면 위의 정리로부터 유계인 $B = (x, y] \in \mathscr{S}_n$에 대해 $\mathbf{P}_X(\partial B) = \mathbf{P}_Y(\partial B) = 0$이면

$$\mathbf{P}_X(B) = \lim_{M\to\infty} \frac{1}{(2\pi)^n} \int_{[-M,M]^n} \left(\prod_{i=1}^n \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} \right) \varphi_X(t) \, d\mu_n(t)$$
$$= \lim_{M\to\infty} \frac{1}{(2\pi)^n} \int_{[-M,M]^n} \left(\prod_{i=1}^n \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} \right) \varphi_Y(t) \, d\mu_n(t)$$
$$= \mathbf{P}_Y(B)$$

이다. 따라서 만약 집합족 $\mathscr{A} = \{B \in \mathscr{S}_n : \mathbf{P}_X(\partial B) = \mathbf{P}_Y(\partial B) = 0\}$가 π−system이고 $\mathscr{B}_n$의 생성자이면 정리 ? ? 로부터 증명이 끝난다. 우선 임의의 $B, B' \in \mathscr{A}$를 생각하면 $\partial(B \cap B') \subseteq \partial B \cup \partial B'$이므로 $\mathscr{A}$가 π−system임은 거의 자명하다. 한편, 임의의 $B = (x, y] \in \mathscr{S}_n$를 택하면 위의 보조정리로부터 $\mathbf{P}_X$−영집합이 아니거나 $\mathbf{P}_Y$−영집합이 아닌 $\mathbb{R}^n$의 축에 수직인 초평면이 가산개이므로 적당한 수열 $\{z_j\}, \{w_j\}$가 존재하여 각 $j \in \mathbb{N}$에 대해 $\mathbf{P}_X(\partial(z_j, w_j]) = \mathbf{P}_Y(\partial(z_j, w_j]) = 0$이고 $z_j \downarrow x, w_j \downarrow y$이다. 이제 각 $j \in \mathbb{N}$에 대해 $B_j = (z_j, w_j] \in \mathscr{A}$로 두면 $B_j \to B$이므로 $B \in \sigma(\mathscr{A})$이고, 곧 $\mathscr{S}_n \subseteq \sigma(\mathscr{A})$에서 $\mathscr{B}_n \subseteq \sigma(\mathscr{A})$인데, 그 역의 포함관계는 자명하므로 $\mathscr{A}$가 $\mathscr{B}_n$의 생성자가 되어 증명이 끝난다. □

위의 따름정리에 의하면 CF와 분포는 정확히 일대일로 대응한다. 한편, 앞서 보인 정리 ? ? 에서 확률벡터의 모든 적률을 알면 (MGF가 존재한다는 가정 하에) 그로부터 CF를 구할 수 있으므로 이상을 종합하면 problem of moments에 어느 정도 만족스러운 답이 된다. 즉, 적률을 알면 CF를 알고, 그런 CF를 가지는 분포는 유일하게 특정된다. 한편, MGF와 CF는 이론적으로 그 역할이 비슷하므로 CF 대신 MGF를 이용하는 방법도 생각해 볼 수 있다. 다음 정리와 따름정리는 이에 대한 이론적인 전개를 담고 있다.

Theorem 2.83 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 이의 모든 적률이 존재하고 적당한 0의 근방 $U \subseteq \mathbb{R}^n$가 존재하여 임의의 $t \in U$에 대해 급수 $\sum_{i=0}^{\infty} \sum_{|\alpha|=i} \mathbf{E}(X^\alpha) t^\alpha / \alpha!$가 절대수렴한다고 하자. 이제 확률공간 $(\Omega', \mathscr{F}', \mathbb{P}')$에서 정의된 rv. $Y : \Omega' \to \mathbb{R}^n$에 대해 Y의 모든 적률이 존재하여 X의 적률과 같다면 $X \equiv Y$이다.

