Complementos de teoría de la medida

Asociación Amarun

Versión 0.0.0.1

Diego Chamorro

22 de noviembre de 2015

Índice general

1. Complementos de teoría de la medida **3**
1.1. Medidas con signo y medidas complejas 4
1.1.1. Definiciones y primeras propiedades 4
1.1.2. Teoremas de descomposición de medidas 11
1.1.3. Variación total de medidas 15
1.1.4. Integración con respecto a una medida con signo o medida compleja 21
1.2. Continuidad absoluta 24
1.2.1. Teorema de Radon-Nikodym 27
1.2.2. Esperanza Condicional 34
1.2.3. Singularidad y Teorema de descomposición de Lebesgue 36
1.3. Aplicaciones diferenciables 39
1.3.1. Imagen de medidas y cambio de variables 39
1.3.2. Diferenciación de medidas 51
1.3.3. Funciones de variación acotada 64
1.3.4. Diferenciación de funciones 71
1.4. Medidas en espacios localmente compactos 75
1.4.1. Introducción 75
1.4.2. Teorema de Representación de Riesz 76
1.4.3. Medidas de Stieltjes sobre $\mathbb{R}$ 80
1.5. Ejercicios 83

Índice alfabético **87**

Capítulo 1

Complementos de teoría de la medida

El objetivo principal del Capítulo 2 del Volumen 1 era la construcción general de *medidas* y la noción de medida desarrollada en esas páginas era la de una aplicación que asigna un valor numérico a cierto tipo de conjuntos. Para ello fue necesario precisar qué tipo de conjuntos iban a ser medidos y qué propiedades básicas se debían exigir para que esta aplicación pueda ser considerada como una medida.

Recordemos rápidamente el camino recorrido para llegar a este objetivo: empezábamos con operaciones finitas de conjuntos, considerando álgebras y funciones aditivas de conjuntos; para luego pasar a operaciones numerables usando para ello la noción de σ-álgebras. Indiquemos que el paso de las operaciones finitas a las operaciones numerables no fue algo trivial y que debimos describir en detalle las particularidades que aparecen en esta transición.

En toda esta construcción, se había exigido que la medida de un conjunto cualquiera sea positiva, eventualmente infinita, o nula. En este capítulo vamos a estudiar lo que sucede cuando se permite que la medida de un conjunto sea negativa o compleja. Esto puede parecer extraño a primera vista, en el sentido que es relativamente natural exigir que una medida sea positiva; sin embargo, si cambiamos ligeramente de punto de vista -y ésta será la idea central de este capítulo- y consideramos las "medidas" como puntos de un espacio con una estructura vectorial, es absolutamente necesario poder hacer ciertas operaciones entre ellas, lo cual nos lleva, casi naturalmente, a tomar en cuenta medidas negativas y, por extensión, medidas complejas.

La primera parte de este capítulo estudiará estos aspectos, es decir, que veremos las precauciones que hay que tomar para poder considerar medidas con signo y medidas complejas. Veremos en particular de qué maneras se puede descomponer una medida con signo de suerte que ésta se pueda expresar de forma simple por medio de medidas positivas. Una vez que habremos sentado las bases de estas descomposiciones, veremos gracias a estos resultados qué estructuras se pueden deducir sobre el espacio de las medidas. Para terminar esta primera sección expondremos cómo construir una integral con respecto a estas medidas con signo y sus análogas complejas. Asumiremos en este capítulo que el lector conoce y está familiarizado con el material expuesto en el Capítulo 2 del primer volumen, sin embargo, en algunas ocasiones y para mayor claridad en la exposición, retomaremos algunos puntos de la teoría de la medida en el marco de las medidas con signo y de las medidas complejas.

En la segunda sección, vamos a estudiar la noción de continuidad absoluta de las medidas. Una vez que disponemos de esta herramienta, podremos presentar el teorema de Radon-Nikodym, que es un resultado de gran importancia en las aplicaciones pues permite expresar cierto tipo de medidas como una integral con respecto a otra medida. Desde cierto punto de vista, este resultado puede considerarse como el inicio de la teoría moderna de probabilidades y veremos una muestra de ello presentando el concepto de esperanza condicional. Finalmente, mostraremos cómo la noción de continuidad absoluta, junto con el concepto de singularidad, nos permitirá descomponer las medidas de manera a evidenciar partes constitutivas de estas medidas.

En la tercera sección estudiaremos los cambios de variables en $\mathbb{R}^n$ y sus relaciones con la medida de Lebesgue. Para ello necesitaremos pedir cierta regularidad y diferenciabilidad a los objetos manipulados y aprovecharemos esta situación para presentar temas relacionados con la diferenciación de medidas y de funciones. Pero antes de exponer el proceso de diferenciación de medidas, será necesario presentar algunos lemas de recubrimiento que también serán utilizados posteriormente. Luego estudiaremos las relaciones existentes entre funciones de variación acotada y el proceso de derivación de medidas.

Finalizaremos este capítulo con una cuarta sección estudiando el teorema de representación de Riesz. Este resultado fundamental será inmediatamente utilizado en los capítulos siguientes y para enunciarlo será necesario desarrollar algunos temas relativos a las medidas definidas en espacios localmente compactos.

1.1. Medidas con signo y medidas complejas

Las medidas estudiadas en el Capítulo 2 del Volumen 1 eran aplicaciones μ definidas sobre un espacio medible $(X, \mathscr{A})$ a valores en el espacio $\overline{\mathbb{R}}_+$; es decir que a cada conjunto $A \in \mathscr{A}$ se le asociaba un número $\mu(A)$ positivo, nulo o infinito. En esta primera sección vamos a generalizar el dominio de definición de las medidas permitiendo que las cantidades $\mu(A)$ sean números negativos o inclusive números complejos.

Esta generalización requiere, evidentemente, fijar algunas manipulaciones para evitar incoherencias y esto será tratado en detalle en la Sección 1.1.1. Una vez que habremos descrito las principales propiedades de las medidas con signo, nos ocuparemos del problema siguiente: ¿es posible descomponer una medida con signo como la suma de dos medidas de signo contrario? La idea que se esconde detrás de esta pregunta tiene que ver con la posibilidad de representar a toda medida con signo μ como $\mu = \mu^+ - \mu^-$ en donde μ^+ y μ^- son dos medidas positivas y a este resultado se lo conoce como el teorema de descomposición de medidas.

Una vez que se dispone de la descomposición $\mu = \mu^+ - \mu^-$, y en analogía completa con la noción de valor absoluto de una función, podemos definir una aplicación $|\mu| = \mu^+ + \mu^-$, llamada la variación de μ, y vamos a ver las numerosas propiedades de esta aplicación. Esllevará a la noción de variación total de una medida y a considerar una estructura de espacio de Banach sobre el conjunto de las medidas de variación total finita. Veremos también cómo considerar análogos sobre los números complejos de todos estos conceptos.

Terminaremos esta primera sección explicando la integración con respecto a las medidas con signo y a las medidas complejas. En este punto insisteremos en algunas particularidades que hay que tener en cuenta cuando se integra con respecto a este tipo de medidas.

1.1.1. Definiciones y primeras propiedades

Lo primero que debemos hacer, para considerar medidas con signo, es ampliar el dominio de definición de las aplicaciones que están definidas sobre familias de conjuntos; exigiendo las mismas propiedades de base explicitadas en el primer volumen.

Definición 1.1.1 (Medidas con signo) *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una función definida sobre la σ-álgebra $\mathscr{A}$ a valores en $[-\infty, +\infty]$.*

1) La función $\mu : \mathscr{A} \longrightarrow [-\infty, +\infty]$ es finitamente aditiva *si verifica*

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i)$$

para toda sucesión finita $(A_i)_{1\leq i\leq n}$ de conjuntos dos a dos disjuntos de $\mathscr{A}$.

2) Diremos que $\mu : \mathscr{A} \longrightarrow [-\infty, +\infty]$ es σ-aditiva *o* numerablemente aditiva *si se tiene*

$$\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mu(A_i) \tag{1.1}$$

para toda sucesión infinita $(A_i)_{i\in\mathbb{N}}$ de conjuntos dos a dos disjuntos de $\mathscr{A}$.

3) Si una función $\mu : \mathscr{A} \longrightarrow [-\infty, +\infty]$ es σ-aditiva *y verifica $\mu(\emptyset) = 0$ entonces decimos que μ es una* medida con signo*. Diremos además que una medida con signo es* finita *si no toma los valores $-\infty$ y $+\infty$.*

Notación: Hablaremos siempre de *medidas con signo* en el sentido de la definición anterior y nos referiremos a *medidas positivas* en el sentido de las medidas consideradas en el primer volumen.

Hagamos algunas observaciones importantes de orden general sobre las medidas con signo que se deducen directamente de la definición y que son necesarias para mantener la coherencia de estos objetos. Si $(X, \mathscr{A})$ es un espacio medible y μ es una medida con signo, entonces:

- para todo $A \in \mathscr{A}$ la suma $\mu(A) + \mu(A^c)$ debe estar bien definida: es decir que esta suma no debe ser de la forma $+\infty + (-\infty)$ ó $-\infty + (+\infty)$ pues estas expresiones no tienen sentido. Además se debe tener $\mu(A) + \mu(A^c) = \mu(X)$. Por lo tanto, si para algún conjunto $A \in \mathscr{A}$ se tiene que $\mu(A) = +\infty$, entonces $\mu(X) = +\infty$ y, si para algún conjunto $A \in \mathscr{A}$ se tiene que $\mu(A) = -\infty$, entonces $\mu(X) = -\infty$.
- Esto implica en particular que una medida con signo solo puede tomar uno de los valores $+\infty$ y $-\infty$ y *no estos dos valores simultáneamente.*
- Como consecuencia de estos dos puntos tenemos que si $A \in \mathscr{A}$ es un conjunto de medida finita, entonces $\mu(B) < +\infty$ para todo subconjunto $\mathscr{A}$-medible $B \subset A$.

Demos ahora tres ejemplos sencillos de medidas con signo.

(i) Sea $(X, \mathscr{A})$ un espacio medible y sean $a \neq b$ dos puntos de X. Para todo conjunto $\mathscr{A}$-medible A, definimos una medida con signo μ escribiendo,

$$\mu(A) = -\delta_a(A) + \delta_b(A).$$

En efecto, vemos sin dificultad que $\mu(\emptyset) = 0$. Se tiene entonces que $\mu(A) = 0$ si $a, b \in A$, que $\mu(A) = -1$ si $a \in A$ y si $b \notin A$; y que $\mu(A) = 1$ si $a \notin A$ y si $b \in A$. En este caso la σ-aditividad es inmediata.

(ii) Observemos que la pequeña variante que consiste en definir la medida $\nu(A) = -\delta_a(A)$ toma solo valores negativos o nulos.

(iii) Sea $(X, \mathscr{A}, \mu)$ un espacio medido con μ una medida positiva y sea $f \in L^1(X, \mathscr{A}, \mu, \mathbb{R})$ una función de módulo integrable. Definimos entonces una nueva función de conjuntos ν sobre $\mathscr{A}$ escribiendo

$$\nu(A) = \int_A f(x) d\mu(x). \tag{1.2}$$

Cuando f es una función positiva se obtiene una medida y esto ya ha sido tratado en la Definición 3.3.1 de medida inducida por una función dada en el Capítulo 3 del primer volumen. En el caso cuando la función f puede tomar valores negativos, esta expresión determina una medida con signo sobre $(X, \mathscr{A})$. Para verificarlo es necesario usar la linealidad de la integral para separar las partes positivas y negativas de la función f y aplicar el teorema de convergencia dominada de Lebesgue. Ver el Ejercicio 1.1 para los detalles.

Es muy interesante notar que esta medida ν es la diferencia de dos medidas positivas ν_1 y ν_2 determinadas por $\nu_1(A) = \displaystyle\int_A f^+ d\mu$ y $\nu_2(A) = \displaystyle\int_A f^- d\mu$, en donde hemos notado $f^+ = \text{máx}(f, 0)$ y $f^- = \text{máx}(-f, 0)$: es decir que podemos escribir $\nu = \nu_1 - \nu_2$. Volveremos a este punto posteriormente.

Continuamos nuestra exposición sobre las propiedades de las medidas con signo. En este caso, cuando las medidas pueden tomar valores en $\overline{\mathbb{R}}$, la propiedad de σ-aditividad (1.1) es una condición muy fuerte que tiene consecuencias importantes.

Proposición 1.1.1 (Convergencia normal de medidas con signo) *Sea μ una medida con signo definida sobre un espacio medible $(X, \mathscr{A})$. Si $(A_n)_{n\in\mathbb{N}}$ es una sucesión de conjuntos disjuntos $\mathscr{A}$-medibles tales que $\mu(\bigcup_{n\in\mathbb{N}} A_n) < +\infty$, entonces la serie*

$$\sum_{n\in\mathbb{N}} \mu(A_n) \tag{1.3}$$

es normalmente convergente.

Antes de pasar a la demostración de este resultado, expliquemos brevemente la provenencia de este hecho: puesto que la reunión de los conjuntos disjuntos A_n no cambia si se modifica la numerotación de los índices, es necesario que la suma sea estable por reordenamiento y esto implica la convergencia normal.

Prueba. Supongamos que $(A_n)_{n\in\mathbb{N}}$ es una sucesión de conjuntos disjuntos $\mathscr{A}$-medibles tales que $\mu(\bigcup_{n\in\mathbb{N}} A_n)$ es un número real. Definimos una sucesión de conjuntos $(B_n)_{n\in\mathbb{N}}$ de la siguiente manera

$$B_n = \begin{cases} A_n & \text{si } \mu(A_n) \geq 0 \\ \emptyset & \text{si } \mu(A_n) < 0, \text{ para todo } n \in \mathbb{N}. \end{cases}$$

Notamos también $A = \bigcup_{n\in\mathbb{N}} A_n$ y $B = \bigcup_{n\in\mathbb{N}} B_n$. Vemos entonces que se tiene

$$\mu(A) = \mu(B \cup (A \setminus B)) = \mu(B) + \mu(A \setminus B)$$

de manera que tanto $\mu(B)$ como $\mu(A \setminus B)$ son números reales. Tenemos además que

$$\mu(B) = \sum_{n\in\mathbb{N}} \mu(B_n) \quad \text{ y } \quad \mu(A \setminus B) = \sum_{n\in\mathbb{N}} \mu(A_n \setminus B_n)$$

son dos series convergentes de números positivos y de números negativos respectivamente. Por lo tanto podemos escribir

$$\sum_{n=0}^{m} |\mu(A_n)| = \sum_{n=0}^{m} |\mu(B_n \cup (A_n \setminus B_n))| = \sum_{n=0}^{m} \mu(B_n) - \sum_{n=0}^{m} \mu(A_n \setminus B_n) \leq \mu(B) - \mu(A \setminus B)$$

para todo $m \in \mathbb{N}$ de manera que se tiene que la serie $\sum_{n\in\mathbb{N}} \mu(A_n)$ converge absolutamente. ■

La Proposición 1.1.1 es un resultado muy fuerte de estabilidad de las medidas con signo. En particular, esta proposición nos permite construir muchos otros ejemplos de medidas con signo, generalizando los ejemplos presentados en la página anterior.

Proposición 1.1.2 (Operaciones entre medidas) *Si μ y ν son dos medidas con signo definidas sobre un mismo espacio medible $(X, \mathscr{A})$, entonces las medidas*

$$\mu + \nu \quad \text{y} \quad \mu - \nu$$

definidas por las fórmulas $(\mu + \nu)(A) = \mu(A) + \nu(A)$ y $(\mu - \nu)(A) = \mu(A) - \nu(A)$ para todo $A \in \mathscr{A}$, son

1) *dos medidas con signo finitas si ambas medidas μ y ν son medidas finitas.*

2) *dos medidas con signo si las medidas μ y ν son medidas con signo y al menos una de ellas es finita.*

Prueba. Empecemos con el primer punto, en donde todos los valores de las medidas μ y ν son números reales finitos. Dado que se tiene $(\mu \pm \nu)(\emptyset) = \mu(\emptyset) \pm \nu(\emptyset) = 0$ y que, para toda familia finita de conjuntos dos a dos disjuntos $A_k \in \mathscr{A}$, se tiene que

$$(\mu \pm \nu)\left(\bigcup_{k=0}^{n} A_k\right) = \mu\left(\bigcup_{k=0}^{n} A_k\right) \pm \nu\left(\bigcup_{k=0}^{n} A_k\right) = \sum_{k=0}^{n} \mu(A_k) \pm \sum_{k=0}^{n} \nu(A_k) = \sum_{k=0}^{n} (\mu \pm \nu)(A_k),$$

se deduce que las aplicaciones $\mu + \nu$ y $\mu - \nu$ son funciones aditivas de conjuntos y por lo tanto lo único que queda por demostrar es la propiedad de σ-aditividad. Aplicamos ahora la Proposición 1.1.1 que nos proporciona la convergencia normal de las series $\sum_{k\in\mathbb{N}} \mu(A_k)$ y $\sum_{k\in\mathbb{N}} \nu(A_k)$, y, puesto que las medidas μ y ν son medidas con signo finitas, se tiene entonces esta propiedad de aditividad numerable y se obtiene que las aplicaciones $\mu + \nu$ y $\mu - \nu$ son medidas finitas.

El segundo caso, es muy similar y solo hay que tener cuidado con la no finitud de una de las medidas, lo cual no causa mayor problema pues en ese caso todas las operaciones están bien definidas: solo una de las medidas μ o ν toma un valor infinito. ■

Vemos pues que si tenemos dos medidas positivas, podemos a partir de estos datos construir nuevas medidas considerando su suma o su resta y esto sugiere la existencia de una estructura de espacio vectorial. Observemos sin embargo que hay restricciones para realizar estas operaciones pues al menos una de las medidas debe ser finita y esto será analizado posteriormente al considerar los espacios de medidas. Veremos además la recíproca un poco más adelante: es decir que explicaremos cómo descomponer toda medida con signo como la diferencia de dos medidas positivas.

Pasemos ahora, con las dos proposiciones siguientes, a la exposición de otras propiedades de las medidas con signo. El primer resultado no es más que una reescritura de las condiciones exigidas para obtener una medida con signo de manera que es enunciado sin demostración; mientras que el segundo resultado nos proporciona estabilidad al pasar al límite cuando se consideran sucesiones de conjuntos.

Proposición 1.1.3 *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida con signo definida sobre este espacio.*

1) *Si $A, B \in \mathscr{A}$ y si $A \cap B = \emptyset$, entonces $\mu(A \cup B) = \mu(A) + \mu(B)$.*

2) *Si $A, B \in \mathscr{A}$, si $A \subset B$ y si $|\mu(B)| < +\infty$, entonces $|\mu(A)| < +\infty$.*

3) *Si $A, B \in \mathscr{A}$, si $A \subset B$ y si $|\mu(A)| < +\infty$, entonces $\mu(B \setminus A) = \mu(B) - \mu(A)$.*

La proposición que sigue es la generalización de un resultado estudiado anteriormente en el primer volumen.

Proposición 1.1.4 (Continuidad de las medidas con signo) *Sea $(X, \mathscr{A})$ un espacio medible y μ una medida con signo definida sobre este espacio. Si $(A_n)_{n\in\mathbb{N}}$ es una sucesión creciente de conjuntos de $\mathscr{A}$ entonces*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mu(A_n). \tag{1.4}$$

Si $(A_n)_{n\in\mathbb{N}}$ es una sucesión decreciente de conjuntos de $\mathscr{A}$ tal que $\mu(A_i)$ es finita para algún índice i entonces

$$\mu\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mu(A_n). \tag{1.5}$$

Prueba. Para verificar estos puntos es suficiente seguir paso a paso la demostración de los resultados presentados en el Capítulo 2 (ver el Teorema 2.2.3) de manera que los detalles quedan a cargo del lector. ■

El siguiente resultado es de utilidad para verificar bajo qué condiciones una función aditiva de conjuntos, que puede tomar valores positivos o negativos, es en realidad una medida con signo.

Proposición 1.1.5 *Sea $(X, \mathscr{A})$ un espacio medible y $\mu : \mathscr{A} \longrightarrow \overline{\mathbb{R}}$ una aplicación que es finitamente aditiva y que verifica $\mu(\emptyset) = 0$.*

1) *si se tiene la identidad $\mu\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \lim\limits_{n\to+\infty} \mu(A_n)$ para toda sucesión creciente $(A_n)_{n\in\mathbb{N}}$ de conjuntos $\mathscr{A}$-medibles o,*

2) *si se tiene la identidad $\lim\limits_{n\to+\infty} \mu(A_n) = 0$ para toda sucesión decreciente $(A_n)_{n\in\mathbb{N}}$ de conjuntos $\mathscr{A}$-medibles tales que $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$,*

entonces la aplicación μ es una medida con signo sobre el espacio $(X, \mathscr{A})$.

Prueba. Lo único que hay que verificar es la propiedad de σ-aditividad de la función μ. Sea pues $(B_n)_{n\in\mathbb{N}}$ una sucesión disjunta de conjuntos $\mathscr{A}$-medibles. Para todo $k \in \mathbb{N}$ definimos el conjunto $A_k = \bigcup_{j=0}^{k} B_j$, de donde se deduce, por la aditividad finita de la aplicación μ, que $\mu(A_k) = \sum_{j=0}^{k} \mu(B_k)$. Por otro lado, si suponemos el punto 1) se tiene que $\mu(\bigcup_{k=0}^{+\infty} A_k) = \lim\limits_{k\to+\infty} \mu(A_k)$; pero como se tiene $\bigcup_{j\in\mathbb{N}} B_j = \bigcup_{k\in\mathbb{N}} A_k$, se deduce que $\mu(\bigcup_{j=0}^{+\infty} B_j) = \sum_{j=0}^{+\infty} \mu(B_j)$.

Supongamos ahora que se tiene el segundo punto. Definimos entonces, para todo $k \in \mathbb{N}$, la sucesión $A_k = \bigcup_{j=k}^{+\infty} B_j$. Entonces la aditividad finita de la aplicación μ implica que

$$\mu(\bigcup_{j=0}^{+\infty} B_j) = \sum_{j=0}^{k} \mu(B_j) + \mu(A_{k+1}),$$

mientras que la condición 2) implica que $\lim_{k\to+\infty} \mu(A_{k+1}) = 0$; de donde se deduce $\mu(\bigcup_{j=0}^{+\infty} B_j) = \sum_{j=0}^{+\infty} \mu(B_j)$. En ambos casos hemos demostrado la σ-aditividad de la aplicación μ, lo que hace de ella una medida con signo. ■

Como se puede ver con estos resultados anteriores, las medidas con signo comparten muchas propiedades de las medidas positivas estudiadas en el primer volumen. El lector podría pensar entonces que las medidas con signo no son más que medidas positivas a las cuales se ha permitido tomar valores negativos. Este no es el caso y es necesario ahora exhibir algunas diferencias que existen entre las medidas con signo y las medidas estudiadas en el primer volumen.

Vemos para empezar que las medidas con signo no verifican la propiedad de subaditividad: si μ es una medida con signo y si $(A_n)_{n\in\mathbb{N}}$ es una familia de conjuntos medibles, no se tiene necesariamente la desigualdad

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n\in\mathbb{N}} \mu(A_n).$$

En efecto, consideremos la medida $\mu = \delta_{-4} - \delta_{-2} + 2\delta_0 - \delta_2 + \delta_4$ y sean los conjuntos $A = [-5, 5]$, $A_0 = [-5, -1]$, $A_1 = [-3, 3]$, $A_2 = [1, 5]$. Se tiene que $A = A_0 \cup A_1 \cup A_2$ pero $\mu(A) = 2$ y $\mu(A_0) = \mu(A_1) = \mu(A_2) = 0$ de manera que no se tiene la desigualdad $\mu(A_0 \cup A_1 \cup A_2) \leq \mu(A_0) + \mu(A_1) + \mu(A_2)$.

Veamos otra diferencia interesante. Si μ es una medida con signo definida sobre un espacio medible $(X, \mathscr{A})$; para decir que un conjunto D es *μ-despreciable* no basta pedir que $D \subset A$ con $\mu(A) = 0$ pues el ejemplo anterior muestra que existen conjuntos medibles de medida nula con subconjuntos de medida positiva. En efecto, se tiene en este caso que $\mu(A_2) = 0$ pero si $B = [3, 5]$ se tiene $B \subset A_2$ y que $\mu(B) = 1$. Es por esta razón que exigiremos una condición más fuerte:

Definición 1.1.2 (Conjunto μ-despreciable) *Sea $(X, \mathscr{A})$ un espacio medible y μ una medida con signo definida sobre él. Diremos que un conjunto $\mathscr{A}$-medible A es* μ-despreciable *si para todo $B \in \mathscr{A}$ subconjunto de A se tiene $\mu(B) = 0$.*

Con estos resultados hemos dado una primera revisión de las propiedades de las medidas con signo y hemos visto algunas diferencias notables con las medidas usuales estudiadas anteriormente.

