A Heuristic Proof of the Radon-Nikodym Theorem

Sung Je Cho*

I. Introduction

The Radon-Nikodym Theorem is among three most important theorems in the modern integration theory. The other two are the Lebesgue's Dominated Convergence Theorem and the Fubini Theorem on the product space. The Lebesgue's Dominated Convergence Theorem deals with the convergence of integrals of point-wise convergent sequences of integrable functions dominated by an integrable function. It covers many important cases of sequences of convergent functions. The Fubini Theorem allows us to reduce the double integral of two dimensions to iterated integrals of one dimension. Also the Fubini Theorem plays a vital role in harmonic analysis. The Radon-Nikodym Theorem concerns about a representation of one measure with respect to another. A little more precisely, if a measure is absolutely continuous with respect to another, then there exists an integrable function such that the measure can be represented as integral of the function with respect to the other measure.

There are several proofs of the Radon-Nikodym Theorem. Proof provided by von Neumann, one of the most prominent mathematician of 20th century, uses the representation theorem of a linear functional on Hilbert space. So to understand his proof, one has to have a solid knowledge on Hilbert space theory. von Neumann's method can be found in [4], for example. Another proof considers the totality of certain functions whose integrals over any measurable sets are less than or equal to the measures of the same sets. Then proof employs some highly technical arguments. See [2] or [3] for

* Seoul National University, Department of Mathematics Education

example. Our proof resembles the proof given in the book of Bartle[1]. Both proofs use the space decomposition into positive and negative sets with respect to the given signed measure. But our proof heavily uses finer space decomposition which resembles the space decomposition when one wants to construct a monotonically increasing sequence of simple functions which converges to the given non-negative measurable function. Of course, the existence of such simple functions are a key to the whole modern integration theory. With the finer space decomposition in hands we only use monotone convergent theorem to find the required function. Thus our proof is more elementary and it gives the readers clearer view on the function under investigation. In this regards, our proof is heuristic. Our approach might be in the same line with the original proof or at least the original idea of proof, but we can not trace the history of proofs of this important theorem. Modern mathematics become more and more abstract. Thus proofs get more and more elegant. In doing so many young mathematicians are likely not to grasp important insights on discoveries or inventions of theorems and their proofs. We hope that this paper would serve a prototype in supplementing such deficiencies resulted from modern abstractions.

II. Preliminaries

In this section we gather some basic terminologies of measure theory and theorems of space decomposition into positive and negative sets. This decomposition is essential to our proof. These facts are well-known and can be found in any modern real analysis book, for example [1]. For completeness we provide adequate explanations.

A collection $\mathfrak{M}$ of subsets of a set X is called a σ-algebra if it satisfies the following:

(1) $\emptyset \in \mathfrak{M}$ and $X \in \mathfrak{M}$.

(2) If $A \in \mathfrak{M}$, then the complement A^c of A is in $\mathfrak{M}$.

(3) If $A_n \in \mathfrak{M}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{M}$.

An element of $\mathfrak{M}$ is called a measurable subset and the pair $(X, \mathfrak{M})$ is usually called a measurable space.

A measure μ is a non-negative extended real-valued function defined on $\mathfrak{M}$ with the following properties:

(1) $\mu(\emptyset)=0$

(2) For any disjoint countable measurable subsets A_n, we have

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$$

The triple $(X, \mathfrak{M}, \mu)$ is called a measure space. It is customary to say that μ is a measure on X. Thus a measure μ can have $\mu(X)=\infty$. A measure μ is said to be finite if $\mu(X) < \infty$. If μ is a measure, then $a\mu$ is also a measure for any positive real number a.

A function λ from $\mathfrak{M}$ into the set of real numbers is called a signed measure if it satisfies the above two conditions about a measure. Thus a finite measure is a signed measure but not the other way around. The difference of any two finite measure is a signed measure. Also it can be proved that any signed measure can be written as a difference of two finite measures. But we do not need this fact.

Let λ be a signed measure on X. A measurable subset P is said to be positive if for any measurable subset A of P we have $\lambda(A) \geq 0$. A measurable subset N is said to be negative if for any measurable subset B of N we have $\lambda(B) \leq 0$. There is a measurable subset which is neither positive nor negative. The following are well-known but perhaps the arrangement is new. For the completeness we provide proofs.