Proof 먼저 $|\alpha| \to \infty$이면 모든 성분이 양수인 임의의 $h \in U$에 대해 $\mathbf{E}(|X|^\alpha) h^\alpha / \alpha! \to 0$임을 보이자. 간결한 논의를 위해 $n = 2$라 하고 모든 성분이 1보다 작은 양수인 임의의 $h \in U$를 택하면 $\sum_{i=0}^{\infty} \sum_{|\alpha|=i} \mathbf{E}(X^\alpha) h^\alpha / \alpha!$가 절대수렴하므로 $i \to \infty$이면 $\sum_{|\alpha|=i} |\mathbf{E}(X^\alpha) h^\alpha / \alpha!| \to 0$이고, 곧 $|\alpha| \to \infty$이면 $\mathbf{E}(X^\alpha) h^\alpha / \alpha! \to 0$이다. ($n = 1$인 경우를 포함한 일반적인 경우에도 이와 비슷하게 하면 된다.) 이로부터 적당한 $M > 0$이 존재하여 임의의 $\alpha \in \mathbb{N}_0^2$에 대해 $|\mathbf{E}(X^\alpha) h^\alpha / \alpha!| \le M$이고, 임의의 $\varepsilon > 0$을 택하면 적당한 $i_0 \in \mathbb{N}$가 존재하여 $|\alpha| \ge i_0$이면 $|\mathbf{E}(X^\alpha) h^\alpha / \alpha!| < h_1 h_2 \varepsilon / 16$이다. 이제 $|\alpha| \ge \max\{K_0 + 2, K_0 + 8M / h_1 h_2 \varepsilon, 8/\varepsilon\}$인 임의의 $\alpha \in \mathbb{N}_0^2$를 택하고 경우를 나누어 생각한다.

만약 α_1, α_2가 모두 짝수라면 $\mathbf{E}(|X|^\alpha) h^\alpha / \alpha! < \varepsilon$임은 자명하다. 만약 α_1은 홀수이지만 α_2는 짝수라면 적당한 $\beta_1 \in \mathbb{N}$에 대해 $\alpha_1 = 2\beta_1 - 1$로 쓸 수 있고, 임의의 $x \in \mathbb{R}$에 대해 $|x|^{\alpha_1} \le x^{2\beta_1} \mathbf{1}_{\{|x| \ge 1\}} + \mathbf{1}_{\{|x| < 1\}} \le x^{2\beta_1} + 1$이므로

$$
\begin{aligned}
\left| \mathbf{E}(|X|^\alpha) \frac{h^\alpha}{\alpha!} \right| &= \left[\int_{\mathbb{R}^n} |x_1|^{\alpha_1} x_2^{\alpha_2} \, d\mathbf{P}_X(x) \right] \frac{h^\alpha}{\alpha!} \\
&\le \left[\int_{\mathbb{R}^n} (x_1^{2\beta_1} + 1) x_2^{\alpha_2} \, d\mathbf{P}_X(x) \right] \frac{h_1^{\alpha_1} h_2^{\alpha_2}}{\alpha_1! \alpha_2!} \\
&= \frac{2\beta_1}{\alpha_1 h_1} \mathbf{E}(X_1^{2\beta_1} X_2^{\alpha_2}) \frac{h_1^{2\beta_1} h_2^{\alpha_2}}{(2\beta_1)! \alpha_2!} + \frac{h_1^{\alpha_1}}{\alpha_1!} \mathbf{E}(X_2^{\alpha_2}) \frac{h_2^{\alpha_2}}{\alpha_2!} \\
&\le \frac{2}{h_1} \mathbf{E}(X_1^{2\beta_1} X_2^{\alpha_2}) \frac{h_1^{2\beta_1} h_2^{\alpha_2}}{(2\beta_1)! \alpha_2!} + \frac{1}{\alpha_1} \mathbf{E}(X_2^{\alpha_2}) \frac{h_2^{\alpha_2}}{\alpha_2!}
\end{aligned}
$$

이다. 그렇다면 $2\beta_1 + \alpha_2 \ge i_0 + 1$에서 $(2/h_1) \mathbf{E}(X_1^{2\beta_1} X_2^{\alpha_2}) h_1^{2\beta_1} h_2^{\alpha_2} / (2\beta_1)! \alpha_2! < h_2 \varepsilon / 8 < \varepsilon / 2$이고 $\alpha_2 \ge i_0$인 경우에는 $(1/\alpha_1) \mathbf{E}(X_2^{\alpha_2}) h_2^{\alpha_2} / \alpha_2! < h_1 h_2 \varepsilon / 16\alpha_1 < \varepsilon / 2$이며 그렇지 않은 경우에는 $\alpha_1 > 8M / h_1 h_2 \varepsilon > 2M / \varepsilon$에서 $(1/\alpha_1) \mathbf{E}(X_2^{\alpha_2}) h_2^{\alpha_2} / \alpha_2! < M / \alpha_1 < \varepsilon / 2$이다. 이상으로부터 α_1은 홀수이지만 α_2는 짝수라면 $\mathbf{E}(|X|^\alpha) h^\alpha / \alpha! < \varepsilon$임을 안다. 반대로 α_1은 짝수이지