Para continuar, es necesario considerar medidas ligeramente distintas y pasamos ahora a la presentación de medidas que pueden tomar valores complejos. Vamos a ver que a partir de medidas con signo finitas es posible construir medidas complejas de la misma manera que a partir de dos números reales es posible construir números complejos.

Definición 1.1.3 (Medidas complejas) *Sea $(X, \mathscr{A})$ un espacio medible. Una* medida compleja *μ es una función definida sobre $\mathscr{A}$ a valores en $\mathbb{C}$ que verifica los dos puntos siguientes*

1) Se tiene $\mu(\emptyset) = 0$.

2) La función $\mu : \mathscr{A} \longrightarrow \mathbb{C}$ es σ-aditiva en el sentido siguiente: se tiene

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n) \tag{1.6}$$

para toda sucesión numerable $(A_n)_{n\in\mathbb{N}}$ de conjuntos dos a dos disjuntos de $\mathscr{A}$.

Observación 1.1 Es muy importante notar que, por definición, una medida compleja toma sus valores en $\mathbb{C}$ y por lo tanto no admite valores infinitos en oposición con las medidas con signo definidas anteriormente. En este sentido, las medidas complejas no constituyen una generalización de las medidas con signo.

Demos, para fijar las ideas, dos ejemplos simples y concretos de tales medidas.

(i) Consideremos pues $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y definamos sobre $\mathcal{B}or(\mathbb{R})$ la función $\mu = -\delta_1 + i\delta_2$. Se tiene entonces para el conjunto $A = [0, 3]$ que $\mu(A) = -1 + i$, de manera que se ha asignado a este conjunto A un valor complejo. El lector verificará sin mayor dificultad que esta aplicación es una medida compleja.

(ii) Consideremos ahora $f : X \longrightarrow \mathbb{C}$ una función que pertenece al espacio $L^1(X, \mathscr{A}, \mu, \mathbb{C})$ con μ una medida positiva. Entonces, de la misma manera que en el ejemplo (ii) de la página 6, se tiene que la medida ν definida por medio de la expresión

$$\nu(A) = \int_A f(x)d\mu(x), \text{ para todo } A \in \mathscr{A}$$

define una medida compleja sobre el espacio medible $(X, \mathscr{A})$.

Observemos que es relativamente sencillo construir tales medidas: en efecto, si μ y ν son dos medidas finitas (reales) definidas sobre un mismo espacio medible $(X, \mathscr{A})$ entonces, para todo $A \in \mathscr{A}$, la fórmula $\pi(A) = \mu(A) + i\nu(A)$ determina una medida compeja π sobre $\mathscr{A}$ y tenemos que la parte real de esta medida está dada por la medida μ, notaremos entonces $\mathfrak{Re}(\pi) = \mu$, mientras que la parte imaginaria está dada por la medida ν, lo cual será notado por $\mathfrak{Im}(\pi) = \nu$.

Es importante recalcar que la convergencia de la serie en la parte derecha de la fórmula (1.6) debe entenderse como el límite de sus sumas parciales. En ese sentido se tiene el análogo para las medidas complejas de la Proposición 1.1.1:

Proposición 1.1.6 *Si μ es una medida compleja definida sobre un espacio medible $(X, \mathscr{A})$ y si $(A_n)_{n\in\mathbb{N}}$ es una sucesión de conjuntos $\mathscr{A}$-medibles dos a dos disjuntos tales que $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right)$ es un número complejo, entonces la serie (1.6) converge normalmente.*

Prueba. Sea μ una medida compleja y consideremos $\mathfrak{Re}(\mu)$ y $\mathfrak{Im}(\mu)$ las partes reales e imaginarias de μ. Entonces tenemos que $\mathfrak{Re}(\mu)$ y $\mathfrak{Im}(\mu)$ son medidas con signo finitas de manera que las cantidades $\mathfrak{Re}(\mu)\left(\bigcup_{n\in\mathbb{N}} A_n\right)$ y $\mathfrak{Im}(\mu)\left(\bigcup_{n\in\mathbb{N}} A_n\right)$ son números reales, de modo que se tiene, por la Proposición 1.1.1, las estimaciones

$$\sum_{n\in\mathbb{N}} |\mu(A_n)| \leq \sum_{n\in\mathbb{N}} |\mathfrak{Re}(\mu)(A_n)| + \sum_{n\in\mathbb{N}} |\mathfrak{Im}(\mu)(A_n)| < +\infty,$$

y de esta manera se obtiene la convergencia normal de esta serie. ■

Este resultado permite definir, de igual forma que en el caso real, varias medidas complejas: si μ y ν son medidas complejas, entonces $\mu \pm \nu$ son medidas complejas.

Proposición 1.1.7 (Operaciones entre medidas complejas) *Sean μ y ν dos medidas complejas definidas sobre el espacio medible $(X, \mathscr{A})$. Para todo $A \in \mathscr{A}$ definimos la suma $\mu + \nu$ por*

$$\mu(A) + \nu(A) = \mathfrak{Re}(\mu(A)) + \mathfrak{Re}(\nu(A)) + i\big(\mathfrak{Im}(\mu(A)) + \mathfrak{Im}(\nu(A))\big).$$

La resta se define de manera similar

$$\mu(A) - \nu(A) = \mathfrak{Re}(\mu(A)) - \mathfrak{Re}(\nu(A)) + i\big(\mathfrak{Im}(\mu(A)) - \mathfrak{Im}(\nu(A))\big).$$

Para la prueba de esta proposición basta adaptar los argumentos utilizados en la demostración de la Proposición 1.1.2 separando las partes reales e imaginarias y usar la Proposición 1.1.6.

Otra particularidad de trabajar con medidas complejas tiene que ver con el hecho de que solo se dispone de una estructura de orden parcial sobre el conjunto de los números complejos y esto tiene algunas consecuencias que hay que tomar en cuenta: por ejemplo no tiene mucho sentido estudiar la subaditividad (finita o numerable) de medidas complejas.

Como vemos, hay algunas similitudes entre las medidas positivas (es decir las medidas consideradas en el primer volumen), las medidas con signo y las medidas complejas. Sin embargo hay que tener algunas precauciones y es necesario aclarar la situación:

- Una medida *positiva* (una medida usual) es una medida con signo sin valores negativos.
- Una medida con signo *finita* es una medida con signo pero también es una medida compleja de parte imaginaria nula.
- Una medida *positiva* que toma valores en $\overline{\mathbb{R}}_+$ no es una medida compleja.
- Una medida con *signo* que toma valores en $\overline{\mathbb{R}}$ no es una medida compleja.

Estos dos últimos puntos son importantes pues, una vez más, no hay que considerar las medidas complejas como una generalización de las medidas positivas.

1.1.2. Teoremas de descomposición de medidas

Hemos visto con la Proposición 1.1.2 y su contraparte para las medidas complejas, la Proposición 1.1.6, que, si disponemos de dos medidas positivas μ y ν, definidas sobre un mismo espacio medible $(X, \mathscr{A})$, entonces es posible construir una nueva medida escribiendo $\pi = \mu - \nu$.

El objetivo de este párrafo es mostrar que toda medida con signo μ se puede descomponer como la diferencia $\mu = \mu^+ - \mu^-$ donde μ^+ y μ^- son medidas positivas. Veremos también cómo aplicar esta descomposición a las medidas complejas. La importancia de disponer de este tipo de descomposiciones es inmediata: podremos usar muchos de los resultados demostrados en el primer volumen y aplicarlos al caso de medidas más generales (con ciertas precauciones, evidentemente).

Una primera etapa para poder descomponer medidas tiene que ver con la clasificación de los conjuntos en el sentido de la siguiente definición.

Definición 1.1.4 (Conjuntos positivos y negativos) *Sea μ una medida con signo definida sobre un espacio medible $(X, \mathscr{A})$.*

1) *Un subconjunto A de X es un conjunto* positivo *si $A \in \mathscr{A}$ y si todo subconjunto $\mathscr{A}$-medible E de A verifica $\mu(E) \geq 0$.*

2) *Recíprocamente, un subconjunto A de X es un conjunto* negativo *si $A \in \mathscr{A}$ y si todo subconjunto $\mathscr{A}$-medible E de A verifica $\mu(E) \leq 0$.*

Demos unos ejemplos de tales conjuntos. Consideremos la recta real $\mathbb{R}$ a la cual dotamos de la medida $\mu = -\delta_{-1} + \delta_1$. El lector observará que el conjunto $\{1\}$ es positivo y que el conjunto $\{-1\}$ es negativo, mientras que el conjunto $A = [-2, 2]$ no es ni un conjunto positivo, ni un conjunto negativo.

Para continuar, necesitaremos el lema a continuación.

Lema 1.1.1 *Sea μ una medida con signo definida sobre un espacio medible $(X, \mathscr{A})$ y sea $A \in \mathscr{A}$ un subconjunto de X que verifica $-\infty < \mu(A) < 0$. Entonces existe un conjunto negativo B contenido en A que verifica $\mu(B) \leq \mu(A)$.*

Prueba. Empecemos fijando $\delta_1 = \sup\{\mu(E) : E \in \mathscr{A},\, E \subset A\}$, vemos que esta cantidad es positiva o nula puesto que se tiene $\delta_1 \geq \mu(0) = 0$. Escojamos ahora un subconjunto $\mathscr{A}$-medible A_1 de A que verifica

$$\mu(A_1) \geq \text{mín}(\frac{1}{2}\delta_1, 1),$$

evidentemente esta cantidad también es positiva o nula. Procedemos ahora por inducción contruyendo sucesiones $(\delta_n)_{n\in\mathbb{N}}$ y $(A_n)_{n\in\mathbb{N}}$ de esta forma:

$$\delta_n = \sup\left\{\mu(E) : E \in \mathscr{A},\, E \subset (A \setminus \bigcup_{i=1}^{n-1} A_i)\right\}$$

y a partir de este valor, fijamos un subconjunto $\mathscr{A}$-medible A_n del conjunto $A \setminus \bigcup_{i=1}^{n-1} A_i$ tal que $\mu(A_n) \geq \text{mín}(\frac{1}{2}\delta_n, 1)$. Ahora definamos A_∞ y B de esta manera: $A_\infty = \bigcup_{n=1}^{+\infty} A_n$ y $B = A \setminus A_\infty$.

Como los conjuntos A_n son disjuntos y verifican $\mu(A_n) \geq 0$, se tiene que $\mu(A_\infty) \geq 0$ y por lo tanto que

$$\mu(A) = \mu(A_\infty) + \mu(B) \geq \mu(B).$$

Queda por verificar la negatividad del conjunto B. Para ello observamos que la finitud de $\mu(A)$ implica la finitud de $\mu(A_\infty)$, y como se tiene $\mu(A_\infty) = \sum_{n\in\mathbb{N}} \mu(A_n)$ entonces necesariamente $\lim\limits_{n\to+\infty} \mu(A_n) = 0$ y por lo tanto se tiene que $\lim\limits_{n\to+\infty} \delta_n = 0$. Dado que por construcción todo subconjunto $\mathscr{A}$-medible E de B verifica $\mu(E) \leq \delta_n$ para todo n se obtiene que $\mu(E) \leq 0$: es decir la negatividad de B. ■

Con este resultado preliminar, podemos enunciar uno de los teoremas más importantes de esta sección que explica cómo se puede descomponer un conjunto por medio de conjuntos positivos y negativos. A este tipo de descomposición se la conoce como la descomposición de Hahn[1] de medidas.

Teorema 1.1.1 (de descomposición de Hahn) *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida con signo sobre $(X, \mathscr{A})$. Existen entonces dos subconjuntos disjuntos P y N de X tales que P es un conjunto positivo de μ, N es un conjunto negativo de μ y se tiene que $X = P \cup N$.*

Diremos entonces que (P, N) es una descomposición de Hahn *de X asociada a la medida μ.*

Demostración. Dado que una medida con signo μ no puede tener entre sus valores $+\infty$ y $-\infty$ simultáneamente, podemos sin pérdida de generalidad suponer que el valor $-\infty$ no pertenece al rango de la medida μ. Para empezar la demostración definimos la cantidad

$$L = \text{ínf}\left\{\mu(A) : A \text{ es un conjunto negativo para } \mu\right\}$$

El conjunto sobre el cual corre el ínfimo no es vacío pues contiene al menos el conjunto vacío que es un conjunto negativo para la medida μ puesto que $\mu(\emptyset) \leq 0$.

Fijemos ahora una sucesión $(A_n)_{n\in\mathbb{N}}$ de conjuntos negativos para μ tales que se tenga $L = \lim\limits_{n\to+\infty} \mu(A_n)$ y consideremos el conjunto $N = \bigcup_{n=1}^{+\infty} A_n$. Tenemos entonces que N es un conjunto

[1] Hans Hahn (1879-1934), matemático austríaco.

negativo para μ: en efecto, todo subconjunto $\mathscr{A}$-medible de N puede expresarse como una unión disjunta de elementos $\mathscr{A}$-medibles que están incluidos en algún conjunto A_n. Observamos entonces que se tiene $L \leq \mu(N) \leq \mu(A_n)$ para todo $n \geq 1$ y por lo tanto se obtiene que $L = \mu(N)$. Además, como la medida μ no alcanza el valor $-\infty$, el valor de $\mu(N)$ debe ser finito.

Para continuar consideramos $P = N^c$ y solo nos queda por verificar que P es un conjunto positivo para μ. Si P está contenido en un conjunto $\mathscr{A}$-medible A tal que $\mu(A) < 0$, entonces por el Lema 1.1.1 tenemos que contiene un conjunto negativo B tal que $\mu(B) < 0$ y por lo tanto $N \cup B$ sería un conjunto negativo tal que

$$\mu(N \cup B) = \mu(N) + \mu(B) < \mu(N) = L.$$

Esto es entonces una contradicción con la definición de la cantidad L y se tiene que P debe ser un conjunto positivo para la medida μ.

Obtenemos de esta manera la descomposición deseada: $X = P \cup N$, con $P \cap N = \emptyset$, en donde P es un conjunto positivo y N es un conjunto negativo para la medida μ. ■

Demos un primer ejemplo de descomposición de Hahn. Sea $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y sea $P \in \mathcal{B}or(\mathbb{R})$ un conjunto tal que su medida de Lebesgue sea finita: $|P| < +\infty$. Consideramos la función $f(x) = \mathbb{1}_P(x) - \mathbb{1}_{P^c}(x)$ para definir una medida con la expresión

$$\nu(A) = \int_A f(x)dx.$$

Vemos entonces, sin mayor problema, que (P, N) con $N = P^c$ es una descomposición de Hahn de $\mathbb{R}$ asociada a la medida ν. En particular $\nu(P)$ es el máximo valor positivo tomado por la medida ν en el sentido que no existe un conjunto $A \subset \mathcal{B}or(\mathbb{R})$ tal que $\nu(A) \geq \nu(P)$, mientras que $\mu(N)$ es su máximo valor negativo.

Es interesante notar que para una misma medida pueden existir más de una descomposición de Hahn, es decir que no existe la unicidad en este tipo de descomposiciones. Por ejemplo si $X = [-1, 1]$, si $\mathscr{A} = \mathcal{B}or([-1, 1])$ y si μ es una medida definida por $\mu(A) = \displaystyle\int_A xdx$ entonces $([0, 1], [-1, 0[)$ y $(]0, 1], [-1, 0])$ son dos descomposiciones de Hahn distintas para la medida μ.

Sin embargo la diferencia entre estas descomposiciones no es más que un conjunto de medida nula para la medida μ. En efecto si (P_1, N_1) y (P_2, N_2) son dos descomposiciones de Hahn para una misma medida μ, entonces $P_1 \cap N_2$ es a la vez un conjunto positivo y negativo para μ y esto implica que cada subconjunto de $P_1 \cap N_2$ es de μ-medida nula. Evidentemente esta propiedad se mantiene para $P_2 \cap N_1$. En este sentido se obtiene en términos prácticos que la descomposición de Hahn es esencialmente única.

El teorema de descomposición de Hahn nos permite representar un conjunto como la unión disjunta de conjuntos positivos y negativos. La etapa siguiente consiste en estudiar las repercusiones de esta descomposición al nivel de la medida de conjuntos generales. Se obtiene entonces el corolario a continuación:

Corolario 1.1.1 *Cada medida con signo definida sobre un espacio medible es la diferencia de dos medidas positivas en donde al menos una de ellas es finita.*

Demostración. Sea μ una medida con signo sobre $(X, \mathscr{A})$. Utilizando el teorema anterior, podemos fijar (P, N) una descomposición de X asociada a la medida μ y definir las siguientes aplicaciones:

$$\mu^+(A) = \mu(A \cap P) \tag{1.7}$$

$$\mu^-(A) = -\mu(A \cap N) \tag{1.8}$$

Por construcción, se tiene que las aplicaciones μ^+ y μ^- son medidas positivas (recordar la Proposición 2.2.7 del primer volumen) tales que $\mu = \mu^+ - \mu^-$. Dado que $+\infty$ y $-\infty$ no pertenecen simultáneamente al rango de valores de μ, al menos uno de los valores $\mu(P)$ y $\mu(N)$ es finito. Esto implica por lo tanto que al menos una de las medidas μ^+ y μ^- es finita. ∎

Podría pensarse que las medidas μ^+ y μ^- definidas por las fórmulas (1.7) y (1.8) dependen explícitamente de la descomposición de Hahn fijada. Veamos otro punto de vista con la definición a continuación.

Definición 1.1.5 (Descomposición de Jordan) *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida con signo definida sobre $\mathscr{A}$. Para todo conjunto medible $A \in \mathscr{A}$ definimos:*

$$\mu^+(A) = \sup\Big\{\mu(B) : B \in \mathscr{A}, B \subset A\Big\}$$

$$\mu^-(A) = \sup\Big\{-\mu(B) : B \in \mathscr{A}, B \subset A\Big\}.$$

Las medidas μ^+ y μ^- son respectivamente las partes positivas *y* negativas *de la medida μ y la representación*

$$\mu = \mu^+ - \mu^-$$

se denomina la descomposición de Jordan[2] *de la medida μ.*

Observamos que en esta definición no entra explícitamente en juego la descomposición de Hahn y, en este sentido, este punto de vista puede parecer ligeramente más general que la caracterización dada con las fórmulas (1.7) y (1.8). Sin embargo es útil explicar la conección entre estas dos definiciones.

Sea pues $(X, \mathscr{A})$ un espacio medible, sea μ una medida con signo definida sobre $\mathscr{A}$ y sea (P, N) una descomposición de Hahn asociada a la medida con signo μ. Si $A \in \mathscr{A}$, entonces para todo subconjunto $\mathscr{A}$-medible B de A se tiene

$$\mu(B) = \mu(B \cap P) + \mu(B \cap N) \leq \mu(B \cap P) \leq \mu(A \cap P)$$

de manera que $\mu(A \cap P) = \sup\{\mu(B) : B \in \mathscr{A}, B \subset A\}$, pero justamente se tiene $\mu(A \cap P) = \mu^+(A)$. Razonando de forma similar se obtiene sin problema la definición anterior.

Es posible generalizar esta descomposición al caso de las medidas complejas.

Definición 1.1.6 (Descomposición de Jordan de medidas complejas) *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida compleja definida sobre $\mathscr{A}$. Como para todo conjunto $A \in \mathscr{A}$ la cantidad $\mu(A)$ es un número complejo, podemos escribir $\mu(A) = \nu_1(A) + i\nu_2(A)$ en donde ν_1 y ν_2 son dos medidas con signo finitas definidas sobre $(X, \mathscr{A})$. Tenemos entonces la* descomposición de Jordan de la medida compleja *μ:*

$$\mu = \nu_1^+ - \nu_1^- + i\nu_2^+ - i\nu_2^-$$

en donde ν_1^+, ν_1^-, ν_2^+ y ν_2^- son medidas positivas finitas definidas sobre $(X, \mathscr{A})$.

[2]Camille Jordan (1838-1922), matemático francés.

Esto muestra que, de una manera muy general, es suficiente considerar las medidas con signo finitas para estudiar las medidas complejas. En este sentido es interesante notar que la parte más importante de comprender, es decir el salto teórico fundamental, reside en el paso de las medidas usuales a las medidas con signo.

1.1.3. Variación asociada a una medida, Variación Total de medidas

Para poder seguir nuestro estudio es necesario presentar un concepto relacionado con el "tamaño" de las medidas. Evidentemente, esta noción de tamaño debe ser definida correctamente para que pueda ser de alguna utilidad. En nuestro caso, vamos a ver que este concepto nos permitirá definir una norma sobre espacios de medidas, es decir que estaremos en capacidad de asignar un número real positivo a cada medida, ya sea ésta una medida con signo o una medida compleja, y veremos además que los espacios normados resultantes son espacios de Banach. A la luz de los resultados expuestos en el primer volumen y en el capítulo anterior, el lector comprenderá claramente el interés de estudiar con detalle las etapas que nos llevarán a la construcción de una tal norma sobre espacios de medidas.

Antes de pasar a la definición de esta norma, llamada la *norma de la variación total*, es necesario definir una medida positiva asociada a la medida en consideración. Para ello vamos a utilizar el razonamiento siguiente: de la misma forma que definimos la cantidad $|f| = f^+ + f^-$ para una función a valores en $\mathbb{R}$, un argumento similar permite considerar, al menos en una primera aproximación, una medida positiva con las características necesarias para nuestros fines.

Empezaremos estudiando medidas con signo para luego pasar a las medidas complejas pues este último caso requiere unas pocas etapas adicionales.

Definición 1.1.7 (Variación, variación total de una medida con signo) *Sea $(X, \mathscr{A})$ un espacio medible dotado de una medida con signo μ.*

1) La variación *de la medida μ es la medida positiva $|\mu|$ definida por la fórmula*

$$|\mu| = \mu^+ + \mu^-, \tag{1.9}$$

en donde μ^+ y μ^- son las medidas positivas asociadas a la descomposición de Jordan de la medida con signo μ.

2) La variación total *de una medida será notada $\|\mu\|_{VT}$ y está determinada por la cantidad*

$$\|\mu\|_{VT} = |\mu|(X). \tag{1.10}$$

Verificar que la aplicación $|\mu|$ definida sobre $(X, \mathscr{A})$ es efectivamente una medida no es muy complicado: basta aplicar la Proposición 1.1.2 a las medidas positivas μ^+ y μ^-.

En el caso cuando $\|\mu\|_{VT} = +\infty$, diremos simplemente que la variación total de la medida μ es infinita y un ejemplo de ello es la medida de Lebesgue dx sobre el espacio medido $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$.

Observación 1.2 Es muy importante notar que

1) no hay que confundir $|\mu(A)|$ con $|\mu|(A)$.

2) se tiene, para todo conjunto $\mathscr{A}$-medible A, las desigualdades

$$|\mu(A)| \leq |\mu|(A) \leq \|\mu\|_{VT}$$

En efecto, sea $(X, \mathscr{A})$ un espacio medible, sea μ una medida con signo definida sobre $\mathscr{A}$ y sean μ^+ y μ^- las medidas positivas asociadas a la descomposición de Jordan de μ. Tenemos entonces:

$$|\mu(A)| = |\mu^+(A) - \mu^-(A)| \leq \mu^+(A) + \mu^-(A) = |\mu|(A),$$

de donde se deduce las desigualdades $|\mu(A)| \leq |\mu|(A) \leq |\mu|(X) = \|\mu\|_{VT}$ válida para todo $A \in \mathscr{A}$ y esta desigualdad puede ser estricta como lo muestra el ejemplo de dos masas de Dirac de signo opuesto: en efecto, si $\mu = -\delta_1 + \delta_{-1}$ entonces para todo conjunto A de la forma $]-a, a[$ con $a > 1$ se tiene $\mu(A) = 0$ y entonces $|\mu(A)| = 0$, pero para este tipo de conjuntos tenemos $|\mu|(A) = \delta_1(A) + \delta_{-1}(A) = 2$ de modo que se tiene $|\mu(A)| < |\mu|(A)$. Veremos un poco más adelante que la medida $|\mu|$ es en realidad la más pequeña medida positiva que verifica esta desigualdad.