Lemma 1. Let λ be a signed measure on X. Suppose that a measurable subset A does not contain any positive subset of positive signed measure.

Then A is a negative set.

Proof. Suppose that A is not a negative set. Then there is a measurable subset E of A with $\lambda(E)>0$. Since E is not a positive set, there are a measurable subset A_1 of E and a smallest natural number n_1 with $\lambda(A_1)< -\frac{1}{n_1}$. Then $\lambda(E\cap A_1^c)=\lambda(E)-\lambda(A_1)>\lambda(E)+\frac{1}{n_1}>\lambda(E)>0$. Hence the set $E\cap A_1^c$ can not be a positive set. If it is, then $E\cap A_1^c$ is a positive set with positive signed measure. This contradicts the hypothesis. By the same token, there are a measurable subset A_2 of $E\cap A_1^c$ and a smallest natural number n_2 with $\lambda(A_2)< -\frac{1}{n_2}$. Using the mathematical induction, we can find measurable subsets A_k of $E\cap A_1^c\cap\cdots\cap A_{k-1}^c$ and a smallest natural number n_k with $\lambda(A_k)< -\frac{1}{n_k}$. Then all of these sets A_k are disjoint. Let $B=\bigcup_{k=1}^{\infty}A_k$, which is a measurable subset of E Then

$$-\infty<\lambda(B)=\sum_{k=1}^{\infty}\lambda(A_k)<\sum_{k=1}^{\infty}-\frac{1}{n_k}<\infty.$$

Thus as $k\to\infty$, $\frac{1}{n_k}\to 0$. Then $\lambda(E\cap B^c)=\lambda(E)-\lambda(B)>\lambda(E)>0$. Since $E\cap B^c$ can not be a positive set, by the same reason as above, there are a measurable subset C of $E\cap B^c$ and a smallest natural number n_j with $\lambda(C)< -\frac{1}{n_j}$, which contradicts to the choices of A_k and natural number n_k. Thus A must be a negative set. This completes the proof.

Theorem (Hahn Decomposition). Let λ be a signed measure on X Then there are a positive set P and a negative set N such that $P\cap N=\emptyset$ and $X=P\cup N$.

Proof. Let $I=\{\lambda(P)|\ \text{positive set } P\}$. Then I is not an empty set, since the empty set is a positive set. If I is not a bounded set, then there are positive sets A_n such that $\lambda(A_n) \to \infty$. Since the union of positive sets is again positive, $A = \bigcup_{n=1}^{\infty} A_n$ is a positive set and $\lambda(A) = \infty$, which contradicts to the fact that λ takes only real numbers. Let α be the supremum of the bounded set I. Then there are positive sets P_n such that $\lambda(P_n) \to \alpha$. Since the union of positive sets is again positive, $B_n = \bigcup_{k=1}^{n} P_k$ is a positive set and $\lambda(B_n) \geq \lambda(P_n)$. Let $P = \bigcup_{n=1}^{\infty} B_n$. Then by the property of measure we have $\lambda(P) = \lim_{n \to \infty} \lambda(B_n) = \alpha$. Then P is a positive set and its complement P^c contains no positive set with a positive signed measure. If P^c contains a positive set C with $\lambda(C) > 0$. Then the set $P \cup C$ is a positive set and $\lambda(P \cup C) > \lambda(P) = \alpha$. Thus α can not be the supremum of I. Hence $N = P^c$ is a negative set by Lemma 1. This completes the proof.

Remark. Decomposition into positive and negative sets are not unique. A pair of positive and negative measurable sets satisfying the conclusion of the above theorem is called a Hahn decomposition of X with respect to the signed measure λ.

III. Heuristic Proof

Let μ be measure on X. One way to get a measure is as follows: Let f be a non-negative measurable function. We define ν by the integral

$$\nu(E) = \int_E f \, d\mu$$

for every measurable subset E. In these two measures μ and ν, we see that if $\mu(E) = 0$ then we have $\nu(E) = 0$ for every measurable subset E. If the above condition holds for two measures ν and μ on X, then we say that the measure ν is absolutely continuous with respect to μ. Thus the measure generated by the integral of a non-negative measurable function with respect to the given measure gives us a pair of absolutely continuous. The Radon-Nikodym Theorem says that the converse holds as well. The Radon-Nikodym Theorem is true for two σ-finite measures(roughly speaking one which can be expressed as a countable sum of finite measures). The essence of the theorem lies in the case of finite measures. In passing we would like to mention that the function f under investigation is uniquely determined μ-almost everywhere. But we do not prove this fact. Now we are ready to present a heuristic proof of the Radon-Nikodym Theorem for finite measure cases. The following technique is very useful in the proof.