만 α_2는 홀수인 경우에도 위와 비슷하게 하면 같은 결론을 얻는다. 마지막으로 α_1, α_2 모두 홀수인 경우에는 위의 과정을 비슷하게 한 번 더 반복하면 되는데, 적당한 $\beta \in \mathbb{N}^2$에 대해 $\alpha = 2\beta - 1$이라 하면 위의 결과로부터

$$\begin{aligned}\mathbf{E}(|X|^{\alpha})\frac{h^{\alpha}}{\alpha!} &\le \frac{2}{h_1}\mathbf{E}(X_1^{2\beta_1}|X_2|^{\alpha_2})\frac{h_1^{2\beta_1}h_2^{\alpha_2}}{(2\beta_1)!\alpha_2!} + \frac{1}{\alpha_1}\mathbf{E}(|X_2|^{\alpha_2})\frac{h_2^{\alpha_2}}{\alpha_2!} \\ &\le \frac{4}{h_1h_2}\mathbf{E}(X^{2\beta})\frac{h^{2\beta}}{(2\beta)!} + \frac{2}{h_1\alpha_2}\mathbf{E}(X_1^{2\beta_1})\frac{h_1^{2\beta_1}}{(2\beta_1)!} + \frac{2}{h_2\alpha_1}\mathbf{E}(X_2^{2\beta_2})\frac{h_2^{2\beta_2}}{(2\beta_2)!} + \frac{1}{\alpha_1\alpha_2}\end{aligned}$$

임을 쉽게 보일 수 있다. 그렇다면 $|2\beta| \ge i_0$에서 $(4/h_1h_2)\mathbf{E}(X^{2\beta})h^{2\beta}/(2\beta)! < \varepsilon/4$이고 $2\beta_1 \ge i_0$인 경우에는 $(2/h_1\alpha_2)\mathbf{E}(X_1^{2\beta_1})h_1^{2\beta_1}/(2\beta_1)! < \varepsilon/8\alpha_2 < \varepsilon/4$이며 그렇지 않은 경우에는 $\alpha_2 \ge 8M/h_1h_2\varepsilon > 8M/h_1\varepsilon$에서 $(2/h_1\alpha_2)\mathbf{E}(X_1^{2\beta_1})h_1^{2\beta_1}/(2\beta_1)! < 2M/h_1\alpha_2 < \varepsilon/4$이다. 비슷하게 $(2/h_2\alpha_1)\mathbf{E}(X_2^{2\beta_2})h_2^{2\beta_2}/(2\beta_2)! < \varepsilon/4$임을 알고, $1/\alpha_1\alpha_2 \le 1/\max\{\alpha_1, \alpha_2\} \le \varepsilon/4$이므로 이상으로부터 α_1과 α_2가 모두 홀수라도 $\mathbf{E}(|X|^{\alpha})h^{\alpha}/\alpha! < \varepsilon$임을 안다. 곧 어떠한 경우에도 $\mathbf{E}(|X|^{\alpha})h^{\alpha}/\alpha! < \varepsilon$이 되어 $|\alpha| \to \infty$이면 $\mathbf{E}(|X|^{\alpha})h^{\alpha}/\alpha! \to 0$임을 안다. 나아가, $h \in U$가 0인 성분을 가지더라도 이가 성립함을 위와 비슷하게 보일 수 있다.

이제 보조정리 ? ? 로부터 임의의 $k \in \mathbb{N}$과 임의의 $t \in \mathbb{R}^n$, 임의의 $h \in U$에 대해

$$\begin{aligned}&\left|\varphi_X(t+h) - \sum_{i=0}^{k}\frac{\mathbf{i}^i}{i!}\int_{\mathbb{R}^n}(h^\top x)^i e^{\mathbf{i}t^\top x}\,d\mathbf{P}_X(x)\right| \\ &\qquad = \left|\mathbf{E}(e^{\mathbf{i}(t+h)^\top X}) - \sum_{i=0}^{k}\frac{1}{i!}\mathbf{E}((\mathbf{i}h^\top X)^i e^{\mathbf{i}t^\top X})\right| \\ &\qquad = \left|\mathbf{E}\left(e^{\mathbf{i}(t+h)^\top X} - \sum_{i=0}^{k}\frac{1}{i!}(\mathbf{i}h^\top X)^i e^{\mathbf{i}t^\top X}\right)\right| \\ &\qquad \le \mathbf{E}\left(\left|e^{\mathbf{i}(t+h)^\top X} - \sum_{i=0}^{k}\frac{1}{i!}(\mathbf{i}h^\top X)^i e^{\mathbf{i}t^\top X}\right|\right) \\ &\qquad = \mathbf{E}\left(|e^{\mathbf{i}t^\top X}|\left|e^{\mathbf{i}h^\top X} - \sum_{i=0}^{k}\frac{1}{i!}(\mathbf{i}h^\top X)^i\right|\right) \\ &\qquad = \mathbf{E}\left(\left|e^{\mathbf{i}h^\top X} - \sum_{i=0}^{k}\frac{1}{i!}(\mathbf{i}h^\top X)^i\right|\right) \\ &\qquad \le \mathbf{E}\left(\frac{(|h|^\top|X|)^k}{k!}\right) \\ &\qquad = \frac{2}{k!}\sum_{|\alpha|=k}\binom{k}{\alpha}\mathbf{E}(|(h_iX_i)^{\alpha}|) \\ &\qquad = 2\sum_{|\alpha|=k}\mathbf{E}(|X|^{\alpha})\frac{h^{\alpha}}{\alpha!}\end{aligned}$$