Pasemos ahora al caso de medidas complejas, aquí no es posible usar la fórmula (1.9) para definir la medida de variación asociada a una medida compleja pues hay que tener ciudado con las partes reales e imaginarias (pensar por ejemplo en la diferencia que existe al definir el valor absoluto de un número real y el módulo de un número complejo). Para ello es necesario utilizar otro punto de vista.

Definición 1.1.8 (Variación, variación total de una medida compleja) *Sea $(X, \mathscr{A})$ un espacio medible dotado de una medida compleja μ.*

1) *Para todo conjunto medible $A \in \mathscr{A}$ definimos la* variación *de la medida compleja μ como*

$$|\mu|(A) = \sup \sum_{j=0}^{+\infty} |\mu(A_j)| \tag{1.11}$$

en donde el supremo corre sobre todos las particiones[3] *$\mathscr{A}$-medibles $(A_j)_{j\in\mathbb{N}}$.*

2) *La* variación total *de una medida compleja será notada $\|\mu\|_{VT}$ y está determinada por la cantidad $\|\mu\|_{VT} = |\mu|(X)$.*

Como vimos, la verificación de que la variación de una medida con signo es una medida positiva no causa mayor dificultad, sin embargo en el caso de las medidas complejas la situación es ligeramente diferente y es necesario el resultado a continuación.

Proposición 1.1.8 *Sea $(X, \mathscr{A})$ un espacio medible dotado de una medida compleja μ. La variación $|\mu|$ asociada a la medida compleja μ es una medida positiva finita definida sobre $(X, \mathscr{A})$.*

Prueba. Notemos que se tiene directamente que $|\mu|(\emptyset) = 0$, de manera que solo hay que verificar la σ-aditividad. Empecemos comprobando la aditividad finita. Sean pues B_1 y B_2 dos conjuntos disjuntos $\mathscr{A}$-medibles y sea $(A_j)_{1\leq j\leq n}$ una partición finita de $B_1 \cup B_2$ por medio de conjuntos $\mathscr{A}$-medibles. Tenemos entonces que

$$\sum_{1\leq j\leq n} |\mu(A_j)| \leq \sum_{1\leq j\leq n} |\mu(A_j \cap B_1)| + \sum_{1\leq j\leq n} |\mu(A_j \cap B_2)| \leq |\mu|(B_1) + |\mu|(B_2).$$

Por definición obtenemos entonces que la cantidad $|\mu|(B_1 \cup B_2)$ está mayorada por $|\mu|(B_1) + |\mu|(B_2)$. Un argumento totalmente similar muestra que

$$|\mu|(B_1) + |\mu|(B_2) \leq |\mu|(B_1 \cup B_2)$$

[3]Recordemos que una *partición* de un conjunto X es una familia disjunta $(A_i)_{i\in I}$ de subconjuntos de X tal que $\bigcup_{i\in I} A_i = X$.

de donde se deduce la identidad $|\mu|(B_1 \cup B_2) = |\mu|(B_1) + |\mu|(B_2)$ y se obtiene de esta manera la aditividad finita de la aplicación $|\mu|$.

Usamos ahora la descomposición de Jordan de la medida compleja μ para escribir que $\mu = \nu_1^+ - \nu_1^- + i\nu_2^+ - i\nu_2^-$ en donde todas estas medidas son medidas finitas. Dado que se tiene, para todo conjunto $\mathscr{A}$-medible la desigualdad

$$|\mu|(A) \leq \nu_1^+(A) + \nu_1^-(A) + \nu_2^+(A) + \nu_2^-(A) \tag{1.12}$$

se obtiene que la variación $|\mu|$ es una aplicación finita.

Para terminar la demostración solo falta verificar que la función aditiva de conjuntos $|\mu|$ es en realidad una medida. Sea pues $(A_n)_{n\in\mathbb{N}}$ una sucesión decreciente de conjuntos $\mathscr{A}$-medibles tal que $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$. Entonces se tiene que $\lim\limits_{n\to+\infty} \nu_i^\pm(A_n) = 0$ con $i = 1, 2$ por la continuidad de las medidas, de manera que con la desigualdad (1.12) se obtiene $\lim\limits_{n\to+\infty} |\mu|(A_n) = 0$. Aplicando la Proposición 1.1.5 obtenemos que la aplicación $|\mu|$ es en realidad una medida, positiva y finita. ■

Observación 1.3 Notemos que la fórmula (1.9) no puede aplicarse directamente en el caso de las medidas complejas, mientras que si se puede utilizar la expresión (1.11) para definir la medida de variación $|\mu|$ asociada a una medida con signo finita μ. En este sentido esta última fórmula es más general.

La medida de variación $|\mu|$ asociada a una medida, con signo o compleja -ambas finitas-, verifica algunas propiedades que conviene explicitar.

Proposición 1.1.9 *Sea $(X, \mathscr{A})$ un espacio medible.*

1) *Si μ es una medida compleja definida sobre $\mathscr{A}$, entonces la variación asociada $|\mu|$ es la más pequeña de todas las medidas positivas ν definidas sobre $\mathscr{A}$ que verifican $|\mu(A)| \leq \nu(A)$ para todo $A \in \mathscr{A}$.*

2) *Si μ es una medida con signo* finita*, entonces μ es una medida compleja (de parte imaginaria nula) y en este caso la variación de μ como medida con signo coincide con la variación como medida compleja.*

Prueba. Sea ν una medida positiva que verifica $|\mu(A)| \leq \nu(A)$ para todo $a \in \mathscr{A}$. Tenemos entonces

$$\sum_{j\in\mathbb{N}} |\mu(A_j)| \leq \sum_{j\in\mathbb{N}} \nu(A_j) = \nu(A)$$

para toda partición $\mathscr{A}$-medible $(A_j)_{j\in\mathbb{N}}$ de A; de donde se deduce el primer punto.

Al ser todas las cantidades finitas, el segundo punto es una consecuencia de la descomposición de Jordan de las medidas complejas. ■

Veamos ahora cómo es posible relacionar la variación de una medida con la descomposición de Jordan.

Proposición 1.1.10 *Consideremos $(X, \mathscr{A})$ un espacio medible y una medida finita con signo μ definida sobre $\mathscr{A}$. Sea $|\mu|$ la variación asociada a la medida μ. Definimos las medidas*

$$\mu^+ = \frac{1}{2}(|\mu| + \mu) \qquad y \qquad \mu^- = \frac{1}{2}(|\mu| - \mu). \tag{1.13}$$

Estas dos medidas corresponden con la descomposición de Jordan dada en la Definición 1.1.5. Además, si $\mu = \nu_1 - \nu_2$ en donde ν_1, ν_2 son dos medidas finitas positivas entonces $\nu_1 \geq \mu^+$ y $\nu_2 \geq \mu^-$.

Prueba. Notemos para empezar que si la medida μ es una medida finita, entonces las aplicaciones μ^+ y μ^- dadas por la fórmula (1.13) están bien definidas por la Proposición 1.1.2 y son medidas finitas.

Para verificar esta proposición es suficiente usar la descomposición de Jordan y se obtiene directamente las identidades deseadas. La minimalidad de la descomposición $|\mu| = \mu^+ + \mu^-$ se deduce de la proposición anterior. ■

Cuando una medida está definida por medio de una integral tenemos el siguiente resultado que relaciona las medidas con signo finitas y las medidas complejas a su variación.

Proposición 1.1.11 *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida positiva definida sobre él. Sea f una función que pertenece al espacio $\mathcal{L}(X, \mathscr{A}, \mu, \mathbb{R})$ o $\mathcal{L}(X, \mathscr{A}, \mu, \mathbb{C})$ y sea ν una medida con signo finita o una medida compleja definida por la expresión $\nu(A) = \displaystyle\int_A f(x)d\mu(x)$. Entonces se tiene la siguiente relación para todo $A \in \mathscr{A}$:*

$$|\nu|(A) = \int_A |f(x)|d\mu(x). \tag{1.14}$$

Prueba. Sea $A \in \mathscr{A}$ y sea $(A_j)_{1\leq j\leq k}$ una sucesión finita de conjuntos disjuntos $\mathscr{A}$-medibles tales que $\displaystyle\bigcup_{j=1}^{k} A_j = A$. Podemos entonces escribir

$$\sum_{j=1}^{k} |\nu(A_j)| = \sum_{j=1}^{k} \left| \int_A f(x)d\mu(x) \right| \leq \sum_{j=1}^{k} \int_A |f(x)|d\mu(x) = \int_A |f(x)|d\mu(x).$$

Dado que $|\nu|(A)$ es el supremo de las sumas que pueden aparecer en la parte izquierda de esta desigualdad tenemos la mayoración

$$|\nu|(A) \leq \int_A |f(x)|d\mu(x).$$

Para obtener la desigualdad recíproca, consideramos una sucesión de funciones simples $\mathscr{A}$-medibles $(g_n)_{n\in\mathbb{N}}$ tales que $|g_n(x)| = 1$ y tal que se tenga $\lim\limits_{n\to+\infty} g_n(x)f(x) = |f(x)|$ para todo $x \in X$. Sean ahora $a_{n,j}$ con $j = 1, ..., k_n$, los valores tomados por las funciones g_n sobre los conjuntos $A_{n,j}$. Entonces, para un conjunto $A \in \mathscr{A}$ cualquiera tenemos

$$\left| \int_A g_n(x)d\mu(x) \right| = \left| \sum_j a_{n,j} \int_{A\cap A_{n,j}} f(x)d\mu(x) \right| = \left| \sum_j a_{n,j}\nu(A \cap A_{n,j}) \right| \leq \sum_j |\nu(A \cap A_{n,j})| \leq |\nu|(A).$$

Ahora, el teorema de convergencia dominada implica que se tiene $\displaystyle\int_A g_n(x)d\mu(x) \underset{n\to+\infty}{\longrightarrow} \int_A |f(x)|d\mu(x)$ de donde se deduce que $\displaystyle\int_A |f(x)|d\mu(x) \leq |\nu|(A)$. Obtenemos de esta manera la igualdad deseada $|\nu|(A) = \displaystyle\int_A |f(x)|d\mu(x)$. ■

Observación 1.4 La fórmula (1.14) permite ver muy claramente que se tiene la desigualdad

$$|\nu(A)| = \left| \int_A f(x)dx \right| \leq \int_A |f(x)|dx = |\nu|(A).$$

Para terminar esta sección, vamos a ver cómo dotar de una estructura al conjunto de medidas definidas sobre un mismo espacio medible $(X, \mathscr{A})$. En efecto, hemos visto con la Proposición 1.1.2 y su versión compleja, la Proposicion 1.1.7, que si μ y ν son dos medidas con signo o complejas definidas sobre el mismo espacio medible $(X, \mathscr{A})$ entonces es posible definir, para todo conjunto $A \in \mathscr{A}$ y todo escalar $c \in \mathbb{K}$, las operaciones siguientes

$$(\mu + \nu)(A) = \mu(A) + \nu(A) \tag{1.15}$$
$$(c\mu)(A) = c\mu(A), \tag{1.16}$$

de manera que las medidas asociadas $\mu + \nu$ y $c\mu$ son medidas con signo o medidas complejas, según sea el caso, y esto sugiere que se dispone de una estructura vectorial. Más precisamente tenemos

Definición 1.1.9 (Espacios de Medidas, Norma de la Variación Total) *Sea $(X, \mathscr{A})$ un espacio medible.*

1) Definimos el conjunto $\mathscr{M}_{\mathbb{R}}(X, \mathscr{A})$ como el conjunto formado por todas las medidas con signo finitas definidas sobre $(X, \mathscr{A})$.

2) De la misma manera definimos $\mathscr{M}_{\mathbb{C}}(X, \mathscr{A})$ como el conjunto formado por todas las medidas complejas definidas sobre $(X, \mathscr{A})$.

Estos dos conjuntos son espacios vectoriales para las operaciones (1.15) y (1.16); además la variación total $\|\cdot\|_{VT}$ es una norma, también llamada la norma de la Variación Total*, sobre estos espacios.*

Insistamos sobre el hecho que, cuando consideramos el espacio de medidas con signo $\mathscr{M}_{\mathbb{R}}(X, \mathscr{A})$, exigimos que estas medidas sean finitas pues de otra manera la suma (1.15) puede estar mal definida. Esta precaución es innecesaria en el caso de las medidas complejas pues todas ellas son finitas por definición.

Vemos entonces que toda medida real de variación total finita, como la masa de Dirac, pertenece al espacio $\mathscr{M}_{\mathbb{R}}(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$, mientras que la medida de Lebesgue, al no ser de variación total finita, no pertenece a este conjunto.

Verifiquemos rápidamente que la variación total $\|\cdot\|_{VT}$ es efectivamente una norma sobre el espacio $\mathscr{M}_{\mathbb{K}}(X, \mathscr{A})$ con $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$. Vemos sin problema, por definición, que $\|\mu\|_{VT} \geq 0$ y que $\|\mu\|_{VT} = 0$ implica que μ es nula sobre todo conjunto $\mathscr{A}$-medible. Dado que la homogenedidad $\|c\mu\|_{VT} = |c|\|\mu\|_{VT}$ no causa mayor problema, nos concentramos en la desigualdad triangular: sean μ y ν dos medidas pertenecientes al espacio $\mathscr{M}_{\mathbb{K}}(X, \mathscr{A})$ y sea $(A_n)_{n\in\mathbb{N}}$ una partición disjunta de conjuntos $\mathscr{A}$-medibles de X. Entonces tenemos

$$\sum_{n\in\mathbb{N}} |(\mu+\nu)(A_n)| \leq \sum_{n\in\mathbb{N}} |\mu(A_n)| + \sum_{n\in\mathbb{N}} |\nu(A_n)| \leq |\mu|(X) + |\nu|(X) = \|\mu\|_{VT} + \|\nu\|_{VT},$$

de donde se deduce $\|\mu + \nu\|_{VT} \leq \|\mu\|_{VT} + \|\nu\|_{VT}$, de manera que $\|\cdot\|_{VT}$ es una norma.

Disponer de una estructura de espacio vectorial normado es sumamente interesante, pero como hemos visto en el Capítulo 1 del primer volumen, esto no siempre es suficiente y es interesante saber si se dispone de mayor estructura sobre este tipo de espacios. En el caso de los espacios definidos anteriormente, tenemos el resultado a continuación que nos permite describir más en detalle sus propiedades:

Teorema 1.1.2 (Completitud de los espacios de medidas) *Sea $(X, \mathscr{A}, \mu)$ un espacio medido. Entonces los espacios $(\mathscr{M}_{\mathbb{R}}(X, \mathscr{A}), \|\cdot\|_{VT})$ y $(\mathscr{M}_{\mathbb{C}}(X, \mathscr{A}), \|\cdot\|_{VT})$ son espacios normados completos, es decir que son espacios de Banach.*

Demostración. Sea $(\mu_n)_{n\in\mathbb{N}}$ una sucesión de Cauchy de medidas, ya sea en $(\mathscr{M}_{\mathbb{R}}(X,\mathscr{A}), \|\cdot\|_{VT})$ o en $(\mathscr{M}_{\mathbb{C}}(X,\mathscr{A}), \|\cdot\|_{VT})$. Vamos a mostrar que si definimos una aplicación (real o compleja) μ sobre $\mathscr{A}$ por medio de la expresión $\mu(A) = \lim\limits_{n\to+\infty}\mu_n(A)$ para todo $A \in \mathscr{A}$, entonces se tiene que μ es una medida (con signo o compleja) que verifica la relación $\lim\limits_{n\to+\infty}\|\mu-\mu_n\|_{VT} = 0$.

Para empezar observamos que la desigualdad $|\mu_m(A)-\mu_n(A)| \leq \|\mu_m-\mu_n\|_{VT}$ implica que, para todo $A \in \mathscr{A}$, la sucesión $(\mu_n(A))_{n\in\mathbb{N}}$ es una sucesión de Cauchy de números reales o complejos, y es por lo tanto una sucesión convergente, lo que justifica la definición que hemos hecho de la aplicación límite μ.

Vemos luego sin mayor problema que $\mu(\emptyset) = 0$ y que se tiene la aditividad finita, de manera que esta aplicación μ es una aplicación finitamente aditiva. Para demostrar que μ es en realidad una medida vamos a utilizar las Proposiciones 1.1.4 y 1.1.5. Para ello veamos que la convergencia de la cantidad $\mu_n(A)$ hacia $\mu(A)$ es uniforme en A. Sea $\varepsilon > 0$ un real y sea $N \in \mathbb{N}$ un entero tal que $\|\mu_m-\mu_n\|_{VT} < \varepsilon$ si $m, n \geq N$. Tenemos entonces que $|\mu_m(A) - \mu_n(A)| < \varepsilon$ para todo $A \in \mathscr{A}$ con $m, n \geq N$, de manera que haciendo $m \longrightarrow +\infty$, se tiene $|\mu(A) - \mu_n(A)| < \varepsilon$ para todo $A \in \mathscr{A}$ con $n \geq N$. Como ε era arbitrario se obtiene la convergencia uniforme de $\mu_n(A)$ hacia $\mu(A)$ para todo $A \in \mathscr{A}$.

Sea ahora $(A_n)_{k\in\mathbb{N}}$ una sucesión decreciente de conjuntos $\mathscr{A}$-medibles tal que $\bigcap_{k\in\mathbb{N}} A_k = \emptyset$ y sea $\varepsilon > 0$ un real. Usando la convergencia uniforme de la cantidad $\mu_n(A)$ hacia $\mu(A)$, podemos fijar un entero N tal que se tenga $|\mu(A) - \mu_n(A)|_{VT} < \varepsilon/2$ para todo $A \in \mathscr{A}$ y para todo $n \geq N$. Usamos ahora la Proposición 1.1.4 para fijar un entero K tal que $|\mu_N(A_k)| < \varepsilon/2$ siempre y cuando $k \geq K$. Tenemos entonces, si $k \geq K$, que

$$|\mu(A_k)| \leq |\mu(A_k) - \mu_N(A_k)| + |\mu_N(A_k)| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$$

de donde se deduce que $\lim\limits_{k\to+\infty}\mu(A_k) = 0$ y por lo tanto que la aplicación límite μ es una medida.

Pasemos ahora a verificar que se tiene $\lim\limits_{n\to+\infty}\|\mu-\mu_n\|_{VT} = 0$. Sea $\varepsilon > 0$ un real y sea $N \in \mathbb{N}$ tal que $\|\mu_m-\mu_n\|_{VT} < \varepsilon$ para todo $m, n \geq N$. Observemos que si $m, n \geq N$ se tiene que cada partición $\mathscr{A}$-medible $(A_j)_{1\leq j\leq k}$ de X verifica

$$\sum_{j=1}^{k} |\mu_m(A_j) - \mu_n(A_j)| \leq \|\mu_m - \mu_n\|_{VT} < \varepsilon$$

y por lo tanto verifica

$$\lim_{m\to+\infty}\sum_{j=1}^{k} |\mu_m(A_j) - \mu_n(A_j)| = \sum_{j=1}^{k} |\mu(A_j) - \mu_n(A_j)| \leq \varepsilon$$

Como la cantidad $\|\mu-\mu_n\|_{VT}$ es el supremo de los números que aparecen en la parte izquierda esta desigualdad, se tiene que $\|\mu-\mu_n\|_{VT} \leq \varepsilon$ siempre y cuando $n \geq N$. De donde se deduce que $\lim\limits_{n\to+\infty}\|\mu-\mu_n\|_{VT} = 0$.

Hemos demostrado que el límite de toda sucesión de Cauchy de medidas con signo o compleja, es una medida con signo o compleja y por lo tanto que los espacios normados $(\mathscr{M}_{\mathbb{R}}(X,\mathscr{A}), \|\cdot\|_{VT})$ y $(\mathscr{M}_{\mathbb{C}}(X,\mathscr{A}), \|\cdot\|_{VT})$ son espacios de Banach. ■

Con este resultado terminamos nuestra presentación sobre las propiedades de las medidas con signo y de las medidas complejas. El lector observará con justa razón que, a la luz del capítulo anterior, una vez que se dispone de una estructura de espacios de Banach, es posible estudiar otro tipo de propiedades (separabilidad, dualidad, reflexividad). Estos temas serán tratados posteriormente y de forma unificada un poco más adelante.

1.1.4. Integración con respecto a una medida con signo o medida compleja

En esta sección damos un rápido recuento de la integración con respecto a las medidas con signo (finitas o no) o con respecto a las medidas complejas. Usando la descomposición de Jordan de las medidas complejas, vemos que es suficiente concentrarse en las medidas con signo pues el paso al caso de las medidas complejas es inmediato.

Para empezar consideramos un espacio medido $(X, \mathscr{A}, \mu)$ en donde μ es una medida con signo. Dado que la noción de funciones medibles y de funciones simples es la misma (recordar las definiciones del Capítulo 3 del primer volumen), podemos definir sin mayor problema la integral con respecto a esta medida. En efecto, para una función simple $f : X \longrightarrow \mathbb{K}$ con $\mathbb{K} = \mathbb{R}, \mathbb{C}$ tal que $f(x) = \sum_{k=1}^{n} \alpha_k \mathbb{1}_{A_k}(x)$ donde $(A_k)_{1\leq k\leq n}$ es una familia disjunta de conjuntos $\mathscr{A}$-medibles y $\alpha_k \in \mathbb{K}$, definimos su integral con respecto a la medida con signo μ de la siguiente forma:

$$\begin{aligned}
\int_X f(x)d\mu(x) &= \sum_{1\leq k\leq n} \alpha_k \mu(A_k) = \sum_{1\leq k\leq n} \alpha_k \mu^+(A_k) - \sum_{1\leq k\leq n} \alpha_k \mu^-(A_k) \\
&= \int_X f(x)d\mu^+(x) - \int_X f(x)d\mu^-(x).
\end{aligned}$$

Como vemos, la descomposición $\mu = \mu^+ - \mu^-$ permite considerar la integral de una función con respecto a una medida con signo como la diferencia de dos integrales usuales y esto nos conduce a la siguiente definición:

Definición 1.1.10 (Integral con respecto a una medida con signo) *Sea $(X, \mathscr{A}, \mu)$ un espacio medido en donde μ es una medida con signo y sea $f : X \longrightarrow \mathbb{R}$ una función $\mathscr{A}$-medible. Definimos entonces la integral de la función f con respecto a la medida con signo μ por medio de la expresión*

$$\int_X f(x)d\mu(x) = \int_X f(x)d\mu^+(x) - \int_X f(x)d\mu^-(x)$$

De forma equivalente, si (P, N) es una descomposición de Hahn del conjunto X asociada a la medida μ y si ε es una función igual a 1 sobre P y -1 sobre N (es decir $\varepsilon(x) = \mathbb{1}_P(x) - \mathbb{1}_N(x)$), definimos

$$\int_X f(x)d\mu(x) = \int_X f(x)\varepsilon(x)d|\mu|(x) \tag{1.17}$$

y en donde $|\mu|$ es la variación asociada a la medida μ.

Esta definición de integral se extiende sin problemas al caso cuando la medida μ es una medida con signo y la función considerada es a valores en $\mathbb{C}$: basta considerar las partes reales e imaginarias para escribir, para una función $f : X \longrightarrow \mathbb{C}$:

$$\int_X f(x)d\mu(x) = \int_X \Re e(f)(x)d\mu(x)+i\int_X \Im m(f)(x)d\mu(x) = \int_X \Re e(f)(x)\varepsilon(x)d|\mu|(x)+i\int_X \Im m(f)\varepsilon(x)d|\mu|(x)$$

El caso cuando tanto la medida como la función son complejas se estudia de forma totalmente similar separando las partes reales e imaginarias: si $f : X \longrightarrow \mathbb{C}$ es una función $\mathscr{A}$-medible y si μ es una medida compleja definida sobre $\mathscr{A}$, entonces dado que $f = \Re e(f) + i\Im m(f)$ y que $\mu = \Re e(\mu) + i\Im m(\mu)$ podemos escribir:

$$\begin{aligned}
\int_X f(x)d\mu(x) &= \int_X \big(\Re e(f) + i\Im m(f)\big)(x)d\big(\Re e(\mu) + i\Im m(\mu)\big)(x) \\
&= \int_X \Re e(f)(x)d\Re e(\mu)(x) - \int_X \Im m(f)(x)d\Im m(\mu)(x) \qquad (1.18)\\
&\quad +i\left[\int_X \Re e(f)(x)d\Im m(\mu)(x) + \int_X \Im m(f)(x)d\Re e(\mu)(x)\right]
\end{aligned}$$

Cada una de estas integrales que aparecen en la expresión (1.18) es una integral con respecto a una medida real con signo: podemos entonces aplicar sin ningún problema la fórmula (1.17) de manera a restringirse únicamente al estudio de integrales con respecto a las medidas positivas usuales estudiadas en el primer volumen.