Lemma 2. Let ν and μ be measures on X. Let c_1, c_2 $(0 < c_1 < c_2)$ be real numbers and $c_3 = \frac{c_1 + c_2}{2}$. Suppose that E is the intersection of a positive set with respect to the signed measure $\nu - c_1\mu$ and a negative set with respect to the signed measure $\nu - c_2\mu$. Then there are two measurable subsets E_1 and E_2 such that

(i) $E_1 \cap E_2 = \emptyset$, $E_1 \cup E_2 = E$

(ii) For any subset $F \subset E_1$, we have $c_1\mu(F) \leq \nu(F) \leq c_3\mu(F)$

(iii) For any subset $G \subset E_2$, we have $c_3\mu(G) \leq \nu(G) \leq c_2\mu(G)$.

Proof. Let N be a negative set with respect to $\nu - c_3\mu$. Let $E_1 = E \cap N$ and $E_2 = E \cap N^c$. Then it is easy to see that two measurable subsets E_1 and E_2 fulfill the stated requirements.

Theorem(The Radon-Nikodym Theorem). Let ν and μ be finite measures on X. Suppose that ν is absolutely continuous with respect to μ. Then there is a non-negative measurable function f such that for every measurable subset E

$$\nu(E) = \int_E f\, d\mu.$$

Proof. First, we decompose X into pair-wise disjoint measurable subset with respect to signed measure $\nu - k\mu$ for $k = 1, 2, 3, \cdots$. Let

$X_{1,1}$ be a negative set of a Hahn decomposition of X with respect to the signed measure $\nu-\mu$. Then for any measurable subset $E\subset X_{1,1}$ we have $0\cdot\mu(E)\le\nu(E)\le 1\cdot\mu(E)$. Let $N_{1,2}$ be a negative set of a Hahn decomposition with respect to the signed measure $\nu-2\mu$. Let $X_{1,2}=N_{1,2}\cap X_{1,1}^c$, the intersection of a positive set with respect to $\nu-\mu$ and negative set with respect to $\nu-2\mu$. Then for any measurable subset $E\subset X_{1,2}$ we have $\mu(E)\le\nu(E)\le 2\mu(E)$. Suppose we have constructed pair-wise disjoint measurable subsets $X_{1,1},\cdots,X_{1,n}$ with

$$(k-1)\mu(E)\le\nu(E)\le k\mu(E)$$

for any measurable subset $E\subset X_{1,k}$ and $k=1,2,\cdots,n$. Let $N_{1,n+1}$ be a negative set of a Hahn decomposition with respect to the signed measure $\nu-(n+1)\mu$. Let $X_{1,n+1}=N_{1,n+1}\cap(X_{1,1}\cup X_{1,2}\cup\cdots\cup X_{1,n})^c$. Then for any measurable subset $E\subset X_{1,n+1}$ we have $n\mu(E)\le\nu(E)\le(n+1)\mu(E)$. Thus by using the mathematical induction we can construct pair-wise disjoint measurable subsets $X_{1,n}$ such that

$$(n-1)\mu(E)\le\nu(E)\le n\mu(E)$$

for any measurable subset $E\subset X_{1,n}$. Let $B=(\bigcup_{n=1}^{\infty}X_{1,n})^c$. Then we have $n\mu(B)\le\nu(B)$ for all natural number n Since $\nu(B)<\infty$, $\mu(B)=0$. Then since ν is absolutely continuous with respect to μ, we have $\nu(B)=0$.