이고 앞선 결론으로부터 $k \to \infty$이면 위 식의 우변이 0으로 수렴하여 정리 ? ? 로부터

$$\begin{aligned}
\varphi_X(t+h) &= \sum_{i=0}^{\infty} \frac{\mathbf{i}^i}{i!} \int_{\mathbb{R}^n} (h^\top x)^i e^{\mathbf{i}t^\top x} d\mathbf{P}_X(x) \\
&= \sum_{i=0}^{\infty} \frac{1}{i!} \sum_{|\alpha|=i} \binom{i}{\alpha} \mathbf{i}^i \int_{\mathbb{R}^n} (h_i x_i)^\alpha e^{\mathbf{i}t^\top x} d\mathbf{P}_X(x) \\
&= \sum_{i=0}^{\infty} \frac{1}{i!} \sum_{|\alpha|=i} \binom{i}{\alpha} \mathbf{i}^i \mathbf{E}(X^\alpha e^{\mathbf{i}t^\top X}) h^\alpha \\
&= \sum_{i=0}^{\infty} \frac{1}{i!} \sum_{|\alpha|=i} \binom{i}{\alpha} \partial^\alpha \varphi_X(t) h^\alpha \\
&= \sum_{i=0}^{\infty} \frac{1}{i!} \mathbf{D}^i \varphi_X(t)(h, \cdots, h)
\end{aligned}$$

이다. 이는 Y에 대해서도 성립하므로 임의의 $t \in \mathbb{R}^n$와 임의의 $h \in U$에 대해 $\varphi_Y(t+h) = \sum_{i=0}^{\infty} \mathbf{D}^i \varphi_Y(t)(h, \cdots, h)/i!$인데, X와 Y의 적률이 모두 같으므로 다시 정리 ? ? 로부터 임의의 $i \in \mathbb{N}_0$에 대해 $\mathbf{D}^i \varphi_X(0) = \mathbf{D}^i \varphi_Y(0)$이고, 곧 U에서 φ_X와 φ_Y가 일치한다. 그런데 위의 결론에서 $t \in \mathbb{R}^n$가 임의의 점이었음을 상기한다면 이는 곧 $\varphi_X = \varphi_Y$임을 의미하여 곧 따름정리 ? ? 로부터 $X \equiv Y$이다. □

Corollary 2.84 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$, $(\Omega', \mathscr{F}', \mathbb{P}')$에서 각각 정의된 rv. $X : \Omega \to \mathbb{R}^n$, $Y : \Omega' \to \mathbb{R}^n$에 대해 M_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재한다고 하자. 그렇다면 $X \equiv Y$일 필요충분조건은 M_Y가 적당한 0의 근방 $V \subseteq \mathbb{R}^n$ 위에서 존재하여 $U \cap V$에서 $M_X = M_Y$인 것이다.

Proof 만약 $X \equiv Y$이면 $\mathbf{P}_X = \mathbf{P}_Y$에서 정리가 자명하므로 역만 성립함을 보이면 된다. 이를 위해 M_Y가 적당한 0의 근방 $V \subseteq \mathbb{R}^n$ 위에서 존재하여 $U \cap V$에서 $M_X = M_Y$라 하자. 그렇다면 정리 ? ? 와 ? ? 로부터 X와 Y의 모든 적률이 존재하여 서로 같으며 $\sum_{i=0}^{\infty} \sum_{|\alpha|=i} \mathbf{E}(X^\alpha) t^\alpha / \alpha!$가 적당한 0의 근방에서 절대수렴하므로 위의 정리로부터 $X \equiv Y$이다. □

Problem of moments에 대한 논의를 마무리하기 전에, 한 가지 유의해야 할 점이 있다. 우리가 적률을 통해 MGF나 CF를 계산할 수 있고, 곧 분포를 유일하게 특정할 수 있지만, 이는 어디까지나 MGF가 존재할 때에만 가능한 것이다. 따라서 어떤 확률벡터의 MGF가 존재하지 않는다면 적률을 모두 알아도 이로부터 분포를 특정지을 수가 없다. 즉, 그러한 적률을 만들어내는 서로다른 분포가 얼마든지 존재할 수 있다.