Esto nos conduce a la siguiente definición.

Definición 1.1.11 (Espacio de funciones integrables) *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida con signo o compleja. Notaremos $\mathcal{I}(X, \mathscr{A}, \mu, \mathbb{K})$ el conjunto de funciones tales que la cantidad (1.17) y/o (1.18) sea finita.*

La principal utilidad de la definición (1.17) y de la fórmula (1.18) tiene que ver con el hecho que estas expresiones nos permiten ver rápidamente algunas de las propiedades de la integral con respecto a una medida con signo o compleja: al menos en una primera aproximación, es suficiente descomponer las integrales en términos de las partes positivas/negativas o reales/imaginarias de las funciones y de las medidas para obtener objetos conocidos.

Observación 1.5 El hecho de descomponer estas integrales es muy útil pues nos permite estudiar estos objetos por medio de integrales conocidas. Pero esta apariencia no debe desorientar al lector y hay que tener mucho cuidado con ciertos resultados que son falsos en el caso de las medidas con signo o en el caso de las medidas complejas.

En efecto, en el caso de las medidas con signo, y menos aún en el caso de las medidas complejas, no se dispone de la importante propiedad de crecimiento: si $f, g : X \longrightarrow \overline{\mathbb{R}}_+$ son dos funciones tales que $f(x) \leq g(x)$ para todo $x \in X$, no se tiene necesariamente que $\displaystyle\int_X f(x)d\mu(x) \leq \int_X g(x)d\mu(x)$ y esta particularidad impide el uso directo de algunos resultados.

En particular:

- El teorema de convergencia monótona de Beppo Levi (ver el Teorema 3.3.1 del primer volumen) falla en el caso de las medidas con signo y de las medidas complejas.
- El lema de Fatou (ver el Teorema 3.3.2 del primer volumen) falla para las medidas con signo.

Veremos en los ejercicios 1.11 y 1.12 al final de este capítulo ejemplos concretos en donde estos resultados no son válidos.

Veamos ahora qué tipo de propiedades de mantienen. Una de ellas, de gran importancia, es el teorema de convergencia dominada de Lebesgue:

Teorema 1.1.3 (convergencia dominada para medidas con signo o complejas) *Sea $(X, \mathscr{A}, \mu)$ un espacio medido en donde μ es una medida con signo o compleja. Sean respectivamente f y $(f_n)_{n\in\mathbb{N}}$ una función y una sucesión de funciones, ambas $\mathscr{A}$-medibles definidas sobre X a valores en $\mathbb{K}$. Hacemos las siguientes hipótesis:*

1) para μ-casi todo $x \in X$ se tiene $\displaystyle\lim_{n\to+\infty} f_n(x) = f(x)$; y

2) existe una función $|\mu|$-integrable $g : X \longrightarrow \overline{\mathbb{R}}_+$ tal que, para todo n, se verifica la desigualdad $|f_n(x)| \leq g(x)$ en μ-casi todas partes.

Entonces f es una función μ-integrable y

$$\lim_{n\to+\infty} \int_X f_n(x)d\mu(x) = \int_X f(x)d\mu(x). \tag{1.19}$$

Demostración. Consideremos primero una medida μ con signo y verifiquemos que la función límite f es μ-integrable. Para ello basta escribir:

$$\left|\int_X f(x)d\mu(x)\right| = \left|\int_X f(x)\varepsilon(x)d|\mu|(x)\right| \leq \int_X |f(x)|d|\mu|(x) \leq \int_X g(x)d|\mu|(x) < +\infty.$$

La identidad (1.19) es inmediata una vez que se tiene la identidad

$$\int_X f_n(x)d\mu(x) = \int_X f_n(x)d\mu^+(x) - \int_X f_n(x)d\mu^-(x)$$

pues entonces basta aplicar el teorema de convergencia dominada de Lebesgue a cada una de las integrales anteriores: pasando al límite se obtiene el resultado deseado.

Para el caso de una medida μ compleja, basta usar la descomposición (1.18) y aplicar el mismo razonamiento a cada una de las partes de esta descomposición. ■

Definición 1.1.12 (Espacios de Lebesgue respecto a una medida con signo) *Sea $(X, \mathscr{A}, \mu)$ un espacio medido en donde μ es una medida con signo. Para todo $0 < p \leq +\infty$ definimos el conjunto de funciones de módulo p-eme integrables con respecto a la medida con signo μ como $L^p(X, \mathscr{A}, |\mu|, \mathbb{K})$. Este espacio puede ser normado por medio de la funcional*

$$\|f\|_{L^p} = \left(\int_X |f(x)|^p d|\mu|dx\right)^{1/p}$$

con las modificaciones usuales si $p = +\infty$, es decir:

$$\|f\|_{L^\infty} = \sup_{x\in X} ess|f(x)| = \text{ínf}\left\{c \in \overline{\mathbb{R}}_+ : |\mu|\big(\{x \in X : |f(x)| > c\}\big) = 0\right\}$$

Esta definición se extiende sin problemas al caso de las medidas complejas.

Observación 1.6 Las diversas propiedades de los espacios de Lebesgue se mantienen al considerar medidas con signo o medidas complejas.

En efecto, cuando se dispone de la definición (1.17) de integral con respecto a una medida con signo y de la descomposición (1.18) en el caso de las medidas complejas, al considerar los espacios de Lebesgue tal como en la Definición 1.1.12 notamos que la medida de variación $d|\mu|$ "anula" las variaciones de signo y las partes imaginarias de las medidas consideradas de manera que toda la problemática se reduce a estudiar los espacios de Lebesgue con respecto a medidas positivas usuales.

En particular tenemos las siguientes propiedades que son enunciadas sin demostración. Rogamos al lector ver los detalles en el Capítulo 4 del primer volumen.

Proposición 1.1.12 *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida (con signo o compleja) definida sobre $\mathscr{A}$. Entonces, los espacios de Lebesgue $L^p(X, \mathscr{A}, |\mu|, \mathbb{K})$ son*

1) espacios de Banach si $1 \leq p \leq +\infty$.

2) espacios separables si $1 \leq p < +\infty$.

1.2. Continuidad absoluta

Sabemos que si $(X, \mathscr{A}, \mu)$ es un espacio medido con μ una medida positiva y si f es una función integrable definida sobre un conjunto X a valores en $\mathbb{R}_+$, entonces es posible a partir de estos dos ingredientes construir una nueva medida ν por medio de la expresión

$$\nu(A) = \int_A f(x)d\mu(x) \qquad (\forall A \in \mathscr{A}). \tag{1.20}$$

Esta expresión puede resultar muy cómoda: en efecto, si una medida ν que admite una descomposición por medio de la fórmula (1.20) será mucho más fácil y directo utilizar esta caracterización de la medida ν si tanto la función f como la medida μ son objetos conocidos y simples de manipular. Este tipo de situaciones son esenciales en la teoría de probabilidades y serán fundamentales en los capítulos siguientes.

Esta medida ν depende, evidentemente, de las propiedades de la función f y de la medida μ. Así por ejemplo si f es una función μ-integrable positiva y si μ es una medida positiva, entonces la medida ν también será una medida positiva. Usando el material desarrollado hasta aquí, el lector no tendrá dificultad en deducir propiedades similares en función de los datos f y μ; en efecto, bajo ciertas condiciones, si la función f toma valores en $\mathbb{R}$ o $\mathbb{C}$, entonces la medida resultante ν es una medida con signo o una medida compleja.

Volvamos ahora a la expresión (1.20) que contiene una particularidad que conviene resaltar: vemos sin mayor dificultad que si A es un conjunto tal que $\mu(A) = 0$, entonces se tiene inmediatamente que $\nu(A) = 0$ puesto que $\nu(A) = \int_X f(x)\mathbb{1}_A(x)d\mu(x)$ y en toda esta sección insistiremos en esta implicación y su recíproca pues en ellas se encierra la noción de base de la *continuidad absoluta* como tendremos la oportunidad de estudiarlo.

Nuestro principal objetivo en esta sección es entonces el siguiente: explicar bajo qué condiciones entre dos medidas μ y ν se tiene la existencia de una función $\mathscr{A}$-medible f tal que la medida ν se pueda expresar en función de la medida μ por medio de la fórmula (1.20). Dicho de otra manera, buscamos los requerimientos necesarios para descomponer una medida ν en los dos elementos f y μ y, en este sentido, deseamos hacer el camino inverso realizado en la Sección 1.1 anterior.

Veremos pues en todo lo que sigue qué es lo que previamente se necesita para obtener esta relación y cuáles son sus consecuencias. Una primera etapa para realizar nuestro objetivo consiste en estudiar, con un poco más de detalle (y con otra terminología), un punto que ha sido tratado en la sección anterior.

Definición 1.2.1 (Producto de una medida por una función) *Sea $(X, \mathscr{A})$ un espacio medible, sea μ una medida positiva definida sobre $\mathscr{A}$ y sea $f : X \longrightarrow \overline{\mathbb{R}}_+$ una función $\mathscr{A}$-medible. El producto de la medida μ por la función f define una nueva medida, notada $(f\mu)$, que está definida por:*

$$(f\mu)(A) = \int_A f(x)d\mu(x) \tag{1.21}$$

para todo $A \in \mathscr{A}$.

Notamos la total similitud de esta medida con la definición de la medida inducida por una función dada en la expresión (1.20). En esta definición nos hemos restringido, por simplicidad, al caso de funciones y medidas positivas y es evidentemente posible considerar variantes con signo o complejas tomando las precauciones necesarias.

La definición de la integral con respecto a la medida $(f\mu)$ no causa ninguna dificultad y está dada de la siguiente manera: si $g : X \longrightarrow \mathbb{R}$ es una función $(f\mu)$-integrable, definimos

$$\int_X g(x)d(f\mu)(x) = \int_X g(x)f(x)d\mu(x).$$

Verificar que todas estas definiciones tienen sentido no es muy complicado (recordar la Definición 3.3.1 del primer volumen y los resultados tratados anteriormente en este capítulo), de manera que dejamos los detalles al lector.

Más allá de la terminología, nos interesamos aquí en las posibles relaciones existentes entre una función f y las medidas μ y $(f\mu)$.

Proposición 1.2.1 *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida σ-finita positiva definida sobre $\mathscr{A}$. Para $f, \psi : X \longrightarrow \overline{\mathbb{R}}_+$ dos funciones $\mathscr{A}$-medibles consideramos $(f\mu)$ y $(\psi\mu)$ el producto entre estas funciones y la medida μ. Si se tiene $(f\mu)(A) = (\psi\mu)(A)$ para todo $A \in \mathscr{A}$, entonces se tiene la igualdad $f = \psi$ en μ-casi todas partes.*

Prueba. Dado que las funciones f y ψ tienen roles similares, por simetría es suficiente mostrar que se tiene la desigualdad $f \leq \psi$ en μ-casi todas partes. Definimos entonces los siguientes conjuntos:

$$A = \{x \in X : \psi(x) < f(x) < +\infty\} \quad \text{y} \quad B = \{x \in X : \psi(x) < f(x) = +\infty\},$$

de manera que se tiene $\{x \in X : \psi(x) < f(x)\} = A \cup B$. Vamos a mostrar que $\mu(A) = \mu(B) = 0$. Para ello consideramos, para $n \geq 1$, los conjuntos

$$A_n = \{x \in X : \psi(x) + 1/n \leq f(x) \leq n\}.$$

Vemos entonces que se tiene $A = \bigcup_{n\geq 1} A_n$, de manera que para mostrar que $\mu(A) = 0$ es suficiente verificar, por la continuidad de las medidas, que $\mu(A_n) = 0$ para todo $n \geq 1$. Procederemos por el absurdo y vamos a suponer que se tiene $\mu(A_n) > 0$ para algún $n \in \mathbb{N}$. En este caso, dado que la medida es σ-finita, debe existir un conjunto $\mathscr{A}$-medible C tal que $C \subset A_n$ y tal que $0 < \mu(C) < +\infty$. Entonces se tiene

$$(f\mu)(C) = \int_X f(x)\mathbb{1}_C(x)d\mu(x) \geq \int_X (\psi(x) + \frac{1}{n})(x)\mathbb{1}_C(x)d\mu(x) = (\psi\mu)(C) + \frac{1}{n}\mu(C)$$

es decir que se tiene la desigualdad $(f\mu)(C) \geq (\psi\mu)(C) + \frac{1}{n}\mu(C)$, pero dado que se tiene por hipótesis que $(f\mu)(C) = (\psi\mu)(C)$, esta desigualdad anterior implica que $(f\mu)(C) = +\infty$. Por otro lado, se tiene

$$(f\mu)(C) = \int_X f(x)\mathbb{1}_C(x)d\mu(x) \leq \int_X n\mathbb{1}_C(x)d\mu(x) = n\mu(C)$$

de donde se deduce que $(f\mu)(C) < +\infty$ y se obtiene de esta manera la contradicción buscada. La prueba que $\mu(B) = 0$ es absolutamente similar y es dejada al lector. ■

Este resultado nos dice que si fijamos una medida σ-finita μ y consideramos la medida $(f\mu)$ entonces la función f está totalmente determinada módulo los conjuntos de μ-medida nula. Este es un primer paso en el estudio de cierto tipo de unicidad en la representación de medidas por medio de la expresión (1.20).

Observación 1.7 Hemos definido el producto de una medida por una función, pero no consideramos el producto de dos medidas.

Para seguir adelante necesitaremos una definición adicional que será la base de los próximos resultados.

Definición 1.2.2 (Continuidad absoluta de medidas, equivalencia de medidas) *Sea $(X, \mathscr{A})$ un espacio medible y sean μ y ν dos medidas positivas definidas sobre $\mathscr{A}$.*

1) *Diremos que la medida ν es* absolutamente continua con respecto a la medida μ*, lo que será notado por $\nu \ll \mu$, si, para todo conjunto $A \in \mathscr{A}$ tal que $\mu(A) = 0$ se tiene $\nu(A) = 0$.*

2) *Diremos que la medida ν es* equivalente *a la medida μ si se tiene $\nu(A) = 0$ si y si solo si $\mu(A) = 0$ para un conjunto $A \in \mathscr{A}$.*

Demos algunos ejemplos. Si $(X, \mathscr{A})$ es un espacio medible, si μ es una medida positiva definida sobre $\mathscr{A}$ y si $f : X \longrightarrow \overline{\mathbb{R}}_+$ es una función medible, entonces se tiene por construcción que $(f\mu) \ll \mu$, es decir que la medida $(f\mu)$ es absolutamente continua con respecto a la medida μ.

Se ve también con este ejemplo que no se tiene necesariamente que $\mu \ll (f\mu)$: en efecto, si consideramos la medida de Lebesgue $\mu = dx$ sobre la recta real $\mathbb{R}$ y fijamos $f(x) = \mathbb{1}_{[0,1]}(x)$, entonces para el conjunto $A = [2, 3]$ se tiene $(f\mu)(A) = \displaystyle\int_A \mathbb{1}_{[0,1]}(x)dx = 0$ pero la medida de Lebesgue del conjunto A es $|A| = 1$.

Mostremos un ejemplo de medidas equivalentes: si $(X, \mathscr{A})$ es un espacio medible y si $f : X \longrightarrow \mathbb{R}$ es una función medible tal que $f > \epsilon > 0$ en μ-casi todas partes, entonces se tiene que las medidas μ y $(f\mu)$ son equivalentes. En efecto, dado que se tiene $(f\mu) \ll \mu$, solo debemos verificar que $\mu \ll (f\mu)$ y procedemos por el absurdo: sea A un conjunto medible tal que $(f\mu)(A) = 0$ pero tal que $\mu(A) \neq 0$, entonces $(f\mu)(A) = \displaystyle\int_A f(x)d\mu(x) \geq \epsilon \int_A d\mu(x) = \mu(A)$, de donde se obtiene la contradicción.

Para terminar con estos ejemplos veamos ahora un ejemplo de medidas que no son equivalentes: consideremos el espacio medible $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y sean δ_0 la masa de Dirac en el punto 0 y dx la medida de Lebesgue, se verifica entonces sin mayor dificultad que estas medidas no son equivalentes: la medida de Lebesgue no carga los puntos.

El siguiente resultado permite caracterizar las medidas positivas finitas que son absolutamente continuas con respecto a medidas positivas arbitrarias.

Proposición 1.2.2 *Sea $(X, \mathscr{A})$ un espacio medible, μ una medida positiva definida sobre $\mathscr{A}$ y sea ν una medida positiva finita definida sobre $\mathscr{A}$. Entonces se tiene $\nu \ll \mu$ si y solo si para todo real positivo $\varepsilon > 0$, existe $\delta > 0$ tal que si $\mu(A) < \delta$ para todo conjunto $\mathscr{A}$-medible A se tiene $\nu(A) < \varepsilon$.*

Esta proposición es interesante pues justifica la denominación de *continuidad absoluta* dada a la noción expresada por $\nu \ll \mu$: en cierto sentido la medida ν está "controlada" por la medida μ. Volveremos a este punto en la Sección 1.3.3.

Prueba. Empecemos suponiendo que para todo ε existe un real δ correspondiente. Sea A un conjunto $\mathscr{A}$-medible tal que $\mu(A) = 0$. Entonces, se tiene para todo δ que $\mu(A) < \delta$ y se tiene $\nu(A) < \varepsilon$ para todo ε, de donde se deduce que $\nu(A) = 0$. Es decir que se tiene la implicación $\mu(A) = 0 \Longrightarrow \nu(A)$, es decir que la medida ν es absolutamente continua con respecto a la medida μ.

Demostremos la recíproca y supongamos ahora que existe un número positivo ε para el cual no existe un real δ correspondiente. Entonces para todo entero k escojemos un conjunto $\mathscr{A}$-medible A_k

que verifica $\mu(A_k) < 1/2^k$ y tal que $\nu(A_k) \geq \varepsilon$ y obtenemos que las desigualdades

$$\mu\left(\bigcup_{k=n}^{+\infty} A_k\right) \leq \sum_{k=n}^{+\infty} \mu(A_k) < \frac{1}{2^{n-1}} \quad \text{y} \quad \nu\left(\bigcup_{k=n}^{+\infty} A_k\right) \geq \nu(A_n) \geq \varepsilon$$

se verifican para todo n. Ahora, si definimos el conjunto $A = \bigcap_{n=1}^{+\infty} \bigcup_{k=n}^{+\infty} A_k$, tenemos que $\mu(A) = 0$ y $\nu(A) \geq \varepsilon$ por la propiedad de continuidad de las medidas. Por lo tanto tenemos un conjunto que satisface $\mu(A) = 0$ pero para el cual se tiene $\nu(A) > 0$, es decir que la medida ν no es absolutamente continua con respecto a μ. ■

1.2.1. Teorema de Radon-Nikodym

En esta subsección enunciamos y demostramos un resultado de gran importancia en el análisis matemático y en las probabilidades. Veremos que el teorema de Radon[4]-Nikodym[5] responde a nuestro objetivo de saber bajo qué condiciones una medida puede escribirse como una integral con respecto a otra medida en el sentido de la fórmula (1.20) y veremos cómo la noción de continuidad absoluta es fundamental en este resultado. Las diversas aplicaciones de este resultado se evidenciarán en los capítulos siguientes.

Teorema 1.2.1 (Radon - Nikodym) *Sea $(X, \mathscr{A})$ un espacio medido y sean μ, ν dos medidas positivas σ-finitas definidas sobre $(X, \mathscr{A})$. Si ν es absolutamente continua con respecto a μ entonces existe una función $\mathscr{A}$-medible $f : X \longrightarrow \mathbb{R}_+$ tal que se tenga la expresión*

$$\nu(A) = \int_A f(x) d\mu(x) \qquad \text{para todo } A \in \mathscr{A}.$$

Además esta función f es única modulo la igualdad μ-casi todas partes.

Demostración. Consideremos para empezar el caso cuando ambas medidas μ y ν son finitas, el caso de medidas σ-finitas será tratado posteriormente. Definimos el conjunto $\mathscr{M}$ como la colección de funciones $\mathscr{A}$-medibles $\varphi : X \longrightarrow \mathbb{R}_+$ que verifican $\displaystyle\int_A \varphi(x) d\mu(x) \leq \nu(A)$ para todo $A \in \mathscr{A}$. Nótese que este conjunto no es vacío pues la función idénticamente nula pertenece a $\mathscr{M}$.

Vamos a descomponer la demostración de este teorema en tres etapas: primero verificaremos que el conjunto $\mathscr{M}$ contiene una función f tal que

$$\int_X f d\mu = \sup\left\{\int_X \varphi d\mu : \varphi \in \mathscr{M}\right\}. \tag{1.22}$$

Una vez que tenemos esta función f, veremos que ésta satisface la identidad $\nu(A) = \displaystyle\int_A f(x) d\mu(x)$ para todo $A \in \mathscr{A}$. Finalmente verificaremos que esta función f puede ser modificada de manera que solo tome valores finitos.

Pasemos pues a la verificación de la primera etapa. Para ello mostramos que si φ_1 y φ_2 pertenecen al conjunto $\mathscr{M}$, entonces $\sup\{\varphi_1, \varphi_2\}$ también pertenece a $\mathscr{M}$. En efecto, si $A \in \mathscr{A}$ es un conjunto medible cualquiera, y si definimos $A_1 = \{x \in A : \varphi_1(x) > \varphi_2(x)\}$ y $A_2 = \{x \in A : \varphi_2(x) \geq \varphi_1(x)\}$, entonces tenemos

$$\int_A \sup\{\varphi_1, \varphi_2\}(x) d\mu(x) = \int_{A_1} \varphi_1(x) d\mu(x) + \int_{A_2} \varphi_2(x) d\mu(x) \leq \nu(A_1) + \nu(A_2) = \nu(A).$$

[4] Johann Radon (1887-1956), matemático austríaco.
[5] Otto Nikodym (1887-1974), matemático polaco.

Ahora escogemos una sucesión $(\varphi_n)_{n\in\mathbb{N}}$ de funciones de $\mathscr{M}$ tales que

$$\lim_{n\to+\infty}\int_X \varphi_n(x)d\mu(x) = \sup\left\{\int_X \varphi(x)d\mu(x) : \varphi\in\mathscr{M}\right\}.$$

Nótese que es posible reemplazar φ_n por $\sup\{\varphi_1,...,\varphi_n\}$ de manera que podemos suponer que la sucesión $(\varphi_n)_{n\in\mathbb{N}}$ es creciente. En este punto definimos $f=\lim\limits_{n\to+\infty}\varphi_n$ y el Teorema de Convergencia Monótona nos asegura que se tiene la expresión

$$\int_A f(x)d\mu(x) = \lim_{n\to+\infty}\int_A \varphi_n(x)d\mu(x)\leq \nu(A)$$

para todo conjunto medible A, de manera que la función f pertenece al conjunto $\mathscr{M}$ y además se tiene que f puede escribirse por medio de la expresión (1.22).