Now we are ready to define a function f_1. Note that $X=B\cup X_{1,1}\cup\cdots\cup X_{1,n}\cup\cdots$. We define the first measurable unction f_1 as follows:

$$f_1(x)=\begin{cases}0, & x\in B\\ n-1, & x\in X_{1,n}\end{cases}.$$

For each $X_{1,n}$, by applying Lemma 2 to $n-1$, n $\frac{2n-1}{2}$, and the signed measure $\nu-\frac{2n-1}{2}\mu$, we can find two disjoint measurable subsets $X_{2,2n-1}$ and $X_{2,2n}$ such that $X_{1,n}=X_{2,2n-1}\cup X_{2,2n}$, $\frac{2n-2}{2}\mu(E)\le\nu(E)\le\frac{2n-1}{2}\mu(E)$

for any $E \subset X_{2,2n-1}$ and $\frac{2n-1}{2}\mu(E) \leq \nu(E) \leq \frac{2n}{2}\mu(E)$ for any $E \subset X_{2,2n}$. Then we define the second measurable function f_2 as follows:

$$f_2(x)=\begin{cases}0, & x\in B\\ \frac{n-1}{2}, & x\in X_{2,n}\end{cases}$$

By applying Lemma 2 to $X_{k,n}$ and the signed measure $\nu-\frac{2n-1}{2^{k-1}}\mu$, we can decompose the set $X_{k,n}$ into two disjoint measurable subsets $X_{k+1,2n-1}$ and $X_{k+1,2n}$ such that $X_{k,n}=X_{k+1,2n-1}\cup X_{k+1,2n}$, $\frac{2n-2}{2^k}\mu(E)\leq\nu(E)\leq\frac{2n-1}{2^k}\mu(E)$ for any $E\subset X_{k+1,2n-1}$, and $\frac{2n-1}{2^k}\mu(E)\leq\nu(E)\leq\frac{2n}{2^k}\mu(E)$ for any $E\subset X_{k+1,2n}$. Then we define measurable function f_{n+1} as follows:

$$f_{n+1}(x)=\begin{cases}0, & x\in B\\ \frac{k-1}{2^n}, & x\in X_{n+1,k}\end{cases}.$$

By the mathematical induction as described above, we can construct a sequence of non-negative measurable functions f_n. By the construction it is obvious to see that the functions f_n are increasing. Each f_n is a countable sum of multiples of characteristic functions. Since for each fixed n all $X_{n,k}$ are disjoint, we have

$$\int_E f_n\,d\mu\leq\nu(E)\leq\int_E f_n\,d\mu+\frac{1}{2^{n-1}}\mu(X)$$

Let $f=\lim_{n\to\infty}f_n$. Since $\mu(X)<\infty$, by applying the Monotone Convergence Theorem, we have for any measurable subset E

$$\int_E f\,d\mu\leq\nu(E)\leq\int_E f\,d\mu$$

Thus we have $\nu(E)=\int_E f\,d\mu$ for every measurable subset E This completes the proof.

References

[1] Robert G. Bartle, The elements of Integration and Lebesgue Measure, John Wiley & Sons, Inc. 1995

[2] Gerald B. Folland, Real Analysis, John Wiley & Sons, Inc. 1999

[3] Paul R. Halmos, Measure Theory, D. Van Nostrand Company Inc., Princeton, N.J., 1950

[4] Walter Rudin, Real and Complex Analysis(3ed ed.), McGraw-Hill Book Company, N.Y. 1987

국문초록

Radon-Nikodym정리의 발견술적 증명

조승제
(서울대학교 사범대학 수학교육과)

Radon-Nikodym정리는 Lebesgue의 수렴정리, Fubini정리와 더불어 현대 적분론의 가장 중요한 정리이다. Lebesgue의 수렴정리는 점별수렴하는 함수들의 적분의 극한과 극한 함수의 적분 사이의 관계를 설명하고 Fubini정리는 이차원 적분과 일차원 반복적분과의 관계를 설명하는 정리이다. Radon-Nikodym정리는 한 측도가 다른 측도에 대하여 절대연속이면 한 측도를 다른 측도의 적분형으로 표현할 수 있다는 정리로 여러 가지의 증명 방법이 있다. 본 논문에서는 적분형으로 표현하기 위하여 필요한 함수를 찾기 위하여 먼저 측도 공간을 잘 세분하고 이들 세분으로부터 단조증가하는 함수들을 찾은 다음 이들 함수의 극한함수가 찾고자하는 함수임을 Lebesgue의 단조증가수렴정리를 이용하여 증명하였다. 이 방법은 기존의 여러 증명보다 초보적이고 직관적이며 다른 고등수학의 이론을 사용하지 않았다. 뿐만 아니라 여기에 제시된 증명 방법에 근거하면 자연스럽게 Radon-Nikodym정리를 예측할 수 있을 것이다.