다음으로 살펴볼 CF에 대한 흥미로운 주제는 연속확률벡터와 CF의 관계이다. 앞서 정리 ? ? 는 어떤 확률벡터가 연속인지를 판단할 수 있는 방법을 제시했는데, 다음 정리는 CF를 이용한 새로운 방법을 하나 제시한다.

Theorem 2.85 (Riemann-Lebesgue) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 연속활률벡터 $X:\Omega \to \mathbb{R}^n$에 대해 $\lim_{|t_1|,\cdots,|t_n|\to\infty} \varphi_X(t) \to 0$이다.

PROOF 임의의 $\varepsilon > 0$을 택하면 f_X가 μ_n−적분가능하므로 정리 ? ? 의 i로부터 적당한 단순함수 $g:\mathbb{R}^n \to \mathbb{R}$가 존재하여 이는 적당한 $a_1, \cdots, a_k \in \mathbb{R}$와 유계인 $B_1, \cdots, B_k \in \mathscr{S}_n$에 대해 $g = \sum_{i=1}^n a_i \mathbf{1}_{B_i}$로 쓸 수 있으며 $\int_{\mathbb{R}^n} |f_X - g|\, d\mu_n < \varepsilon$이고, 곧 임의의 $t \in \mathbb{R}^n$에 대해

$$\begin{aligned}\left|\int_{\mathbb{R}^n} [f_X(x) - g(x)] e^{\mathbf{i}t^\top x}\, d\mu_n(x)\right| &\le \int_{\mathbb{R}^n} |[f_X(x) - g(x)] e^{\mathbf{i}t^\top x}|\, d\mu_n(x) \\ &= \int_{\mathbb{R}^n} |f_X(x) - g(x)|\, d\mu_n(x) \\ &< \frac{\varepsilon}{2}\end{aligned}$$

이다. 이제 각 $i \le k$에 대해 적당한 $x_i, y_i \in \mathbb{R}^n$에 대해 $B_i = \prod_{j=1}^n (x_i^j, y_i^j]$라 하자. 그렇다면 Fubini의 정리로부터 0을 성분으로 갖지 않는 임의의 $t \in \mathbb{R}^n$에 대해

$$\begin{aligned}\left|\int_{\mathbb{R}^n} g(x) e^{\mathbf{i}t^\top x}\, d\mu_n(x)\right| &= \left|\int_{\mathbb{R}^n} \sum_{i=1}^k a_i \mathbf{1}_{B_i} e^{\mathbf{i}t^\top x}\, d\mu_n(x)\right| \\ &= \left|\sum_{i=1}^k a_i \int_{B_i} \prod_{j=1}^n e^{\mathbf{i}t_j x_j}\, d\mu_n(x)\right| \\ &= \left|\sum_{i=1}^k a_i \prod_{j=1}^n \int_{x_i^j}^{y_i^j} e^{\mathbf{i}t_j x_j}\, d\mu_1(x_j)\right| \\ &= \left|\sum_{i=1}^k a_i \prod_{j=1}^n \frac{e^{\mathbf{i}t_j y_i^j} - e^{\mathbf{i}t_j x_i^j}}{\mathbf{i}t_j}\right| \\ &\le \sum_{i=1}^k |a_i| \prod_{j=1}^n \left|\frac{e^{\mathbf{i}t_j y_i^j} - e^{\mathbf{i}t_j x_i^j}}{\mathbf{i}t_j}\right| \\ &\le \sum_{i=1}^k |a_i| \prod_{j=1}^n \frac{|e^{\mathbf{i}t_j y_i^j}| + |e^{\mathbf{i}t_j x_i^j}|}{|t_j|} \\ &= 2^n \sum_{i=1}^k |a_i| \Big/ \prod_{j=1}^n |t_j|\end{aligned}$$