Ataquemos ahora la segunda etapa, es decir que deseamos demostrar que esta función f verifica la identidad $\nu(A)=\displaystyle\int_A f(x)d\mu(x)$ para todo $A\in\mathscr{A}$. Dado que $f\in\mathscr{M}$, podemos definir una medida positiva sobre $\mathscr{A}$ por medio de la expresión $\nu_0(A)=\nu(A)-\displaystyle\int_A f(x)d\mu(x)$. Vamos a verificar que ν_0 es la medida nula procediendo por el absurdo. Supongamos pues que la medida ν_0 es positiva y que no es idénticamente nula, entonces como la medida μ es finita existe un número $\varepsilon>0$ tal que

$$\nu_0(X)>\varepsilon\mu(X). \tag{1.23}$$

Sea ahora (P,N) una descomposición de Hahn de la medida con signo $\nu_0-\varepsilon\mu$; observamos que para todo $A\in\mathscr{A}$ se tiene $\nu_0(A\cap P)\geq\varepsilon\mu(A\cap P)$ y por lo tanto se tiene

$$\nu(A)=\int_A fd\mu+\nu_0(A)\geq\int_A fd\mu+\nu_0(A\cap P)\geq\int_A fd\mu+\varepsilon\mu(A\cap P)=\int_A\big(f+\varepsilon\mathbb{1}_P\big)d\mu. \tag{1.24}$$

Nótese que $\mu(P)>0$, pues como se ha supuesto la continuidad absoluta de la medida ν con respecto a la medida μ se obtiene que si $\mu(P)=0$ entonces $\nu_0(P)=0$ y por lo tanto se tendría

$$\nu_0(X)-\varepsilon\mu(X)=(\nu_0-\varepsilon\mu)(X)=(\nu_0-\varepsilon\mu)(P\cup N)=(\nu_0-\varepsilon\mu)(N)\leq 0$$

lo que sería una contradicción con (1.23). A partir de estos cálculos, por la fórmula (1.24) se obtiene que $f+\varepsilon\mathbb{1}_P$ pertenece al conjunto $\mathscr{M}$ y verifica $\displaystyle\int_X(f+\varepsilon\mathbb{1}_P)d\mu>\int_X f(x)d\mu(x)$ lo cual es una contradicción con (1.22) y por lo tanto se deduce que ν_0 es idénticamente nula. Con esto hemos demostrado que para todo $A\in\mathscr{A}$ se tiene la relación $\nu(A)=\displaystyle\int_A f(x)d\mu(x)$.

Para la última etapa usamos el Corolario 3.2.9-1) del Volumen 1 que nos dice que la función integrable f solo puede tomar valores infinitos sobre conjuntos de μ-medida nula, de modo que es posible redefinirla para que solo tome valores finitos. Con esto hemos terminado la verificación del teorema en el caso en que ambas medidas μ y ν son finitas.

En el caso en que estas medidas sean σ-finitas procedemos de la siguiente manera. En efecto, bajo esta hipótesis adicional tenemos que el conjunto X puede escribirse como la unión de una sucesión $(B_n)_{n\in\mathbb{N}}$ de conjuntos disjuntos $\mathscr{A}$-medibles de medida finita para μ y ν. Entonces, para todo $n\in\mathbb{N}$, la primera parte de la demostración nos proporciona funciones $\mathscr{A}$-medibles $f_n: X\longrightarrow\mathbb{R}_+$ tales que $\nu(A)=\displaystyle\int_A f_n(x)d\mu(x)$ para todo subconjunto $\mathscr{A}$-medible de B_n. Si definimos la función $f: X\longrightarrow\mathbb{R}_+$

que coincide sobre cada conjunto B_n con las funciones f_n, obtenemos de esta manera la función deseada.

Pasemos ahora estudiar la unicidad de esta función f. Es decir que si $\psi : X \longrightarrow \mathbb{R}_+$ es otra función que verifica

$$\nu(A) = \int_A \psi(x) d\mu(x) \qquad \text{para todo } A \in \mathscr{A}$$

deseamos saber las relaciones existentes entre f y ψ. Dado que la medida μ es σ-finita, podemos aplicar la Proposición 1.2.1 para obtener que $f = \psi$ en μ-casi todas partes, lo que termina la demostración del teorema. ■

Antes de continuar, es necesario mostrar un ejemplo en donde, al no disponer de las buenas hipótesis, no se tiene la conclusión del teorema: sean pues el conjunto $X = \{0, 1\}$, la σ-álgebra $\mathscr{A} = \mathcal{P}(X)$ y la medida $\mu(\{0\}) = 1$, $\mu(\{1\}) = +\infty$. Vemos entonces sin mayor problema que esta medida no es σ-finita. Si consideramos la medida ν definida por $\nu(\{0\}) = \nu(\{1\}) = 1$ obtenemos que ν es finita y además se tiene que estas dos medidas son equivalentes pues el conjunto vacío es el único conjunto que es de medida nula para μ y para ν, es decir que se tiene $\nu \ll \mu$ y $\mu \ll \nu$.

Como se tiene que $\nu \ll \mu$, vemos por un lado que no existe una función f definida sobre X tal que se tenga $\nu(\{1\}) = \displaystyle\int_{\{1\}} f(x) d\mu(x)$ puesto que $\nu(\{1\}) = 1$ y que esta integral es igual a 0 si $f(1) = 0$ e igual a $+\infty$ si $f(1) \neq 0$. Dado que se tiene también que $\mu \ll \nu$, observamos por otro lado que $\mu = (\mathbb{1}_{\{0\}} + \infty \times \mathbb{1}_{\{1\}}) \times \nu$, de manera que no existe una función f sobre X tal que $\mu = (f\nu)$.

Este pequeño ejemplo muestra la necesidad de la hipótesis de σ-finitud de las medidas que intervienen en el teorema de Radon-Nikodym.

En toda la Sección 1.2 hemos trabajado hasta ahora con medidas positivas y es necesario generalizar la noción de continuidad absoluta y el teorema de Radon-Nikodym a medidas más generales.

Definición 1.2.3 (Continuidad absoluta para medidas con signo o medidas complejas) *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida positiva definida sobre $\mathscr{A}$. Si ν es una medida con signo o una medida compleja, diremos que ν es absolutamente continua con respecto a μ, lo que será notado $\nu \ll \mu$, si la medida de variación $|\nu|$ de ν es absolutamente continua con respecto a μ en el sentido de la Definición 1.2.2.*

A raíz de esta definición, no es difícil ver que una medida con signo ν es absolutamente continua con respecto a una medida μ si y solo si las medidas positivas ν^+ y ν^- son absolutamente continuas con respecto a μ. De la misma manera, una medida compleja ν es absolutamente continua con respecto a una medida μ si y solo si las medidas ν_1^+, ν_1^-, ν_2^+ y ν_2^- asociadas a su descomposición de Jordan son absolutamente continuas con respecto a μ.

Proposición 1.2.3 *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida positiva definida sobre μ. Una medida ν con signo o compleja definida sobre $\mathscr{A}$ es absolutamente continua con respecto a μ si y solo si cada conjunto $A \in \mathscr{A}$ que verifica $\mu(A) = 0$ también verifica $\nu(A) = 0$.*

Este resultado, dejado al lector como ejercicio, permite usar la definición anterior de continuidad absoluta de medidas sin pasar por la medida de variación.

Observación 1.8 Es importante recordar que $\nu(A) = 0$ no es equivalente a $|\nu|(A) = 0$. Ver la Definición 1.1.2 y la Observación 1.2 página 15. Nótese además que siempre se tiene $\nu(A) \leq |\nu|(A)$, de manera que toda medida con signo es absolutamente continua con respecto a su variación. En efecto, la mayoración $\nu(A) \leq |\nu|(A)$ implica por la Proposición 1.2.2 que $\nu \ll |\nu|$.

Con estos preliminares podemos enunciar una versión más general del teorema de Radon-Nikodym.

Teorema 1.2.2 (de Radon-Nikodym versión con signo y compleja) *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida positiva σ-finita definida sobre él. Sea además ν una medida σ-finita con signo o una medida compleja σ-finita definida sobre $\mathscr{A}$. Si ν es absolutamente continua con respecto a μ, entonces existe una función f que pertenece al espacio $\mathcal{L}(X, \mathscr{A}, \mu, \mathbb{R})$, o al espacio $\mathcal{L}(X, \mathscr{A}, \mu, \mathbb{C})$ en el caso complejo, que verifica $\nu(A) = \displaystyle\int_A f(x)d\mu(x)$ para todo $A \in \mathscr{A}$. Esta función es única en μ-casi todas partes.*

Demostración. Si ν es una medida compleja absolutamente continua con respecto a la medida positiva μ, entonces ésta medida ν puede descomponerse como $\nu = \nu_1^+ - \nu_1^- + i\nu_2^+ - i\nu_2^-$ en donde las medidas ν_1^+, ν_1^-, ν_2^+ y ν_2^- son medidas positivas absolutamente continuas con respecto a μ. Para cada una de estas medidas, el Teorema 1.2.1 proporciona funciones f_1^+, f_1^-, f_2^+ y f_2^- tales que se tiene las relaciones

$$\nu_1^\pm(A) = \int_A f_1^\pm(x)d\mu(x) \qquad \text{y} \qquad \nu_2^\pm(A) = \int_A f_2^\pm(x)d\mu(x)$$

para todo $A \in \mathscr{A}$. La función f buscada se construye entonces de la siguiente manera: $f = f_1^+ - f_1^- + if_2^+ - if_2^-$. El caso de una medida ν con signo es totalmente similar.

Para estudiar la unicidad, basta proceder como en el Teorema 1.2.1; en el caso en que la medida ν sea compleja basta separar las partes reales y imaginarias de la función f para obtener el resultado deseado. ■

Con este teorema, vemos bajo qué condiciones se puede relacionar dos medidas μ y ν (ya sean estas medidas positivas, medidas con signo o medidas complejas) por medio de la expresión

$$\nu(A) = \int_A f d\mu.$$

Insistamos que además de la noción de σ-finitud, es necesario disponer de la noción de absoluta continuidad para obtener este resultado.

Una vez que disponemos de esta representación por medio de una integral de una medida, podemos estar tentados en generalizar el teorema fundamental del cálculo que relaciona una función F definida de forma integral por $F = \displaystyle\int f dx$ con su derivada f. De esta manera tenemos la definición a continuación:

Definición 1.2.4 (Derivada de Radon-Nikodym) *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida positiva σ-finita definida sobre $\mathscr{A}$. Sea además ν una medida con signo o una medida compleja σ-finita sobre $\mathscr{A}$ y supongamos que ν es absolutamente continua con respecto μ. Definimos la* derivada de Radon-Nikodym de ν con respecto a μ, *que será notada $D_\mu(\nu)$ o también $\frac{d\nu}{d\mu}$, a la función $\mathscr{A}$-medible f definida sobre X que verifica*

$$\nu(A) = \int_A f(x)d\mu(x) \qquad\qquad (\forall A \in \mathscr{A}).$$

Dicho de otra manera tenemos $\frac{d\nu}{d\mu} = f$ y esta derivada es única en μ-casi todas partes.

Es muy importante observar que la derivada de Radon-Nikodym de una medida no es una medida, sino una función; de manera que, si bien se puede hacer una analogía con el teorema fundamental del cálculo y la noción de derivada Radon-Nikodym, esta analogía es por el momento puramente superficial y será necesario precisarla. Notemos además que la *existencia* de esta función está justamente dada por el teorema de Radon-Nikodym, en particular, esta definición de derivada está dada en un sentido "integral" y veremos más tarde de qué manera se puede obtener estas medidas por medio de "límites". Volveremos a esta noción de derivada de una medida en la Sección 1.3.2 donde se estudiará este concepto con más detalle.

Damos ahora algunas propiedades relacionadas con este concepto de derivada de medidas.

Proposición 1.2.4 *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida positiva σ-finita definida sobre $\mathscr{A}$. Sea además ν una medida con signo o una medida compleja σ-finita sobre $\mathscr{A}$ y supongamos que ν es absolutamente continua con respecto μ. Suponemos además que la derivada de Radon-Nikodym $\frac{d\nu}{d\mu}$ de la medida ν con respecto a la medida μ existe. Entonces:*

1) si la medida ν es positiva, la derivada de Radon-Nikodym $\frac{d\nu}{d\mu}$ es una función positiva en μ-casi todas partes.

2) si la medida ν es una medida con signo y si $\nu = \nu^+ - \nu^-$ es la descomposición de Jordan de esta medida, entonces las derivadas de Radon-Nikodym $\frac{d\nu^+}{d\mu}$ y $\frac{d\nu^-}{d\mu}$ de las medidas $\nu^\pm$ con respecto a la medida μ existen. Además, si (P, N) es una descomposición de Hahn del conjunto X con respecto a la medida μ se tiene

$$\frac{d\nu^+}{d\mu} = \mathbb{1}_P \frac{d\nu}{d\mu} \quad y \quad \frac{d\nu^-}{d\mu} = -\mathbb{1}_N \frac{d\nu}{d\mu} \quad \text{en } \mu\text{-casi todas partes.}$$

3) Si $\left(\frac{d\nu}{d\mu}\right)^+$ y $\left(\frac{d\nu}{d\mu}\right)^-$ denotan las partes positivas y negativas de la derivada de Radon-Nikodym $\frac{d\nu}{d\mu}$ entonces se tienen las relaciones siguientes en μ-casi todas partes:

$$\frac{d\nu^+}{d\mu} = \left(\frac{d\nu}{d\mu}\right)^+ \quad y \quad \frac{d\nu^-}{d\mu} = \left(\frac{d\nu}{d\mu}\right)^-.$$

Prueba.

1) Por definición tenemos para todo conjunto $A \in \mathscr{A}$ que $\displaystyle\int_A \frac{d\nu}{d\mu} d\mu = \nu(A)$, y como la medida μ es positiva se deduce sin problema la positividad en μ-casi todas partes de la derivada de Radon-Nikodym $\frac{d\nu}{d\mu}$.

2) Dado que, por el Corolario 1.1.1, si (P, N) es una descomposición de Hahn del conjunto X, entonces tenemos las siguientes identidades para todo $A \in \mathscr{A}$

$$\nu^+(A) = \nu(A \cap P) = \int_{A\cap P} \frac{d\nu}{d\mu} d\mu = \int_A \mathbb{1}_P \frac{d\nu}{d\mu} d\mu$$

$$\nu^-(A) = -\nu(A \cap N) = \int_{A\cap N} -\frac{d\nu}{d\mu} d\mu = \int_A -\mathbb{1}_N \frac{d\nu}{d\mu} d\mu,$$

se obtiene que las funciones $\mathbb{1}_P \frac{d\nu}{d\mu}$ y $-\mathbb{1}_N \frac{d\nu}{d\mu}$ son derivadas de Radon-Nikodym de las medidas ν^+ y ν^- respectivamente. La unicidad está dada en μ-casi todas partes por el Teorema de Radon-Nikodym 1.2.1.

3) Definimos los conjuntos $\mathcal{P} = \{x \in X : \frac{d\nu}{d\mu}(x) \geq 0\}$ y $\mathcal{N} = \{x \in X : \frac{d\nu}{d\mu}(x) < 0\}$ de manera que se tiene en μ-casi todas partes

$$\left(\frac{d\nu}{d\mu}\right)^+ = \mathbb{1}_{\mathcal{P}} \frac{d\nu}{d\mu} \quad \text{y} \quad \left(\frac{d\nu}{d\mu}\right)^- = -\mathbb{1}_{\mathcal{N}} \frac{d\nu}{d\mu}.$$

Se tiene que los conjuntos $\mathcal{P}$ y $\mathcal{N}$ son $\mathscr{A}$-medibles y verifican $\mathcal{P} \cap \mathcal{N} = \emptyset$ y $\mathcal{P} \cup \mathcal{N} = X$. Vamos a verificar que $(\mathcal{P}, \mathcal{N})$ es una descomposición de Hahn del conjunto X con respecto a la medida ν: para ello observamos que, por construcción del conjunto $\mathcal{P}$, se tiene que para todo conjunto $\mathscr{A}$-medible $A \subset \mathcal{P}$, la medida de A es positiva pues $\nu(A) = \int_A \frac{d\nu}{d\mu} d\mu \geq 0$ y esto muestra que el conjunto $\mathcal{P}$ es un conjunto positivo de X con respecto a la medida ν. Razonando de manera similar se obtiene que el conjunto $\mathcal{N}$ es un conjunto negativo de X con respecto a la medida ν y por lo tanto que $(\mathcal{P}, \mathcal{N})$ es una descomposición de Hahn y podemos escribir las relaciones

$$\frac{d\nu^+}{d\mu} = \mathbb{1}_{\mathcal{P}} \frac{d\nu}{d\mu} = \left(\frac{d\nu}{d\mu}\right)^+ \quad \text{y} \quad \frac{d\nu^-}{d\mu} = -\mathbb{1}_{\mathcal{N}} \frac{d\nu}{d\mu} = \left(\frac{d\nu}{d\mu}\right)^-,$$

que son válidas en μ-casi todas partes. ■

Proposición 1.2.5 *Sean μ una medida positiva σ-finita y ν una medida σ-finita con signo definidas sobre el espacio medible $(X, \mathscr{A})$. Si la derivada de Radon-Nikodym $\frac{d\nu}{d\mu}$ de la medida ν con respecto a la medida μ existe, entonces para toda función $\mathscr{A}$-medible $\varphi : X \longrightarrow \mathbb{K}$ se tiene la identidad*

$$\int_X \varphi d\nu = \int_X \varphi \frac{d\nu}{d\mu} d\mu$$

Prueba. Empecemos considerando el caso cuando la medida ν es positiva. Sea $\varphi = \mathbb{1}_A$ donde $A \in \mathscr{A}$ y entonces dado que la derivada $\frac{d\nu}{d\mu}$ existe, podemos escribir

$$\int_X \varphi d\nu = \int_X \mathbb{1}_A d\nu = \nu(A) = \int_A \frac{d\nu}{d\mu} d\mu = \int_X \mathbb{1}_A \frac{d\nu}{d\mu} d\mu = \int_X \varphi \frac{d\nu}{d\mu} d\mu$$

de manera que se obtiene el resultado deseado en este caso particular. El caso donde φ es una función medible positiva se deduce del Teorema de Convergencia Monótona, pues todas las cantidades que intervienen son positivas por el primer punto de la Proposición 1.2.4. Ahora si la función φ es cualquiera, podemos descomponerla como $\varphi = \varphi^+ - \varphi^-$ y se tiene por lo anterior que

$$\int_X \varphi^+ d\nu = \int_X \varphi^+ \frac{d\nu}{d\mu} d\mu \quad \text{y} \quad \int_X \varphi^- d\nu = \int_X \varphi^- \frac{d\nu}{d\mu} d\mu$$

de manera que si una de las diferencias $\int_X \varphi^+ d\nu - \int_X \varphi^- d\nu$ o $\int_X \varphi^+ \frac{d\nu}{d\mu} d\mu - \int_X \varphi^- \frac{d\nu}{d\mu} d\mu$ existe, la otra existe inmediatamente de manera que se obtiene la fórmula buscada en el caso de que la función φ sea a valores reales.

Pasemos ahora al caso cuando la medida ν es una medida σ-finita con signo. Consideremos $\nu = \nu^+ - \nu^-$ la descomposición de Jordan de la medida ν. Tenemos entonces que las medidas $\nu^\pm$ son σ-finitas y la existencia de la derivada $\frac{d\nu}{d\mu}$ implica por la Proposición 1.2.4 que $\left(\frac{d\nu}{d\mu}\right)^\pm$ son las derivadas de Radon-Nikodym de las medidas $\nu^\pm$ con respecto a la medida μ. Como las medidas $\nu^\pm$ son positivas, podemos aplicar el resultado apenas demostrado para obtener las relaciones

$$\int_X \varphi d\nu^\pm = \int_X \varphi \frac{d\nu^\pm}{d\mu} d\mu = \int_X \varphi \left(\frac{d\nu}{d\mu}\right)^\pm d\mu$$

válidas para toda función medible φ. A partir de esto, si una de las diferencias $\int_X \varphi d\nu^+ - \int_X \varphi d\nu^-$ o $\int_X \varphi \left(\frac{d\nu}{d\mu}\right)^+ d\mu - \int_X \varphi \left(\frac{d\nu}{d\mu}\right)^- d\mu$ existe, la otra también existe y son iguales; lo que termina la demostración. ■

Proposición 1.2.6 (Linealidad)

1) Sean μ una medida positiva σ-finita y ν una medida σ-finita con signo definidas sobre el espacio medible $(X, \mathscr{A})$. Si la derivada $\frac{d\nu}{d\mu}$ existe, entonces para todo $\alpha \in \mathbb{R}$, $\alpha \neq 0$, se tiene que $\frac{d(c\nu)}{d\mu}$ existe y se tiene la identidad

$$\frac{d(c\nu)}{d\mu} = c\frac{d\nu}{d\mu}$$

en μ-casi todas partes.

2) Si π es otra medida σ-finita con signo definida sobre el espacio medible $(X, \mathscr{A})$ tal que la medida $(\nu + \pi)$ está correctamente definida sobre el espacio $(X, \mathscr{A})$ y si las derivadas $\frac{d\nu}{d\mu}$ y $\frac{d\pi}{d\mu}$ existen entonces se tiene que la derivada $\frac{d(\nu+\pi)}{d\mu}$ existe y se tiene la relación

$$\frac{d(\nu + \pi)}{d\mu} = \frac{d\nu}{d\mu} + \frac{d\pi}{d\mu}$$

en μ-casi todas partes.

Prueba. El primer punto es directo pues si $\alpha \neq 0$ entonces podemos escribir

$$\int_A \alpha\frac{d\nu}{d\mu}d\mu = \alpha\int_A \frac{d\nu}{d\mu}d\mu = \alpha\nu(A) = (\alpha\nu)(A),$$

de manera que $\alpha\frac{d\nu}{d\mu}$ es una derivada de Radon-Nikodym de la medida $(\alpha\nu)$ con respecto a la medida μ, pero por la unicidad de tales funciones se obtiene la relación buscada.

Para el segundo punto, dado que las derivadas $\frac{d\nu}{d\mu}$ y $\frac{d\pi}{d\mu}$ existen, se tiene que la suma $\frac{d\nu}{d\mu} + \frac{d\pi}{d\mu}$ es una función que está bien definida y que es medible. Esto nos permite escribir

$$\int_A \left(\frac{d\nu}{d\mu} + \frac{d\pi}{d\mu}\right) d\mu = \int_A \frac{d\nu}{d\mu}d\mu + \int_A \frac{d\pi}{d\mu}d\mu = \nu(A) + \pi(A) = (\nu + \pi)(A)$$

de donde se deduce que $\frac{d\nu}{d\mu} + \frac{d\pi}{d\mu}$ es una derivada de Radon-Nikodym de la medida $(\nu+\pi)$ con respecto a la medida μ, dado que la unicidad está asegurada se obtiene la linealidad de la derivada. ■

Proposición 1.2.7 (Regla de la Cadena) *Sean μ y π dos medidas σ-finitas positivas y sea ν una medida σ-finita con signo, todas ellas definidas sobre el espacio medible $(X, \mathscr{A})$. Si las derivadas $\frac{d\nu}{d\mu}$ y $\frac{d\mu}{d\pi}$ exisiten entonces $\frac{d\nu}{d\pi}$ existe y se tiene la relación*

$$\frac{d\nu}{d\pi} = \frac{d\nu}{d\mu}\frac{d\mu}{d\pi}.$$

Prueba. Como las medidas μ y π son σ-finitas, se tiene que las funciones $\frac{d\nu}{d\mu}$ y $\frac{d\mu}{d\pi}$ están bien definidas y se puede considerar el producto de funciones $\frac{d\nu}{d\mu}\frac{d\mu}{d\pi}$ como una función medible. Sea ahora $A \in \mathscr{A}$, tenemos entonces aplicando la Proposición 1.2.5 a la función $\mathbb{1}_A\frac{d\nu}{d\mu}$:

$$\int_A \frac{d\nu}{d\mu}\frac{d\mu}{d\pi}d\pi = \int_X \mathbb{1}_A\frac{d\nu}{d\mu}\frac{d\mu}{d\pi}d\pi = \int_X \mathbb{1}_A\frac{d\nu}{d\mu}d\mu = \int_A \frac{d\nu}{d\mu}d\mu = \nu(A)$$

y esto muestra que $\frac{d\nu}{d\mu}\frac{d\mu}{d\pi}$ es la derivada de Radon-Nikodym de ν con respecto a la medida π. ■

Para terminar esta sección, damos un resultado que explica algunas propiedades de la derivada de Radon-Nikodym de una medida ν con respecto a su variación $|\nu|$.

Proposición 1.2.8 *Sea ν una medida finita con signo o una medida compleja definida sobre un espacio medible $(X, \mathscr{A})$. Entonces la derivada de Radon-Nikodym $\dfrac{d\nu}{d|\nu|}$ de ν con respecto a su variación $|\nu|$ verifica la identidad*

$$\left|\frac{d\nu}{d|\nu|}\right| = 1$$

en $|\nu|$-casi todas partes.