이므로 $\lim_{|t_1|,\cdots,|t_n|\to\infty} |\int_{\mathbb{R}^n} g(x) e^{\mathbf{i}t^\top x}\, d\mu_n(x)| = 0$에서 적당한 $M > 0$이 존재하여 $|t_1|, \cdots, |t_n| \ge M$이면 $|\int_{\mathbb{R}^n} g(x) e^{\mathbf{i}t^\top x}\, d\mu_n(x)| < \varepsilon/2$이다. 이상으로부터 $|t_1|, \cdots, |t_n| \ge M$이면 $|\varphi_X(t)| = |\int_{\mathbb{R}^n} f_X(x) e^{\mathbf{i}t^\top x}\, d\mu_n(x)| \le |\int_{\mathbb{R}^n} [f_X(x) - g(x)] e^{\mathbf{i}t^\top x}\, d\mu_n(x)| + |\int_{\mathbb{R}^n} g(x) e^{\mathbf{i}t^\top x}\, d\mu_n(x)| < \varepsilon$이 되어 증명이 끝난다. □

심지어 CF를 이용해 PDF를 바로 구할 수도 있다.

Theorem 2.86 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 연속확률벡터 $X : \Omega \to \mathbb{R}$에 대해 $\int_{\mathbb{R}^n} |\varphi_X| \, d\mu_n < \infty$이면

$$f_X(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-\mathbf{i}t^\top x} \varphi_X(t) \, d\mu_n(t)$$

이다.

이제 앞서 기댓값에 대한 다양한 부등식을 소개할 때 미처 소개하지 못한 부등식을 하나 소개하고 MGF와 CF에 대한 내용을 마무리하자. 이 부등식도 tail probability를 bounding 하는 부등식이다.

Theorem 2.87 (Chernoff's bound) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 0의 적당한 근방 $U \in \mathbb{R}$에서 M_X가 존재하면 임의의 $x \in \mathbb{R}$에 대해 $\mathbb{P}\{X \geq x\} \leq \inf_{0 \leq t \in U} e^{-xt} M_X(t)$이다. 비슷하게, 임의의 $x \in \mathbb{R}$에 대해 $\mathbb{P}\{X \leq x\} \leq \inf_{0 \geq t \in U} e^{-xt} M_X(t)$이다.

Proof 임의의 $x, y \in \mathbb{R}$와 임의의 음이 아닌 $t \in U$에 대해 $\mathbf{1}_{[x,\infty)}(y) \leq e^{t(y-x)}$이므로

$$\begin{aligned}
\mathbb{P}\{X \geq x\} &= \mathbf{P}_X([x, \infty)) \\
&= \int_{\mathbb{R}} \mathbf{1}_{[x,\infty)} \, d\mathbf{P}_X \\
&\leq \int_{\mathbb{R}} e^{t(y-x)} \, d\mathbf{P}_X(y) \\
&= \mathbf{E}(e^{t(X-x)}) \\
&= e^{-xt} M_X(t)
\end{aligned}$$

에서 $\mathbb{P}\{X \geq x\} \leq \inf_{0 \leq t \in U} e^{-xt} M_X(t)$이고, 이와 비슷하게 $\mathbb{P}\{X \leq x\} \leq \inf_{0 \geq t \in U} e^{-xt} M_X(t)$임을 보일 수 있다. □

이번 절의 후반부에서는 MGF와 관련되거나 혹은 유사한 함수들을 간단히 살펴본다. 먼저 첫째로 볼 것은 MGF에 로그를 취하여 얻는 CGF, 그리고 이의 Taylor 전개의 계수로 주어지는 누율이다.

Definition 2.88 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 M_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재하면 rv. X의 **누율생성함수(cumulant generating function)**를 $C_X : U \to \mathbb{R}$로 쓰고 $C_X = \log M_X$로 정의한다. 특별히, $n \geq 2$인 경우 C_X를 rv. $X_1, \cdots, X_n$의 **결합누율생성함수(joint cumulant generating function)**라 하기도 한다.

Proposition 2.89 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 C_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재한다면 C_X는 U에서 해석적이다.

Proof 가정으로부터 M_X가 U 위에서 존재하고 정리 ? ? 로부터 이는 U에서 해석적이다. 한편, 함수 $x \mapsto \log x$가 $(0, \infty)$에서 해석적이므로 $C_X = \log M_X$도 U에서 해석적이다. □

Definition 2.90 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 C_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재한다고 하자. 이제 $\alpha \in \mathbb{N}_0^n$에 대해 $C_X(t)$의 0에서의 Taylor 전개에서 $t^\alpha/\alpha!$의 계수를 X의 $|\alpha|$**차 (결합)누율**($|\alpha|$**th (joint) cumulant**)이라 하고 $c_{\alpha,X}$ 혹은 간단히 c_α로 쓴다.

Theorem 2.91 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 C_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재한다고 하면 임의의 $\alpha \in \mathbb{N}_0^n$에 대해 $c_{X,\alpha} = \partial^\alpha C_X(0)$이다.