Prueba. Notamos para empezar que se tiene $\nu \ll |\nu|$ por la observación 1.8. Luego aplicamos la Proposición 1.1.11 con $f = \dfrac{d\nu}{d|\nu|}$ y $\mu = |\nu|$ para obtener la relación

$$|\nu|(A) = \int_A \left|\frac{d\nu}{d|\nu|}\right| d|\nu|$$

para todo $A \in \mathscr{A}$. Tenemos entonces que $\left|\frac{d\nu}{d|\nu|}\right|$ es una derivada de Radon-Nikodym de $|\nu|$ con respecto a $|\nu|$, pero como la función constante 1 también es una derivada de Radon-Nikodym, se deduce que $\left|\frac{d\nu}{d|\nu|}\right| = 1$ en μ-casi todas partes. ■

Como vemos, la derivada de medidas posee muchas propiedades similares a la noción usual de derivadas de funciones. Sin embargo hay que tener cuidado pues al no hacer los cálculos con calma pueden haber algunas sorpresas. Ver el Ejercicio 1.9 para más detalles al respecto.

Con estos resultados hemos terminado nuestra primera exposición sobre el teorema de Radon-Nikodym. Adelantándonos un poco, veremos cómo utilizar esta resultado en el estudio de la dualidad de los espacios de Lebesgue L^p y, dado que este punto será estudiado en detalle en el capítulo siguiente, deseamos mostrar en la sección a continuación otra aplicación de este teorema.

1.2.2. Esperanza Condicional

Hasta el momento hemos estudiado algunas relaciones existentes entre dos medidas definidas sobre un mismo y solo espacio medible $(X, \mathscr{A})$, es decir, que en toda la teoría desarrollada hasta ahora, tanto el conjunto X como la σ-álgebra $\mathscr{A}$ estaban fijados. Existen sin embargo situaciones en donde es necesario considerar familias de espacios medibles y vamos a presentar en este párrafo algunas propiedades que aparecen cuando se consideran diferentes σ-álgebras y diferentes medidas definidas sobre ellas. Es importante señalar que ésta forma de proceder es una de las diversas formas de iniciar la teoría de probabilidades, de manera que hay muchísimo material adicional que, lastimosamente, queda fuera del objetivo de estas pocas páginas.

Sea pues $(X, \mathscr{A})$ un espacio medible y consideremos $\mathscr{B}$ una σ-álgebra contenida en $\mathscr{A}$ (hablaremos entonces de una *sub-σ-álgebra*). Tenemos por la teoría desarrollada en el Volumen 1 los puntos a continuación:

- si μ es una medida sobre $\mathscr{A}$, entonces la restricción $\mu_{|\mathscr{B}}$ de μ a la σ-álgebra $\mathscr{B}$ es también una medida.
- si la medida μ es finita se tiene que la medida $\mu_{|\mathscr{B}}$ también es finita.
- si $f : X \longrightarrow \mathbb{K}$, con $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$, es una función $\mathscr{B}$-medible, entonces f también es una función $\mathscr{A}$-medible. En efecto, si la imagen inversa por f de todo conjunto boreliano de $\mathbb{K}$ pertenece a $\mathscr{B}$, entonces esta imagen inversa también pertenece a $\mathscr{A}$.

Podría pensarse que, al considerar sub-σ-álgebras, la teoría desarrollada hasta ahora no varía y se mantiene. Es interesante notar sin embargo que algunas propiedades no siempre se mantienen al pasar a sub-σ-álgebras. En particular si la medida μ es σ-finita, la medida $\mu_{|\mathscr{B}}$ no es necesariamente σ-finita: basta para ello considerar el espacio medible $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$: sabemos que la medida de Lebesgue es σ-finita sobre este espacio pero se tiene que la restricción de la medida de Lebesgue a la sub-σ-álgebra $\mathscr{B} = \{\emptyset, \mathbb{R}\}$ no es σ-finita.

Esto muestra que hay propiedades muy importantes que no se mantienen al pasar al sub-σ-álgebras y por lo tanto hay que tener un poco de cuidado cuando se trabaja sobre familias de σ-álgebras.

Pasamos ahora al resultado más importante de esta subsección en donde se hace variar las σ-álgebras.

Teorema 1.2.3 *Sea $(X, \mathscr{A}, \mu)$ un espacio medido con μ una medida finita. Sea $\mathscr{B} \subset \mathscr{A}$ una sub-σ-álgebra y sea f una función que pertenece al espacio $L^1(X, \mathscr{A}, \mu, \mathbb{R})$. Entonces:*

1) existe una función $h \in L^1(X, \mathscr{B}, \mu_{|\mathscr{B}}, \mathbb{R})$ tal que se tenga la fórmula

$$\int_X f(x)g(x)d\mu(x) = \int_X g(x)h(x)d\mu(x) \tag{1.25}$$

para toda función $\mathscr{B}$-medible $g : X \longrightarrow \mathbb{R}$ que pertenece al espacio $L^\infty(X, \mathscr{B}, \mu_{|\mathscr{B}}, \mathbb{R})$.

2) si h' es una función en $L^1(X, \mathscr{B}, \mu_{|\mathscr{B}}, \mathbb{R})$ tal que

$$\int_A f(x)d\mu(x) = \int_A h'(x)d\mu(x)$$

para todo conjunto $A \in \mathscr{B}$, entonces se tiene que $h = h'$ en μ-casi todas partes.

Antes de pasar a la demostración de este resultado, es importante explicar brevemente la expresión (1.25). El lector podría confundirse con la mezcla de σ-álgebras y podría cuestionar la validez de esta expresión. En efecto, en la integral de la parte izquierda de esta fórmula intervienen la función f que es $\mathscr{A}$-medible, la función g que es $\mathscr{B}$-medible y la medida μ que está definida sobre $\mathscr{A}$; mientras que en la integral de la derecha, las funciones g y h son $\mathscr{B}$-medibles y son integradas con respecto a μ. Todo se simplifica al recordar que una función $\mathscr{B}$-medible también es $\mathscr{A}$-medible si $\mathscr{B} \subset \mathscr{A}$, de manera que todo tiene sentido en esta fórmula.

Demostración. Verificamos cada uno de estos puntos por separado.

1) Observamos para empezar que si f es una función positiva μ-integrable y si μ es una medida finita, entonces el producto $(f\mu)$ de la función f por la medida μ, también es una medida finita sobre $\mathscr{A}$ de manera que las restricciones $\mu_{|\mathscr{B}}$ y $(f\mu)_{|\mathscr{B}}$ son igualmente medidas finitas sobre $\mathscr{B}$. Se tiene además que $(f\mu)_{|\mathscr{B}} \ll \mu_{|\mathscr{B}}$ puesto que si $A \in \mathscr{B}$ y $\mu(A) = 0$, entonces se tiene $f\mathbb{1}_A = 0$ en μ-casi todas partes de modo que $(f\mu)(A) = \displaystyle\int_A f(x)d\mu(x) = 0$.

   En este punto invocamos el teorema de Radon-Nikodym y lo aplicamos a las medidas finitas $(f\mu)_{|\mathscr{B}}$ y $\mu_{|\mathscr{B}}$ para obtener una función $\mathscr{B}$-medible h definida sobre X a valores en $\mathbb{R}_+$ tal que se tenga la identidad $(f\mu)_{|\mathscr{B}} = (h\mu_{|\mathscr{B}})$. Entonces, al considerar la integral de una función g que es $\mathscr{B}$-medible, que está definida sobre el conjunto X y que es $(f\mu)$-integrable obtenemos

$$\begin{aligned}\int_X g(x)f(x)d\mu(x) &= \int_X g(x)d(f\mu)(x) = \int_X g(x)d(f\mu)_{|\mathscr{B}}(x) = \int_X g(x)d(h\mu_{|\mathscr{B}})(x) \\ &= \int_X g(x)h(x)d\mu_{|\mathscr{B}}(x) = \int_X g(x)h(x)d\mu(x),\end{aligned}$$

de donde se deduce el primer punto en este caso. Para tratar el caso cuando la función f puede tomar valores negativos, la descomponemos en $f = f^+ - f^-$ y el razonamiento anterior nos proporciona dos funciones postivas h_1 y h_2 tales que

$$\int_X f^+(x)g(x)d\mu(x) = \int_X g(x)h_1(x)d\mu(x) \qquad \text{y} \qquad \int_X f^-(x)g(x)d\mu(x) = \int_X g(x)h_2(x)d\mu(x)$$

para toda función positiva g $\mathscr{B}$-medible. Dado que las funciones f^+ y f^- son μ-integrables, se obtiene que las funciones h_1 y h_2 son también μ-integrables: basta para ello considerar $g = \mathbb{1}_X$. De esta manera se tiene que $h = h_1 - h_2$ pertenece a $L^1(X, \mathscr{B}, \mu_{|_{\mathscr{B}}})$, de manera que por la linealidad de la integral se obtiene

$$\int_X f(x)g(x)d\mu(x) = \int_X g(x)h(x)d\mu(x)$$

para toda función $g \in L^\infty(X, \mathscr{B}, \mu_{|_{\mathscr{B}}}, \mathbb{R})$.

2) Dado que la expresión $\displaystyle\int_A f(x)d\mu(x) = \int_A h'(x)d\mu(x)$ es válida para todo conjunto $A \in \mathscr{B}$, entonces se tiene $(f\mu)_{|_{\mathscr{B}}} = (h'\mu)_{|_{\mathscr{B}}}$. Tomando en la fórmula (1.25) la función g igual a la función indicatriz de conjuntos de la σ-álgebra $\mathscr{B}$ se obtiene que $(f\mu)_{|_{\mathscr{B}}} = (h\mu)_{|_{\mathscr{B}}}$ y por lo tanto tenemos que $(h\mu_{|_{\mathscr{B}}}) = (h'\mu_{|_{\mathscr{B}}})$ de donde se deduce que $h = h'$ en $\mu_{|_{\mathscr{B}}}$-casi todas partes por la Proposición 1.2.1. ■

Definición 1.2.5 (Esperanza Condicional) *A la función h que interviene en la fórmula (1.25) se la denomina la* esperanza condicional *de f con respecto a $\mathscr{B}$ y se la denota por $\mathbb{E}[f|\mathscr{B}]$.*

Lo que se ha hecho aquí con una sub-σ-álgebra puede generalizarse a una familia de sub-σ-álgebras, todas ellas contenidas en una σ-álgebra $\mathscr{A}$ y, como se ha dicho anteriormente, éste puede considerarse como el punto de partida de la teoría de probabilidades, o en todo caso, el punto en donde el análisis y las probabilidades se separan para formar ramas distintas de las matemáticas. A falta de continuar en esta vía, nos limitamos a dar un ejemplo muy sencillo de cálculo de esperanza condicional.

Sea $(X, \mathscr{A}, \mu)$ un espacio probabilizado, sea $f \in L^1(X, \mathscr{A}, \mu, \mathbb{R})$ y sea $\mathscr{B} \subset \mathscr{A}$ una sub-σ-álgebra. Si la función f es $\mathscr{B}$-medible entonces se tiene la identidad $\mathbb{E}[f|\mathscr{B}] = f$. En efecto, si f es $\mathscr{B}$-medible y consideramos la identidad (1.25), la única función h que es $\mathscr{B}$-medible y que verifica esta expresión es justamente f, de donde se obtiene la relación $\mathbb{E}[f|\mathscr{B}] = f$. Así mismo, un razonamiento similar nos muestra que $\mathbb{E}[f|\mathscr{A}] = f$. Para mayor detalles, el lector puede consultar [?].

1.2.3. Singularidad y Teorema de descomposición de Lebesgue

Pasamos ahora a otros conceptos que permiten dar más informaciones sobre una medida. Hemos visto el tipo de propiedades que se pueden obtener cuando una medida ν es absolutamente continua con respecto a otra medida μ. Nos proponemos ahora estudiar una descomposicón de medidas ν de la forma $\nu = \nu_a + \nu_s$ en donde ν_a es una medida absolutamente continua y ν_s es una medida *singular* en un sentido que será definido a continuación.

Definición 1.2.6 (Medidas Concentradas sobre un conjunto) *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida positiva definida sobre él. Diremos que la medida μ está* concentrada sobre el conjunto $\mathscr{A}$-medible E *si $\mu(E^c) = 0$.*

En el caso de medidas con signo o de medidas compleja μ, diremos que tales medidas están concentradas sobre el conjunto $\mathscr{A}$-medible E si la variación $|\mu|$ de μ está concentrada sobre E; es decir si todo subconjunto $\mathscr{A}$-medible A de E^c verifica $\mu(A) = 0$.

Quizás el ejemplo más sencillo de medida concentrada sobre un conjunto está dado en el espacio medible $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ por la masa de Dirac δ_a en un punto $a \in \mathbb{R}$: sin niguna dificultad se observa que esta medida está concentrada sobre el punto a. En este mismo marco, vemos que para un punto cualquiera $x \in \mathbb{R}$, la medida de Lebesgue está concentrada en el conjunto $\mathbb{R} \setminus \{x\}$. Esto muestra en particular que no hay necesariamente unicidad en el conjunto sobre el cual está concentrada una medida.

Definición 1.2.7 (Medidas mutuamente singulares) *Si μ y ν son dos medidas positivas, con signo o complejas, definidas sobre un mismo espacio medible $(X, \mathscr{A})$, diremos que μ y ν son* mutuamente singulares *si existe un conjunto $\mathscr{A}$-medible E tal que μ está concentrada sobre E y ν está concentrada sobre E^c. Notaremos esta relación por $\mu \perp \nu$.*

Por abuso de lenguaje, y en caso de que no exista confusión, también diremos que una medida ν es *singular con respecto a* μ o más simplemente que ν y μ son *singulares*. Vemos en particular que si $\mu \perp \nu$ entonces $\nu \perp \mu$.

Demos unos ejemplos de medidas mutuamente singulares. Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida con signo definida sobre $\mathscr{A}$. Entonces las partes positivas μ^+ y negativas μ^- de μ son mutuamente singulares. En efecto, por definición estas dos medidas están concentradas sobre los conjuntos positivos P y negativos N que aparecen en la descomposición de Hahn de la medida μ. Dado que estos conjuntos son disjuntos, se obtiene que las medidas μ^+ y μ^- son mutuamente singulares.

Un segundo ejemplo está dado en $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ al considerar la medida de Lebesgue dx y una medida discreta μ definida sobre un conjunto finito. Estas dos medidas son singulares puesto que la medida de Lebesgue no carga los puntos.

Las nociones de medidas singulares y medidas absolutamente continuas son en cierto sentido complementarias como nos lo indica el siguiente resultado.

Lema 1.2.1 *Sea μ una medida positiva definida sobre un espacio medible $(X, \mathscr{A})$ y sea ν una medida positiva, finita con signo o compleja. Si $\nu \ll \mu$ y $\nu \perp \mu$, entonces se tiene que $\nu = 0$.*

Prueba. Dado que $\nu \perp \mu$, si μ está concentrada en un conjunto $E \in \mathscr{A}$, es decir si $\mu(E^c) = 0$, entonces ν está concentrada sobre E^c y se tiene $\nu(E) = 0$. Pero como $\nu \ll \mu$, se tiene que $\mu(E^c) = 0$ implica $\nu(E^c) = 0$, de donde se deduce que $\nu = 0$. ■

Tenemos pues que una medida ν no puede ser a la vez absolutamente continua con respecto a una medida μ y ser singular con respecto a esta misma medida μ y esto muestra que estas nociones reflejan informaciones muy distintas.

Sin embargo, se tiene que una medida ν puede, bajo ciertas condiciones, descomponerse como una suma $\nu = \nu_a + \nu_s$ en donde ν_a y ν_s son medidas absolutamente continuas y singulares con respecto a una medida μ como nos lo indica el siguiente teorema:

Teorema 1.2.4 (de descomposición de Lebesgue) *Sea μ una medida positiva definida sobre un espacio medible $(X, \mathscr{A})$ y sea ν una medida positiva σ-finita, una medida con signo finita o una medida compleja definida sobre $\mathscr{A}$. Entonces existen dos únicas medidas positivas, con signo o complejas ν_a y ν_s definidas sobre $\mathscr{A}$ tales que*

1) ν_a es absolutamente continua con respecto a μ, es decir que $\nu_a \ll \mu$,

2) ν_s es singular con respecto a μ: se tiene $\nu_s \perp \mu$,

3) se tiene la relación $\nu = \nu_a + \nu_s$.

La descomposición $\nu = \nu_a + \nu_s$ *se denomina la* descomposición de Lebesgue *de la medida* ν *con respecto a la medida* μ. *Las medidas* ν_a *y* ν_s *se denominan las partes* absolutamente continuas *y las* partes singulares *de* ν *respectivamente.*

Demostración Empezamos con el caso cuando la medida ν es positiva y finita. Definimos entonces el conjunto

$$\mathcal{N}_\mu = \{B \in \mathscr{A} : \mu(B) = 0\}$$

y consideramos una sucesión $(B_j)_{j\in\mathbb{N}}$ de conjuntos de $\mathcal{N}_\mu$ tal que

$$\lim_{j\to+\infty} \nu(B_j) = \sup\{\nu(B) : B \in \mathcal{N}_\mu\}.$$

Definimos ahora el conjunto $N = \bigcup_{j\in\mathbb{N}} B_j$ y definimos las medidas ν_a y ν_s por medio de las expresiones

$$\nu_a(A) = \nu(A \cap N^c) \qquad \text{y} \qquad \nu_s(A) = \nu(A \cap N).$$

Vemos entonces que se tiene por construcción que $\nu = \nu_a + \nu_s$. Ahora, la σ-aditividad de la medida μ muestra que $\mu(N) = 0$ y por lo tanto se tiene que ν_s es una medida singular con respecto a la medida μ. Observamos para continuar que dado que se tiene que

$$\nu(N) = \sup\{\nu(B) : B \in \mathcal{N}_\mu\},$$

entonces cada subconjunto $\mathscr{A}$-medible B de $\mathcal{N}^c$ que verifica $\mu(B) = 0$ también verifica $\nu(B) = 0$. En efecto, si éste no fuera el caso, se tendría que $N \cup B$ pertenecería a $\mathcal{N}_\mu$ y verificaría $\nu(N \cup B) > \nu(N)$. A partir de esto se obtiene la continuidad absoluta de la medida ν_a con respecto a la medida μ.

En el caso en que la medida ν es finita con signo o es una medida compleja podemos aplicar la construcción anterior a la medida finita $|\nu|$. De esta manera obtenemos un conjunto de μ-medida nula N tal que la descomposición de Lebesgue de $|\nu|$ está dada por $|\nu|_a(A) = |\nu|(A \cap N^c)$ y $|\nu|_s(A) = |\nu|(A \cap N)$. A partir de esta descomposición, no es difícil verificar que las medidas con signo o complejas ν_a y ν_s determinadas por $\nu_a(A) = \nu(A \cap N^c)$ y $\nu_s(A) = \nu(A \cap N)$ son una descomposición de Lebesgue de ν.

Pasemos ahora a demostrar que esta descomposición es única. Sean pues $\nu = \nu_a + \nu_s$ y $\nu = \overline{\nu}_a + \overline{\nu}_s$ dos descomposiciones de Lebesgue de la medida ν. Supongamos primero que la medida ν es finita con signo, compleja o positiva finita; en este caso se tendría la relación

$$\nu_a - \overline{\nu}_a = \overline{\nu}_s - \nu_s,$$

ahora dado que se tiene que $(\nu_a - \overline{\nu}_a) \ll \mu$ y $(\overline{\nu}_s - \nu_s) \perp \mu$ se deduce, por el Lema 1.2.1, que $\nu_a - \overline{\nu}_a = \overline{\nu}_s - \nu_s = 0$, de donde se obtiene que $\nu_a = \overline{\nu}_a$ y que $\overline{\nu}_s = \nu_s$, es decir que tenemos la unicidad de la descomposición en el caso de medidas finitas.

En el caso en que la medida ν es σ-finita fijamos una partición $(B_n)_{n\in\mathbb{N}}$ de subconjuntos $\mathscr{A}$-medibles de X que son de medida finita con respecto a ν. Basta entonces aplicar el razonamiento anterior a las restricciones de las medidas $\nu_a, \overline{\nu}_a, \nu_s$ y $\overline{\nu}_s$ sobre los subconjuntos $\mathscr{A}$-medibles de los conjuntos B_n y de esta manera se obtiene la unicidad de la descomposición de Lebesgue. ■

Mostremos ahora un ejemplo en donde se muestra la importancia de la hipótesis de σ-finitud de la medida ν. Consideremos el espacio medible $([0,1], \mathcal{B}or([0,1]))$. En el rol de μ, fijamos la medida de Lebesgue y en el rol de ν consideramos la medida de conteo, que en este contexto no es σ-finita. Vamos a ver que no existen dos medidas ν_a y ν_s tales que $\nu = \nu_a + \nu_s$, $\nu_a \ll \mu$ y $\nu_s \perp \mu$. En efecto, en caso de existir tales medidas, como $\nu_s \perp \mu$ existiría un conjunto $S \in \mathcal{B}or([0,1])$ tal que $\nu_s([0,1] \setminus S) = \mu(S) = 0$, de modo que $S \neq [0,1]$. Ahora, si x es un punto del conjunto $[0,1] \setminus S$, entonces $\nu_s(x) = 0$ y dado que $\nu_a(\{x\}) = 0$, pues $\mu(\{x\}) = 0$, se obtiene que $\nu(\{x\}) = \nu_a(\{x\}) + \nu_s(\{x\}) = 0$, pero esto es una contradicción con la definición de la medida de conteo ν, y se deduce que esta descomposición no puede existir en estas condiciones.

Como hemos visto en toda esta Sección 1.2, la noción de continuidad absoluta es el concepto fundamental que nos permite obtener una serie de resultados que nos proporcionan informaciones muy precisas sobre las medidas. Estos resultados serán puestos en acción en aplicaciones en los capítulos que siguen.

1.3. Aplicaciones diferenciables

Sabemos desde los primeros cursos de cálculo diferencial que existe una relación muy estrecha entre la integral de funciones y la derivada de funciones. Tenemos en particular el primer teorema fundamental del cálculo que nos dice lo siguiente: si $f : [a, b] \longrightarrow \mathbb{R}$ es una función continua y si definimos $F = \displaystyle\int f dx$, entonces tenemos la relación $\frac{d}{dx}F = F' = f$. Esta identidad ha sido demostrada en el Volumen 1 en la Sección 3.5.2 en donde hemos discutido rápidamente el hecho que en general, el proceso de primitivación no corresponde necesariamente a la operación inversa de la derivación. El objetivo de esta sección es estudiar estas relaciones basándonos en las herramientas desarrolladas hasta ahora.

Evidentemente, al hablar de *derivadas* será necesario disponer de estructuras suplementarias de manera que la mayoría de resultados que expondremos a continuación serán planteados en el marco de $\mathbb{R}^n$. Empezaremos considerando los importantes resultados relativos a los cambios de variable, en donde interviene el concepto de matriz Jacobiana, para luego concentrarnos en la derivación de funciones y de medidas.

1.3.1. Imagen de medidas y cambio de variables

Muchos problemas del cálculo diferencial pueden resolverse de manera mucho más simple al realizar un cambio de variable conveniente y es por lo tanto necesario demostrar estos precedimientos usuales en el marco de la integral de Lebesgue. Daremos primero unos resultados de orden general para luego pasar al caso del espacio euclídeo $\mathbb{R}^n$ y terminaremos con algunas aplicaciones importantes.

A) Resultados generales

Hemos visto con la Proposición 2.2.2 del primer volumen que es posible transportar la estructura de σ-álgebra de un conjunto a otro por medio de una aplicación. Este procedimiento se aplica también a medidas y tenemos la siguiente definición:

Definición 1.3.1 (Imagen de una medida) *Sean $(X, \mathscr{A})$ y $(Y, \mathscr{B})$ dos espacios medibles, consideremos $\theta : X \longrightarrow Y$ una función $(\mathscr{A}, \mathscr{B})$-medible y sea μ una medida positiva definida sobre el espacio medible $(X, \mathscr{A})$. Definimos una nueva medida ν definida sobre $(Y, \mathscr{B})$, llamada la* imagen de la medida μ con respecto a la aplicación θ*, por medio de la expresión*

$$\nu(B) = \mu(\theta^{-1}(B)) \quad \textit{para todo } B \in \mathscr{B}. \tag{1.26}$$

Utilizando las propiedades de la imagen inversa de las aplicaciones y la σ-aditividad de la medida μ se obtiene sin problema que la aplicación de conjuntos $\theta(\mu)$ es efectivamente una medida sobre el espacio medible $(Y, \mathscr{B})$. Vemos además que las medidas μ y ν tienen la misma masa total, es decir que se tiene $\nu(Y) = \mu(X)$. Esta propiedad es la base del cambio de variable como tendremos la oportunidad de verlo.

Demos un ejemplo muy sencillo. Sea $\theta : x \longmapsto 2x^2$ una función definida de $[0,1]$ sobre el conjunto $[0,2]$, ambos conjuntos dotados de sus respectivas σ-álgebras borelianas. Si μ es la medida de Lebesgue sobre el conjunto $[0,1]$, podemos definir una nueva medida ν sobre el conjunto $[0,2]$ usando la expresión (1.26). Así si consideramos el conjunto $[1,2]$, su medida ν está dada por $\nu([1,2]) = \mu(\theta^{-1}([1,2])) = \mu([\frac{1}{\sqrt{2}},1]) = 1 - 1/\sqrt{2}$.

Estudiemos ahora la integral de funciones con respecto a las medidas que se obtienen por medio de la definición anterior.

Proposición 1.3.1 *Sean $(X,\mathscr{A})$ y $(Y,\mathscr{B})$ dos espacios medibles, sea $\theta : X \longrightarrow Y$ una función $(\mathscr{A},\mathscr{B})$-medible y sea μ una medida positiva definida sobre el espacio medible $(X,\mathscr{A})$. Si $f : Y \longrightarrow \overline{\mathbb{R}}$ es una función positiva $\mathscr{B}$-medible, entonces f es $\mu(\theta^{-1})$-integrable si y solo si $f\circ\theta$ es μ-integrable y definimos la integral de f con respecto a la medida imagen de μ por medio de la aplicación θ por*

$$\int_Y f d\nu = \int_X f\circ\theta d\mu. \tag{1.27}$$

Prueba. Estudiemos primero el caso cuando la función f es positiva. Vemos sin problema que la función compuesta $f\circ\theta$ es $\mathscr{A}$-medible. Empezamos suponiendo que $f = \mathbb{1}_B$ es la función indicatriz de un conjunto $B \in \mathscr{B}$, observamos entonces que se dispone de la identidad siguiente $\theta(\mathbb{1}_B)(x) = \mathbb{1}_{\theta^{-1}(B)}(x)$. A partir de esta relación, utilizando la definición de la medida imagen ν, tenemos $\nu(B) = \mu(\theta^{-1}(B))$, es decir $\displaystyle\int_Y \mathbb{1}_B d\nu = \int_X \mathbb{1}_{\theta^{-1}(B)} d\mu = \int_X \mathbb{1}_B \circ \theta d\mu$, y de esta manera se obtiene el resultado deseado para funciones indicatrices de conjuntos $\mathscr{B}$-medibles. A partir de esta identidad pasamos sin dificultad a las funciones simples positivas $f = \displaystyle\sum_{j=1}^{n} \alpha_j \mathbb{1}_{B_j}$ y a las funciones medibles positivas arbitrarias por medio del teorema de convergencia monótona de Beppo Levi.

Para tratar el caso cuando f toma valores negativos, podemos descomponer esta función en sus partes positivas y negativas $f = f^+ - f^-$ y, utilizando la linealidad de la integral y aplicando por separado los argumentos anteriores, se obtiene el resultado deseado. ■

Lo interesante de la expresión (1.27) es que nos permite cambiar el espacio de definición de las medidas: la medida ν está definida sobre $(Y,\mathscr{B})$ y no sobre $(X,\mathscr{A})$, así mismo, la función f está definida sobre Y y no sobre X.

Este resultado se generaliza a funciones a valores en $\mathbb{C}$ puesto que este caso se obtiene separando las partes reales y las partes imaginaras.

Esto constituye una primera situación de un *cambio de variable* y es oportuno verlo por medio de unos pocos ejemplos concretos.

(i) Sea la función $\theta : \mathbb{R} \longrightarrow \mathbb{R}$ definida por $\theta(x) = -x$. Vemos entonces que, por definición de medida imagen, se tiene la identidad $\lambda(\theta^{-1})(A) = \lambda(A)$ para todo $A \in \mathcal{B}or(\mathbb{R})$ y donde λ es la medida de Lebesgue sobre $\mathbb{R}$. A partir de esto vemos que una función $f : \mathbb{R} \longrightarrow \mathbb{R}$ es Lebesgue-integrable si y solo si la función $x \longmapsto f(-x)$ es Lebesgue-integrable y se tiene entonces

$$\int_{\mathbb{R}} f(x)dx = \int_{\mathbb{R}} f(-x)dx.$$

(ii) Para un $\tau \in \mathbb{R}$ consideramos la función $\theta_\tau : \mathbb{R} \longrightarrow \mathbb{R}$ definida por $\theta_\tau(x) = x + \tau$. Dado que también se tiene la identidad $\lambda(\theta^{-1})(A) = \lambda(A)$ para todo $A \in \mathcal{B}or(\mathbb{R})$, vemos que una función medible f es Lebesgue-integrable si y solo si $f(x+\tau)$ es Lebesgue-integrable y se tiene

$$\int_{\mathbb{R}} f(x)dx = \int_{\mathbb{R}} f(x+\tau)dx.$$

Notemos que este resultado no es más que un caso particular de la Proposición 3.2.21 del primer volumen.

(iii) Sea ahora μ una medida positiva definida sobre el espacio medible $(X, \mathscr{A})$ y sea $\theta : X \longrightarrow \mathbb{R}$ una aplicación $\mathscr{A}$-medible. Consideramos la función $f(t) = |t|^p$ definida sobre $\mathbb{R}$ con $1 \leq p < +\infty$, si definimos una nueva medida por $\nu = \mu(\theta^{-1})$, se tiene para todo $1 \leq p < +\infty$ la fórmula

$$\int_X |\theta(x)|^p d\mu(x) = \int_{\mathbb{R}} |t|^p d\nu(t).$$

Esta expresión es una aplicación directa de la Proposición 1.3.1 con $f = |\cdot|^p$.

Antes de pasar al cambio de variables en el espacio euclideo, damos un último resultado de orden general.

Proposición 1.3.2 *Sean $(X, \mathscr{A})$ y $(Y, \mathscr{B})$ dos espacios medibles y sea $\theta : X \longrightarrow Y$ un isomorfismo de espacios medibles, es decir que θ es una aplicación $(\mathscr{A}, \mathscr{B})$-medible, biyectiva y se tiene que θ^{-1} es una aplicación $(\mathscr{B}, \mathscr{A})$-medible. Si $f : X \longrightarrow \overline{\mathbb{R}}_+$ es una función positiva $\mathscr{A}$-medible entonces se tiene la fórmula:*

$$\theta(f\mu) = (f \circ \theta^{-1})(\theta\mu).$$

Prueba. Sea pues un conjunto $B \in \mathscr{B}$ y consideremos el producto $\nu = (\theta\mu)$, tenemos entonces

$$(\theta(f\mu))(B) = \int_{\theta^{-1}(B)} f d\mu = \int_X (\mathbb{1}_B \circ \theta) d\mu = \int_X (\mathbb{1}_B \times f \circ \theta^{-1}) \circ \theta d\mu = \int_B f \circ \theta^{-1} d\nu.$$

■

Esta proposicion es interesante pues nos permitirá estudiar integrales asociadas a estas medidas por medio de los cambios de variable usuales.

B) Cambio de variables en $\mathbb{R}^n$

Pasamos ahora a estudiar los cambios de variable en el espacio euclídeo $\mathbb{R}^n$ de manera que en toda esta sección consideraremos su estructura natural de espacio medido, es decir que consideraremos el espacio $(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n), dx)$ en donde $\mathcal{B}or(\mathbb{R}^n)$ es la σ-álgebra de los borelianos asociados a la topología usual de $\mathbb{R}^n$ y en donde dx es la medida usual de $\mathbb{R}^n$, que será notada indistintamente por dx o por λ_n. La medida de un conjunto boreliano A será notada $\lambda_n(A)$ o también $|A|$.

Una vez que hemos fijado estas notaciones, para continuar nuestra exposición será necesario recordar algunos conceptos del álgebra lineal.

Sea $n \geq 1$ un entero, notaremos $\mathcal{M}_{n\times n}(\mathbb{R})$ el conjunto de todas las matrices reales de dimensión $n \times n$, notaremos la matriz identidad por $I_{n\times n}$ o simplemente I cuando la dimensión esté clara y notaremos las columnas de una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ por $A_1, A_2, ..., A_n$.

Sea $D : \mathcal{M}_{n\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ una función definida sobre el conjunto de todas las matrices reales. Escribiremos $D(A_1, ..., A_n) = D(A)$ cuando sea necesario insistir en la relación de la aplicación D con respecto a las columnas de la matriz A. Recordamos que una aplicación $D : \mathcal{M}_{n\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ es *multilineal* si las aplicaciones $A_i \longmapsto D(A_1, ..., A_n)$ son lineales para cada i y para cada A_j con $1 \leq i \neq j \leq n$. Diremos además que esta aplicación D es *alternada* si $D(A) = 0$ cada vez que dos columnas de A son iguales. Finalmente una aplicación D es un *determinante* si es multilineal, alternada y si $D(I_{n\times n}) = 1$. Recordemos finalmente que existe un único determinante sobre $\mathcal{M}_{n\times n}(\mathbb{R})$ que será notado det.

Algunas propiedades bien conocidas de los determimantes se encuentran en el siguiente lema.

Lema 1.3.1 *Sea n un entero positivo y sea $\mathcal{M}_{n\times n}(\mathbb{R})$ el conjunto de todas las matrices reales. Entonces*

1) si $det(A) \neq 0$ la matriz es inversible,

2) para todo $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$ se tiene $det(A \times B) = det(A)det(B)$, en particular, si A es una matriz inversible se tiene la relación

$$det(A^{-1}) = \frac{1}{det(A)}. \tag{1.28}$$

3) $det(A)$ es un polinomio en la componentes de A.

La demostración de este resultado puede encontrarse en cualquier libro de álgebra lineal.

Si $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es una aplicación lineal y si A es la matriz de T con respecto a una base de $\mathbb{R}^n$ y si B es la matriz de T con respecto a otra base de $\mathbb{R}^n$, entonces existe una matriz inversible U tal que $A = UBU^{-1}$. Se deduce de esto que

$$det(A) = det(U)det(B)det(U^{-1}) = det(B)$$

y por lo tanto podemos definir el determinante de la aplicación T como $det(T)$ de forma independiente de la base utilizada para caracterizar la aplicación T.

Ya hemos visto en la Sección 2.4.3 del Volumen 1 algunas propiedades de la medida de Lebesgue con respecto a dos tipos muy particulares de aplicaciones lineales que son la traslación $T_\tau(A) = \tau + A$ y la dilatación $\delta_\alpha(A) = \alpha A$ para todo $A \in \mathcal{B}or(\mathbb{R}^n)$. El resultado a continuación muestra cómo se modifica la medida de Lebesgue de conjuntos borelianos bajo la acción de aplicaciones lineales generales.

Proposición 1.3.3 *Consideremos el espacio medido $(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n), \lambda_n)$ y sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una aplicación lineal inversible. Entonces para todo conjunto boreliano $A \in \mathcal{B}or(\mathbb{R}^n)$ se tiene la fórmula*

$$\lambda_n\big(T(A)\big) = |det(T)|\lambda_n(A). \tag{1.29}$$

Prueba. Empecemos observando que las aplicaciones T y T^{-1} son continuas y por lo tanto medibles con respecto a la σ-álgebra $\mathcal{B}or(\mathbb{R}^n)$, de manera que el conjunto $T(A)$ es boreliano si y solo si A es un boreliano. Luego vemos que la aplicación lineal T tiene una matriz asociada (en una cierta base) y entonces la acción de T sobre un vector $x \in \mathbb{R}^n$ se transcribe por medio de operaciones matriciales. Ahora, puesto que T es inversible, existen aplicaciones lineales $T_1, ..., T_n$ tales que $T = T_1 \circ \cdots \circ T_n$ y tales que cada T_k con $1 \leq k \leq n$ opera sobre un vector x en una de las formas siguientes:

i) una componente de x es multiplicada por un factor no nulo y las otras no son modificadas,

ii) dos componentes de x son intercambiadas y las otras no son modificadas,

iii) para algún i y j la componente x_i es reemplazada por $x_i + x_j$ y las otras no son modificadas.

El lector observará que estas tres operaciones $i) - iii)$ corresponden a las operaciones matriciales usuales que se enseñan en los primeros cursos de cálculo matricial cuando se estudia el problema de invertir una matriz.

Para continuar, dado que tenemos la relación $det(T) = det(T_1) \cdots det(T_n)$, sólo debemos mostrar que se tiene

$$\lambda_n(T_k(A)) = |det(T_k)|\lambda_n(A) \tag{1.30}$$

para todo $1 \leq k \leq n$ y todo conjunto boreliano A. Esto nos permite concentrarnos únicamente en los casos $i) - iii)$ mencionados anteriormente.

Supongamos primero que T_k es uno de los dos primeros casos $i)$ o $ii)$. Entonces la fórmula (1.30) es fácil de verificar para un cubo A cuyos ejes son paralelos a los ejes de coordenadas. En efecto, en el caso en que una componente es multiplicada por un factor $\alpha \neq 0$ entonces tenemos que $\lambda_n(T_k(A)) = |\alpha|\lambda_n(A)$ y no es muy difícil ver que en este caso se tiene que $|det(T_k)| = |\alpha|$. El caso $ii)$ es totalmente similar, solo que aquí $|det(T_k)| = 1$. Una vez que hemos obtenido estos resultados en el caso de un cubo, por la regularidad de la medida de Lebesgue, podemos extender este resultado a los conjuntos abiertos o a los conjuntos borelianos cualquiera.

Supongamos ahora que T_k verifica el caso $iii)$. Existe por lo tanto dos índices i y j tales que si $x = (x_1, ..., x_n)$ entonces la i-ésima coordenada de $T_k(x)$ es $x_i + x_j$ mientras que las otras coordenadas permanecen iguales. Veamos entonces $\mathbb{R}^n$ como el producto de $\mathbb{R}$ (que corresponde a la i-ésima coordenada) con $\mathbb{R}^{n-1}$ y sea A un subconjunto boreliano de $\mathbb{R}^n$. No es difícil verificar que para cada $y \in \mathbb{R}^{n-1}$ las secciones en y de A y de $T_k(A)$ son traslaciones la una de la otra y por lo tanto tienen la misma medida de Lebesgue. Se deduce entonces por la teoría de las medidas producto que $\lambda_n(A) = \lambda_n(T_k(A))$. Dado que en este caso se tiene $det(T_k) = 1$, hemos terminado la demostración. ■

Veamos un ejemplo simple. Consideremos la aplicación lineal $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por la expresión $T(x_1, x_2) = (2x_1 + x_2, x_2)$, se tiene que la matriz asociada a la aplicación T es

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}.$$

Vemos sin problema que la imagen del cubo $C = [0, 2] \times [0, 2]$ bajo la acción de la aplicación T es el paralelogramo de vértices $((0, 0); (2, 2); (6, 2); (4, 0))$. A partir de esto se puede verificar sencillamente la fórmula (1.29) y se tiene $\lambda_2(T(C)) = |det(T)| \times \lambda_2(C) = 2 \times 4 = 8$.

Si bien las aplicaciones lineales constituyen una familia fundamental de aplicaciones, es necesario considerar otro tipo de aplicaciones y, para continuar nuestra exposición, es necesario presentar algunos hechos acerca de las derivadas de funciones vectoriales. Recordaremos pues unas definiciones y unos resultados bien conocidos cuyas demostraciones pueden encontrarse en cualquier libro de cálculo vectorial.

Definición 1.3.2 (Función diferenciable) *Sean $(E, \|\cdot\|_E)$ y $(F, \|\cdot\|_F)$ dos espacios de Banach. Sean U un abierto de E y $x_0 \in U$. Una función $\varphi : U \longrightarrow F$ es* diferenciable *en x_0 si existe una aplicación lineal continua $D_\varphi : E \longrightarrow F$ tal que*

$$\lim_{x \to x_0} \frac{\|\varphi(x) - \varphi(x_0) - D_\varphi(x - x_0)\|_F}{\|x - x_0\|_E} = 0. \tag{1.31}$$

La aplicación D_φ es la derivada de φ en el punto x_0 y la notamos $\varphi'(x_0)$. Notaremos además $\mathcal{C}^1(U,F)$ el conjunto de funciones diferenciables para todo $x \in U \subset E$ a valores en F.

En este contexto, la regla de la cadena tiene la siguiente forma

Proposición 1.3.4 (Regla de la Cadena) *Sean $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y), (Z, \|\cdot\|_Z)$ tres espacios de Banach y sean U y V dos subconjuntos abiertos de X y Y. Si $x_0 \in U$, si $\varphi : U \longrightarrow Y$ es una aplicación diferenciable en x_0 que verifica $\varphi(U) \subset V$ y si $\psi : Y \longrightarrow Z$ es una aplicación diferenciable en $\varphi(x_0)$, entonces la aplicación $\psi \circ \varphi$ es diferenciable en x_0 y se tiene además*

$$(\psi \circ \varphi)'(x_0) = \psi'(\varphi(x_0))\varphi'(x_0).$$

Volvamos ahora al espacio $\mathbb{R}^n$ que dotaremos de la norma $\|\cdot\|_\infty$ definida por $\|x\|_\infty = \max\limits_{1\leq i\leq n} |x_i|$. Vemos en particular que en el espacio $(\mathbb{R}^n, \|\cdot\|_\infty)$ si $A = (a_{ij})_{1\leq i,j\leq n}$ es la matriz de una aplicación lineal $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ con respecto a la base canónica de $\mathbb{R}^n$ entonces T es continua y su norma está dada por la expressión

$$\|T\| = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|. \tag{1.32}$$

Necesitaremos también los dos lemas a continuación.

Lema 1.3.2 *Sea U un subconjunto abierto de $\mathbb{R}^n$ y sea $\varphi : U \longrightarrow \mathbb{R}^n$ una función de clase $\mathcal{C}^1(U,\mathbb{R}^n)$. Entonces $\varphi = (\varphi_1, \cdots, \varphi_n)$ es diferenciable en cada punto de U y la matriz de φ' con respecto a la base canónica de $\mathbb{R}^n$ está dada por $\left(\frac{\partial \varphi_i}{\partial x_j}\right)_{1\leq i,j\leq n}$.*

Lema 1.3.3 (Teorema del Valor Medio) *Sea U un subconjunto abierto de $\mathbb{R}^n$ y sea $\varphi : U \longrightarrow \mathbb{R}^n$ una aplicación diferenciable en todo punto de U. Sean x_0 y x_1 dos puntos de U, definimos para todo $t \in [0,1]$ los puntos y del segmento que une x_0 con x_1 por la fórmula $y = tx_0 + (1-t)x_1$. Si todos estos puntos x_0, x_1 pertenecen a U, si todos los puntos del segmento que une x_0 con x_1 pertenecen a U y si se tiene $\|\varphi'(y)\|_\infty \leq C$ para todos los puntos de este segmento; entonces*

$$\|\varphi(x_1) - \varphi(x_0)\|_\infty \leq C\|x_1 - x_0\|_\infty.$$

Estos dos lemas son resultados clásicos del análisis vectorial.

Pasamos ahora a uno de los objetos más importantes cuando se trata de estudiar el cambio de variables en $\mathbb{R}^n$.

Definición 1.3.3 (Matriz Jacobiana, Jacobiano) *Sea U un subconjunto abierto de $\mathbb{R}^n$. Si la aplicación $\varphi : U \longrightarrow \mathbb{R}^n$ es una función de clase $\mathcal{C}^1(U,\mathbb{R}^n)$ de componentes $\varphi_1, ..., \varphi_n$, definimos entonces la* matriz Jacobiana *de φ por*

$$J_\varphi(x) = \begin{bmatrix} \frac{\partial \varphi_1}{\partial x_1} & \cdots & \frac{\partial \varphi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_n}{\partial x_1} & \cdots & \frac{\partial \varphi_n}{\partial x_n} \end{bmatrix}. \tag{1.33}$$

El Jacobiano[6] *de la función φ está determinado por el determinante de esta matriz y será notado $det(J_\varphi(x))$.*

[6]Carl Jacobi (1804-1851), matemático alemán.

Podemos entonces enunciar el teorema más relevante de esta sección.

Teorema 1.3.1 (de cambio de variables) *Sea $\mathbb{R}^n$ dotado de su estructura de espacio medido natural $(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n), \lambda_n)$. Sean U, V dos abiertos de $\mathbb{R}^n$ y sea φ una biyección de U sobre V tal que tanto φ como φ^{-1} sean de clase $\mathcal{C}^1(U, V)$. Entonces, para todo subconjunto boreliano A de U se tiene la fórmula*

$$\lambda_n\big(\varphi(A)\big) = \int_A |det(J_\varphi)(x)|dx. \tag{1.34}$$

Si además $f : V \longrightarrow \mathbb{R}$ es una función boreliana medible, entonces

$$\int_V f(t)dt = \int_U f\big(\varphi(x)\big)|det(J_\varphi)(x)|dx. \tag{1.35}$$

Observemos que, bajo las hipótesis del teorema, dado que se tiene la identidad $\varphi^{-1}\big(\varphi(x)\big) = x$, la regla de la cadena implica que se tiene la identidad $\big(\varphi^{-1}\big)'\big(\varphi(x)\big)\varphi'(x) = I$ para todo $x \in U$ y por lo tanto la aplicación $\varphi'(x)$ es inversible. Esto implica que $det(J_\varphi)(x)$ no es nulo sobre U y muestra que estas dos expresiones anteriores tienen sentido.

Es interesante recalcar que la matriz Jacobiana mediante su determinante expresa en la fórmula (1.34) la deformación del volumen por medio de la función φ. Un ejemplo de ello se ha visto rápidamente en la página 43 al considerar una aplicación lineal.

Observación 1.9 En virtud de toda la teoría desarrollada sobre la integral, se tiene que las expresiones (1.34) y (1.35) son equivalentes.

En efecto, si f es la función indicatriz de un subconjunto boreliano B de V y si $A = \varphi^{-1}(B)$, entonces (1.34) se deduce de (1.35): por un lado se tiene que $\displaystyle\int_V \mathbb{1}_B(t)dt = \lambda_n(\varphi(A))$ y por otro lado se tiene que $\mathbb{1}_B(\varphi(x)) = \mathbb{1}_A(x)$. Recíprocamente, la linealidad de la integral y el teorema de convergencia monótona implican que se tiene (1.35) a partir de la expresión (1.34) para todas las funciones borelianas positivas y el caso de funciones borelianas generales se deduce de la descomposición $f = f^+ - f^-$.

Esto muestra que basta concentrarse en demostrar la fórmula (1.34) y para ello necesitaremos un par de lemas.

Lema 1.3.4 *Sea U un conjunto abierto de $\mathbb{R}^n$ y sea $\varphi : U \longrightarrow \mathbb{R}^n$ una aplicación diferenciable. Sea $\varepsilon > 0$ un real, K un cubo contenido en U cuyos ejes son paralelos a los ejes de coordenadas y tal que se tiene la estimación*

$$\|\varphi'(x) - I\|_\infty \leq \varepsilon \qquad \textit{para todo } x \in K.$$

Entonces $\varphi(K)$ la imagen de K por medio de la aplicación φ verifica

$$\lambda_n(\varphi(K)) \leq (1 + \varepsilon)^n \lambda_n(K).$$

Prueba. Sea x_0 el centro del cubo K y sea b la longitud de los lados de K. Se tiene entonces la siguiente estimación para todo $x \in K$: $\|x - x_0\|_\infty \leq b/2$ y entonces por hipótesis podemos aplicar el Lema 1.3.3 a la función $x \longmapsto \varphi(x) - x$ de manera que para todo $x \in K$ se tiene

$$\|(\varphi(x) - x) - (\varphi(x_0) - x_0)\|_\infty \leq \varepsilon \|x - x_0\|_\infty$$

es decir

$$\|\varphi(x) - \varphi(x_0)\|_\infty \leq (1 + \varepsilon)\|x - x_0\|_\infty \leq \frac{1}{2}(1 + \varepsilon)b.$$

Luego las imágenes bajo la aplicación φ de los puntos del cubo C pertenecen al cubo cerrado de ejes paralelos a los ejes de coordenadas cuyo centro está en $\varphi(x_0)$ y cuyos lados tienen una longitud $(1+\varepsilon)b$. Puesto que este cubo tiene una medida $(1+\varepsilon)^n b^n$, y dado que el cubo K tiene medida b^n, se obtiene el lema. ■

Lema 1.3.5 *Sean U, V y φ como en el Teorema 1.3.1. Si $C > 0$ es un número positivo y A es un subconjunto boreliano de U tal que $|J_\varphi(x)| \leq C$ para todo $x \in A$, entonces $\lambda_n(\varphi(A)) \leq C\lambda_n(A)$.*

Prueba. Suponemos primero que W es un subconjunto abierto de U tal que $\overline{W}$ es compacto y está incluido en U y se tiene $|J_\varphi(x)| < C$ para todo $x \in W$. Sea $\varepsilon > 0$ un real, puesto que $\overline{W}$ es un subconjunto compacto de U y dado que φ y φ^{-1} son de clase $\mathcal{C}^1$, entonces los valores de $(\varphi^{-1})'(\varphi(x))$ son acotados si $x \in W$. De igual manera las funciones que envían x a las componentes de $\varphi'(x)$ (es decir las derivadas parciales de φ) son uniformemente continuas sobre W. Entonces, podemos fijar un número positivo M tal que se tenga la estimación

$$\|(\varphi^{-1})'(\varphi(x))\|_\infty \leq M \tag{1.36}$$

para todo $x \in W$ y podemos fijar un número positivo $\delta > 0$ tal que la desigualdad

$$\|\varphi'(x) - \varphi'(x_0)\|_\infty \leq \frac{\varepsilon}{M} \tag{1.37}$$

es verificada siempre y cuando x y x_0 pertenecen a W y satisfacen $\|x - x_0\|_\infty \leq \delta$.