Proof 이는 누율의 정의로부터 자명하다. □

CGF의 기본적인 성질은 MGF의 성질로부터 거의 자명하게 얻어진다.

Theorem 2.92 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 C_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재하면 다음이 성립한다.

i. $C_X(0) = 0$.
ii. CGF C_X는 볼록하다.
iii. 임의의 $a \in \mathbb{R}$와 임의의 $b \in \mathbb{R}^n$에 대해 C_{aX+b}가 적당한 0의 근방 $V \subseteq \mathbb{R}^n$ 위에서 존재하고, $at \in U$인 $t \in V$에 대해 $C_{aX+b}(t) = C_X(at) + t^\mathsf{T} b$이다.

Proof i. 이는 $C_X(0) = \log M_X(0) = 0$에서 자명하다.
ii. 정리 ? ? 의 iii으로부터 이는 자명하다.
iii. 정리 ? ? 의 iv로부터 C_{aX+b}가 적당한 0의 근방 $V \subseteq \mathbb{R}^n$에서 존재하고 $at \in U$인 $t \in V$에 대해 $C_{aX+b}(t) = \log M_{aX+b}(t) = \log e^{t^\mathsf{T} b} M_X(at) = C_X(at) + t^\mathsf{T} b$이다. □

앞서 배운 적률이 확률변수의 거듭제곱의 기댓값으로 깔끔하게 주어진 것에 비해, 누율과 기댓값을 깔끔하게 연결짓는 것은 쉽지 않다. 다만, 몇몇 낮은 차수의 누율의 경우 특별한 의미를 가진다. 그리고 때로는 이러한 관계를 잘 사용하면 기댓값이나 분산 등을 복잡한 계산을 피해 쉽고 간단하게 구할 수 있다.

Proposition 2.93 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 C_X가 적당한 0의 근방 $U \subseteq \mathbb{R}$ 위에서 존재한다고 하면 $c_{X,1} = \mathbf{E}(X)$이고 $c_{X,2} = \mathbf{Var}(X)$이다.

Proof 이는 정리 ? ? 으로부터 $C_X'(0) = (\log M_X)'(0) = (M_X'/M_X)(0) = \mathbf{E}(X)$이고

$$\begin{aligned} C_X''(0) &= (\log M_X)''(0) \\ &= \left(\frac{M_X'}{M_X}\right)'(0) \end{aligned}$$

$$
\begin{aligned}
&= \left[\frac{M_X'' M_X - (M_X')^2}{M_X^2}\right](0) \\
&= \mathbf{E}(X^2) - [\mathbf{E}(X)]^2 \\
&= \mathbf{Var}(X)
\end{aligned}
$$

이므로 자명하다. □

Proposition 2.94 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 C_X가 적당한 0의 근방 $U \subseteq \mathbb{R}$ 위에서 존재하며 $\mathbf{Var}(X) \neq 0$이라 하자. 이제 $Z = (X - \mu_X)/\sigma_X$라 하면 C_Z가 적당한 0의 근방 $V \subseteq \mathbb{R}$ 위에서 존재하고 $c_{Z,1} = 0$, $c_{Z,2} = 1$, $c_{Z,3} = \mathbf{Skew}(X)$, $c_{Z,4} = \mathbf{Kurt}(X)$이다.

Proof 정리 ? ? 의 iii으로부터 적당한 0의 근방 $V \subseteq \mathbb{R}$에서 C_Z가 존재함을 안다. 이제 정리 ? ? 으로부터

$$
\begin{aligned}
&M_Z(0) = 1 \\
&M_Z'(0) = \mathbf{E}\left(\frac{X - \mu_X}{\sigma_X}\right) = \frac{\mathbf{E}(X) - \mu_X}{\sigma_X} = 0 \\
&M_Z''(0) = \mathbf{E}\left(\left(\frac{X - \mu_X}{\sigma_X}\right)^2\right) = \frac{\mathbf{Var}(X)}{\sigma_X^2} = 1, \\
&M_Z'''(0) = \mathbf{E}\left(\left(\frac{X - \mu_X}{\sigma_X}\right)^3\right) = \mathbf{Skew}(X) \\
&M_Z^{(4)}(0) = \mathbf{E}\left(\left(\frac{X - \mu_X}{\sigma_X}\right)^4\right) = \mathbf{Kurt}(X) + 3
\end{aligned}
$$

이다. 그렇다면 명제 ? ? 로부터 $c_{Z,1} = \mathbf{E}(Z) = 0$, $c_{Z,2} = \mathbf{Var}(Z) = 1$이고

$$
\begin{aligned}
C_Z'''(0) &= (\log M_Z)'''(0) \\
&= \left(\frac{M_Z'}{M_Z}\right)''(0) \\
&= \left[\frac{M_Z'' M_Z - (M_Z')^2}{M_Z^2}\right]'(0) \\
&= \left\{\frac{(M_Z''' M_Z + M_Z'' M_Z' - 2M_Z' M_Z'') M_Z^2 - [M_Z'' M_Z - (M_Z')^2] 2 M_Z M_Z'}{M_Z^4}\right\}(0) \\
&= \left[\frac{M_Z''' M_Z^2 - 3 M_Z'' M_Z' M_Z + 2(M_Z')^3}{M_Z^3}\right](0) \\
&= \mathbf{Skew}(X)
\end{aligned}
$$

이며

$$\begin{aligned}
C_Z^{(4)}(0) &= (\log M_Z)^{(4)}(0) \\
&= \left(\frac{M_Z'}{M_Z}\right)'''(0) \\
&= \left[\frac{M_Z''M_Z - (M_Z')^2}{M_Z^2}\right]''(0) \\
&= \left[\frac{M_Z'''M_Z^2 - 3M_Z''M_Z'M_Z + 2(M_Z')^3}{M_Z^3}\right]'(0) \\
&= \left\{\frac{\begin{array}{l}[M_Z^{(4)}M_Z^2 + M_Z'''2M_ZM_Z' - 3M_Z'''M_Z'M_Z - 3(M_Z'')^2M_Z \\ \quad - 3M_Z''(M_Z')^2 + 6(M_Z')^2M_Z'']M_Z^3 - [M_Z'''M_Z^2 - 3M_Z''M_Z'M_Z + 2(M_Z')^3]3M_Z^2M_Z'\end{array}}{M_Z^6}\right\}(0) \\
&= \left[\frac{M_Z^{(4)}M_Z^3 - 4M_Z'''M_Z'M_Z^2 + 12M_Z''(M_Z')^2M_Z - 3(M_Z'')^2M_Z^2 - 6(M_Z')^4}{M_Z^4}\right](0) \\
&= \mathbf{Kurt}(X)
\end{aligned}$$

이다. □

Notes

1 여기서 $\lim_{x_i\to-\infty} F_X(x) = 0$은 임의의 $\cdots, x_{i-1}, x_{i+1}, \cdots \in \mathbb{R}$와 임의의 $\varepsilon > 0$에 대해 적당한 $M \in \mathbb{R}$이 존재하여 $x_i < M$인 임의의 $x_i \in \mathbb{R}$에 대해 $F_X(x) < \varepsilon$이라는 의미이다.

2 여기서 $\lim_{x_1,\cdots,x_n\to\infty} F_X(x) = 1$은 임의의 $\varepsilon > 0$에 대해 적당한 $M \in \mathbb{R}$이 존재하여 $x \geq M\mathbf{1}$인 임의의 $x \in \mathbb{R}^n$에 대해 $|F_X(x) - 1| < \varepsilon$이라는 의미이다.

3 만약 $x \in \mathbb{R}^n$에 대해 F가 x에서 symmetric differentiable하지 않으면 $F^{(s)}(x)$의 값을 0과 같은 dummy value로 정하는 관례를 전제하였다. 따름정리 ? ? 로부터 이러한 점의 집합이 영집합을 이루고, PDF는 거의 어디서나 같은 함수를 하나로 볼 때 유일하므로 정리의 성립여부는 이때의 dummy value의 선택과 무관하다.

4 만약 $t \in \mathbb{R}^n$에 대해 이가 0을 성분으로 가지면 $\prod_{i=1}^n (e^{-\mathbf{i}t_ix_i} - e^{-\mathbf{i}t_iy_i})/\mathbf{i}t_i$의 값을 0과 같은 dummy value로 정하는 관례를 전제하였다. 이러한 t의 집합이 영집합을 이루므로 적분의 결과는 이때의 dummy value의 선택과 무관하다.

References

[1] Sheldon Ross, *A First Course in Probability*, Prentice Hall, ????.
[2] Patrick Billingsley, *Probability and Measure 3rd Edition*, Wiley, 1995.
[3] 김우철,『수리통계학』, 민영사, 2012.
[4] Mathematics Stack Exchange, Available: https://math.stackexchange.com.
[5] Wolfram MathWorld, Available: http://mathworld.wolfram.com.
[6] Wikipedia, Available: https://en.wikipedia.org.
[7] 나무위키, Available: https://namu.wiki.