Sabemos que el conjunto W es la unión numerable de una familia $(K_i)_{i\in\mathbb{N}}$ de cubos semi-abiertos disjuntos de ejes paralelos a los ejes de coordenadas y, subdividiendo estos cubos, es posible suponer que cada uno de estos cubos tiene los lados de longitud 2δ. Sea K un tal cubo y sea x_0 su centro y definamos $g : U \longrightarrow \mathbb{R}^n$ de la siguiente forma

$$g(x) = [(\varphi^{-1})'(\varphi(x_0))]\varphi(x).$$

La regla de la cadena implica que para cada $x \in U$ se tiene

$$g'(x) - I = [(\varphi^{-1})'(\varphi(x_0))]\varphi'(x) - I = [(\varphi^{-1})'(\varphi(x_0))](\varphi'(x) - \varphi'(x_0)).$$

Así, con las desigualdades (1.36) y (1.37), obtenemos la estimación para todo $x \in K$:

$$\|g'(x) - I\|_\infty \leq \|(\varphi^{-1})'(\varphi(x_0))\|_\infty \|\varphi'(x) - \varphi'(x_0)\|_\infty \leq M\frac{\varepsilon}{M} = \varepsilon.$$

Ahora, aplicamos el Lema 1.3.4 y tenemos que $\lambda_n(g(K)) \leq (1+\varepsilon)^n \lambda_n(K)$.

Pero por otro lado, dado que $\varphi(K) = \varphi'(x_0)(g(K))$, la Proposición 1.3.3 implica que se tiene $\lambda_n(\varphi(K)) = |det(\varphi'(x_0))|\lambda_n(g(K))$. Por lo tanto se tienen las estimaciones

$$\lambda_n(\varphi(K)) = |det(\varphi'(x_0))|\lambda_n(g(K)) \leq C\lambda_n(g(K)) \leq C(1+\varepsilon)^n \lambda_n(K).$$

Dado que K era un cubo arbitrario de la colección K_i y que ε era un real positivo cualquiera, se obtiene que la desigualdad

$$\lambda_n(\varphi(W)) = \sum_i \lambda_n(\varphi(K_i)) \leq \sum_i C(1+\varepsilon)^n \lambda_n(K_i) = C(1+\varepsilon)^n \lambda_n(W)$$

es válida para todo $\varepsilon > 0$, es decir $\lambda_n(\varphi(W)) \leq C\lambda_n(W)$.

Ahora suponemos que W es un subconjunto abierto de U tal que $|J_\varphi(x)| < C$ se tiene para todo $x \in W$. Podemos entones escojer una sucesión creciente de conjuntos abiertos $(W_n)_{n\in\mathbb{N}}$ tal que $W = \bigcup_{n\in\mathbb{N}} W_n$ y tal que la cerradura de cada W_n es compacta y contenida en U. Entonces cada W_n satisface $\lambda_n(\varphi(W_n)) \leq C\lambda_n(W_n)$ y tenemos:

$$\lambda_n(\varphi(W)) \leq \lim_{n\to\infty} \lambda_n(\varphi(W_n)) \leq \lim_{n\to\infty} C\lambda_n(W_n) = C\lambda_n(W). \tag{1.38}$$

Finalmente, sea A un subconjunto boreliano de U tal que $|J_\varphi(x)| < C$ para todo $x \in A$ y sea $\varepsilon > 0$ un real. Si W es un subconjunto abierto de U que contiene A y si $W_{C+\varepsilon}$ está definido por

$$W_{C+\varepsilon} = \{x \in W : |J_\varphi(x)| < C + \varepsilon\},$$

entonces la desigualdad (1.38), con C reemplazado por $C + \varepsilon$ implica que

$$\lambda_n(\varphi(A)) \leq \lambda_n(\varphi(W_{C+\varepsilon})) \leq (C+\varepsilon)\lambda_n(W_{C+\varepsilon}) \leq (C+\varepsilon)\lambda_n(W);$$

dado que ε es arbitrario y que la medida λ_n es regular obtenemos el lema. ■

Una vez que disponemos de estos dos lemas, podemos atacar con toda comodidad la demostración del teorema de cambio de variable.

Demostración del Teorema 1.3.1. Supongamos primero que A es un subconjunto boreliano de U tal que $\lambda_n(A) < +\infty$. Para cada $j \geq 1$ definimos los conjuntos $A_{j,k}$ con $k = 1, 2, ...$ de la siguiente forma

$$A_{j,k} = \left\{x \in A : \frac{k-1}{j} \leq |J_\varphi(x)| < \frac{k}{j}\right\}.$$

Vemos entonces que el Lema 1.3.5 implica que $\lambda_n(\varphi(A_{j,k})) \leq \frac{k}{j}\lambda_n(A_{j,k})$. Además, si $k > 1$ entonces se tiene[7] $|J_{\varphi^{-1}}(y)| \leq \frac{j}{(k-1)}$ para todo $y \in \varphi(A_{j,k})$. Nótese en particular que si $y = \varphi(x)$ entonces

$$|J_{\varphi^{-1}}(y)J_\varphi(x)| = |det((\varphi^{-1})'(y) \circ \varphi'(x))| = |det(I)| = 1.$$

Ahora, si aplicamos el Lema 1.3.5 a φ^{-1} obtenemos la estimación

$$\lambda_n(\varphi^{-1}(\varphi(A_{j,k}))) \leq \frac{j}{k-1}\lambda_n(\varphi(A_{j,k}))$$

y por lo tanto se obtiene

$$\frac{k-1}{j}\lambda_n(A_{j,k}) \leq \lambda_n(\varphi(A_{j,k})) \leq \frac{k}{j}\lambda_n(A_{j,k})$$

para todo k. Además se tiene que

$$\frac{k-1}{j}\lambda_n(A_{j,k}) \leq \int_{A_{j,k}} |J_\varphi(x)|dx \leq \frac{k}{j}\lambda_n(A_{j,k})$$

para todo k. De esta manera, por las dos fórmulas precedentes, obtenemos

$$\left|\lambda_n(\varphi(A_{j,k})) - \int_{A_{j,k}} |J_\varphi(x)|dx\right| \leq \frac{k}{j}\lambda_n(A_{j,k}) - \frac{k-1}{j}\lambda_n(A_{j,k}) = \frac{1}{j}\lambda_n(A_{j,k}).$$

Pero dado que se tiene $A = \bigcup_{k\geq 1} A_{j,k}$ podemos escribir

$$\left|\lambda_n(\varphi(A)) - \int_A |J_\varphi(x)|dx\right| \leq \sum_{k\geq 1} \frac{1}{j}\lambda_n(A_{j,k}) = \frac{1}{j}\lambda_n(A),$$

[7] Recordar la fórmula (1.28).

de manera que haciendo tender $j \longrightarrow +\infty$ se obtiene la identidad $\lambda_n(\varphi(A)) = \int_A |J_\varphi(x)|dx$. Hemos entonces demostrado (1.34) en el caso cuando $\lambda_n(A) < +\infty$. Si A es un subconjunto boreliano cualquiera de U, se lo puede expresar como la unión creciente de una sucesión $(A_k)_{k\in\mathbb{N}}$ de conjuntos borelianos de medida finita y tomando el límite sobre k en la relación $\lambda_n(\varphi(A_k)) = \int_{A_k} |J_\varphi(x)|dx$ se obtiene

$$\lambda_n(f(A)) = \int_A |J_f(x)|dx$$

lo que completa la demostración de (1.34). ■

C) Aplicaciones y Cambios de Variable Usuales

En los primeros años de estudios se enseñan ciertos cambios de variable que son de gran utilidad para resolver y calcular diversos problemas. En esta sección recordamos estos cambios de variable y damos algunas aplicaciones concretas del teorema general de cambio de medidas.

• **Cambio de coordenadas polares en $\mathbb{R}^2$.** Sean ρ y θ dos reales tales que $0 < \rho < +\infty$ y $0 < \theta < 2\pi$. Definimos una aplicación φ del conjunto abierto $]0, +\infty[\times]0, 2\pi[$ sobre el conjunto abierto $\mathbb{R}^2 \setminus D$ (en donde $D = [0, +\infty[$ es una recta de $\mathbb{R}^2$) de la siguiente manera:

$$\begin{array}{rcl} \varphi :]0, +\infty[\times]0, 2\pi[& \longrightarrow & \mathbb{R}^2 \setminus D \\ (\rho, \theta) & \longmapsto & \big(\rho\cos(\theta), \rho\sin(\theta)\big) \end{array} \tag{1.39}$$

Vemos entonces que es posible relacionar todo punto (x, y) del conjunto $\mathbb{R}^2 \setminus D$ con un punto (ρ, θ) del conjunto $]0, +\infty[\times]0, 2\pi[$ por medio de la relación $x = \rho\cos(\theta)$ y $y = \rho\sin(\theta)$. Además, no es difícil verificar que esta aplicación y su inversa son de clase $\mathcal{C}^1$ de manera que podemos calcular sin problema su matriz Jacobiana que está dada por

$$J_\varphi(\rho, \theta) = \begin{bmatrix} \cos(\theta) & -\rho\sin(\theta) \\ \sin(\theta) & \rho\cos(\theta) \end{bmatrix} \qquad \text{y se tiene además que } det(J_\varphi) = \rho.$$

Finalmente, dado que la recta D es de medida de Lebesgue λ_2-nula, podemos aplicar el Teorema 1.3.1 para obtener la fórmula

$$\int_{\mathbb{R}^2} f(x, y)dxdy = \int_0^{+\infty}\int_0^{2\pi} f\big(\rho\cos(\theta), \rho\sin(\theta)\big)\rho\, d\rho d\theta. \tag{1.40}$$

Esta fórmula expresa el cambio de coordenadas de las coordenadas cartesianas (x, y) a las coordenadas polares (ρ, θ).

• **Cambio de coordenadas esféricas en $\mathbb{R}^3$.**

Este ejemplo es una generalización del caso anterior. Consideramos las variables reales (ρ, θ, φ) tales que $0 < \rho < +\infty$, $0 < \theta < \pi$ y $0 < \varphi < 2\pi$; definimos las relaciones siguientes:

$$\begin{cases} x = \rho\sin(\theta)\cos(\varphi) \\ y = \rho\sin(\theta)\sin(\varphi) \\ z = \rho\cos(\theta). \end{cases}$$

Vemos entonces que estas relaciones definen un difeomorfismo φ del cilindro abierto

$$\Delta = \{(\rho, \theta, \varphi) : 0 < \rho < +\infty,\ 0 < \theta < \pi,\ 0 < \varphi < 2\pi, \}$$

sobre el espacio $\mathbb{R}^3$ privado del semi-plano cerrado de coordenadas $y = 0, x \geq 0$. La matriz Jacobiana de esta transformación es

$$J_\varphi(\rho, \theta, \varphi) = \begin{bmatrix} \sin(\theta)\cos(\varphi) & \rho\cos(\theta)\cos(\varphi) & -\rho\sin(\theta)\sin(\varphi) \\ \sin(\theta)\sin(\varphi) & \rho\cos(\theta)\sin(\varphi) & \rho\sin(\theta)\cos(\varphi) \\ \cos(\theta) & -\rho\sin(\theta) & 0 \end{bmatrix} \quad \text{y se tiene que } det(J_\varphi) = \rho^2\sin(\theta).$$

Obtenemos entonces la fórmula

$$\int_{\mathbb{R}^3} f(x, y, z)dxdydz = \int_0^{+\infty}\int_0^{\pi}\int_0^{2\pi} f\big(\rho\sin(\theta)\cos(\varphi), \rho\sin(\theta)\sin(\varphi), \rho\cos(\theta)\big)\rho^2\sin(\theta)\, d\rho d\theta d\varphi.$$

• **Cambio de coordenadas esféricas en $\mathbb{R}^n$.**

Es posible extender los resultados anteriores utilizando el siguiente cambio de variable

$$\begin{aligned} x_1 &= \rho\cos(\theta_1) \\ x_2 &= \rho\sin(\theta_1)\cos(\theta_2) \\ x_3 &= \rho\sin(\theta_1)\sin(\theta_2)\cos(\theta_3) \\ &\cdots \\ x_{n-1} &= \rho\sin(\theta_1)\sin(\theta_2)\cos(\theta_3)\cdots\sin(\theta_{n-2})\cos(\theta_{n-1}) \\ x_n &= \rho\sin(\theta_1)\sin(\theta_2)\cos(\theta_3)\cdots\sin(\theta_{n-2})\sin(\theta_{n-1}) \end{aligned}$$

que permite obtener un difeomorfismo del abierto definido por las restricciones $0 < \rho < +\infty$, $0 < \theta_1, ..., \theta_{n-2} < \pi$ y $0 < \theta_{n-1} < 2\pi$ sobre el espacio $\mathbb{R}^n$ al cual se ha quitado un conjunto cerrado. El determinante Jacobiano es entonces

$$det(J_\varphi) = \rho^{n-1}\cos^{n-2}(\theta_1)\cos^{n-3}(\theta_2)\cdots\cos(\theta_{n-2})$$

Obtenemos finalmente la fórmula siguiente:

$$\int_{\mathbb{R}^n} f(x)dx = \int_0^{+\infty}\int_0^{\pi}\cdots\int_0^{2\pi} f(\rho\cos(\theta_1), ..., \rho\sin(\theta_1)\cdots\sin(\theta_{n-1}))\rho^{n-1}d\rho d\theta_1\cdots d\theta_{n-1} \qquad (1.41)$$

• **Medida invariante sobre la esfera unidad.**

Consideramos aquí la esfera unidad $\mathbb{S}^{n-1}$ definida como el subconjunto de $\mathbb{R}^n$ dado por la condición $\{x \in \mathbb{R}^n : |x| = (x_1^2 + \cdots + x_n^2)^{\frac{1}{2}} = 1\}$. Notaremos φ el homeomorfismo $(\xi, \rho) \longmapsto \rho\xi$ definido de $\mathbb{S}^{n-1} \times \mathbb{R}_+^*$ sobre $\mathbb{R}^n \setminus \{0\}$.

Si A es una parte de $\mathbb{S}^{n-1}$, notaremos $\Gamma(A)$ el sector cónico definido por $\varphi(A\times]0, 1]) = \{\rho\xi : 0 < \rho \leq 1, \xi \in A\}$. Vemos que si A es una parte boreliana de la esfera $\mathbb{S}^{n-1}$, entonces se tiene que $\varphi(A\times]0, 1])$ es una parte boreliana de $\mathbb{R}^n$ puesto que es la imagen recíproca del boreliano $A\times]0, 1]$, subconjunto de $\mathbb{S}^{n-1} \times \mathbb{R}_+^*$, por medio de la aplicación continua φ^{-1}.

Definición 1.3.4 (medida invariante sobre la esfera unidad) *Sea $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ la esfera unidad de $\mathbb{R}^n$. Para todo subconjunto boreliano A de $\mathbb{S}^{n-1}$ definimos la* medida invariante *sobre $\mathbb{S}^{n-1}$ por*

$$\sigma(A) = n\lambda_n(\Gamma(A)).$$

Esta medida es *invariante* en el sentido que para toda transformación ortogonal T se tiene $T(\overline{B}(0,1)) = \overline{B}(0,1)$ y por la Proposición 1.3.3 se tiene $\lambda_n(T(\overline{B}(0,1))) = |det(T)|\lambda_n(\overline{B}(0,1))$, y como $|det(T)| = 1$ se obtiene la invariancia de la medida σ.

En particular se tiene la siguiente fórmula de integración

$$\int_{\mathbb{R}^n} f(x)dx = \int_0^{+\infty} \left(\int_{\mathbb{S}^{n-1}} f(\rho\xi)d\sigma(\xi) \right) \rho^{n-1} d\rho \tag{1.42}$$

válida para toda función integrable $f : \mathbb{R}^n \longrightarrow \mathbb{K}$. El lector notará sin problema la semejanza entre esta expresión y la fórmula (1.41).

Veamos ahora unos pocos ejemplos en dónde estos cambios de variable permiten resolver problemas de manera particularmente cómoda y directa.

(i) **Integral de una Gaussiana.** El primer ejemplo es un cálculo totalmente clásico en donde se evalua la integral

$$I = \int_{-\infty}^{+\infty} e^{-x^2} dx.$$

Para calcular esta integral razonamos de la siguiente manera: por el teorema de Fubini observamos que se tiene la identidad

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dxdy = \left(\int_{-\infty}^{+\infty} e^{-x^2}dx\right)\left(\int_{-\infty}^{+\infty} e^{-y^2}dy\right) = I^2.$$

Utilizando ahora el cambio en coordenadas polares (1.39) y la fórmula (1.40) obtenemos

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dxdy = \int_0^{+\infty}\int_0^{2\pi} \rho e^{-\rho^2}\, d\rho d\theta = 2\pi \int_0^{+\infty} \rho e^{-\rho^2} d\rho = \pi.$$

Tenemos entonces que $I^2 = \pi$ de donde se deduce la fórmula clásica:

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}.$$

(ii) **Integral de una función radial.** Diremos que una función $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función *radial* si $f(x) = f(|x|)$ para todo $x \in \mathbb{R}^n$. Si f es una función integrable, a partir de la fórmua (1.42) tenemos que

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)dx &= \int_0^{+\infty} \left(\int_{\mathbb{S}^{n-1}} f(\rho\xi)d\sigma(\xi) \right) \rho^{n-1} d\rho = \int_0^{+\infty} \left(\int_{\mathbb{S}^{n-1}} f(|\rho\xi|)d\sigma(\xi) \right) \rho^{n-1} d\rho \\ &= \int_0^{+\infty} f(\rho) \left(\int_{\mathbb{S}^{n-1}} d\sigma(\xi) \right) \rho^{n-1} d\rho = s_n \int_0^{+\infty} f(\rho)\rho^{n-1} d\rho\end{aligned}$$

en donde hemos notado $s_n = \displaystyle\int_{\mathbb{S}^{n-1}} d\sigma(\xi)$ el área (o superficie) de la esfera unidad $\mathbb{S}^{n-1}$. La identidad

$$\int_{\mathbb{R}^n} f(x)dx = s_n \int_0^{+\infty} f(\rho)\rho^{n-1} d\rho \tag{1.43}$$

válida para las funciones radiales es de gran utilidad en muchas situaciones pues las funciones radiales juegan un rol muy importante.

(iii) **Area de la esfera unidad en $\mathbb{R}^n$.** Hemos visto en el primer volumen, al final de la Sección 3.4, cómo calcular la medida λ_n de la bola unidad $B^n = \overline{B}(0,1)$. Habíamos obtenido en particular que

$$v_n = \lambda_n(B^n) = \frac{\pi^{\frac{n}{2}}}{\Gamma(n/2+1)}$$

en donde Γ es la función de Euler definida por $\Gamma(\alpha) = \displaystyle\int_0^{+\infty} e^{-t}t^{\alpha-1}dt$ para $\alpha > 0$.

Vamos ahora a calcular el área o superficie s_n de la esfera unidad $\mathbb{S}^{n-1}$. Dado que la función indicatriz $\mathbb{1}_{\overline{B}(0,1)}$ es una función radial, podemos usar la fórmula (1.43) para obtener

$$v_n = \int_{\mathbb{R}^n} \mathbb{1}_{\overline{B}(0,1)}(x)dx = s_n \int_0^1 \rho^{n-1}d\rho = \frac{s_n}{n}$$

de modo que se tiene $s_n = nv_n$. Finalmente usando la propiedad $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ válida para todo $\alpha > 0$, deducimos la expresión siguiente

$$s_n = \frac{2\pi^{\frac{n}{2}}}{\Gamma(n/2)}.$$

1.3.2. Diferenciación de medidas

Al estudiar el teorema de Radon-Nikodym hemos visto que existe un concepto dado en la Definición 1.2.4 en la página 30 que hemos llamado la *Derivada de Radon-Nikodym* de una medida. Sin embargo, esta derivada estaba dada de manera totalmente formal por medio de una definición, es decir que su existencia estaba garantizada y no habíamos dado un procedimiento para obtener esta derivada. Es entonces necesario desarrollar más este tema y ver en qué sentido se puede relacionar esta derivada con las diferentes operaciones existentes sobre las medidas y así presentar las propiedades más relevantes de estas derivadas. En particular, vamos a relacionar esta noción de derivada con la integración y estaremos en medida de enunciar un análogo del teorema fundamental del cálculo en el marco general de las medidas.

Para llevar a cabo este objetivo, necesitaremos algunas definiciones y herramientas. Recordemos que si X es un espacio topológico separado, entonces una medida μ definida sobre una σ-álgebra $\mathscr{A}$ que contiene los borelianos de X es *regular* si verifica las condiciones siguientes:

- cada subconjunto compacto K de X es de medida finita ($\mu(K) < +\infty$).
- cada conjunto $A \in \mathscr{A}$ verifica $\mu(A) = \inf\{\mu(U);\ A \subset U \quad \text{con} \quad U \text{ abierto}\}$
- cada subconjunto abierto U verifica $\mu(U) = \sup\{\mu(K);\ K \subset U \quad \text{con} \quad K \text{ compacto}\}$.

Una de las principales propiedades de las medidas regulares reside en la posibilidad de aproximar la medida de un conjunto cualquiera por los conjuntos usuales en topología como los conjuntos abiertos y cerrados.

Para nuestros fines necesitaremos el siguiente concepto.

Definición 1.3.5 (Medida de Radon) *Sea X un espacio topológico separado. Diremos que una medida μ es una* medida de Radon *si es una medida definida sobre los borelianos de X y además es una medida regular.*

Dicho de otra manera, una medida de Radon, no es más que una medida boreliana regular. El ejemplo clásico de medida de Radon constituye la medida de Lebesgue sobre la σ-álgebra de los borelianos $\mathcal{B}or(\mathbb{R}^n)$. Observamos en particular que en el caso de $\mathbb{R}^n$, una medida de Radon definida sobre la σ-álgebra de los borelianos es automáticamente una medida σ-finita puesto que el espacio euclideo $\mathbb{R}^n$ puede ser recubierto por una familia numerable de compactos, cuya medida es finita. Para más detalles y propiedades sobre las medidas regulares ver la Sección 2.4.2 del Volumen 1.

A) Lemas de Recubrimiento

En esta subsección vamos a dar dos resultados conocidos como *lemas de recubrimiento*. Este tipo de lemas son de inmensa utilidad en el análisis y poseen muchísimas consecuencias como tendremos la oportunidad de verlo muy pronto. El primer resultado que presentamos, dado por el lema de Vitali[8], permite estudiar cómodamente la medida de Lebesgue sobre $\mathbb{R}^n$; mientras que el segundo, dado por el lema de Besicovitch[9], permite trabajar con medidas de Radon más generales.

Teorema 1.3.2 (Lema de recubrimiento de Vitali) *Sea $\mathscr{C}$ una colección de bolas cerradas de $\mathbb{R}^n$ de radio estrictamente positivo tal que*

$$\sup_{B\in\mathscr{C}} diam(B) < +\infty$$

Entonces existe una familia numerable $\mathcal{G}$ de bolas disjuntas de $\mathscr{C}$ tal que se tenga la inclusión

$$\bigcup_{B\in\mathscr{C}} B \subset \bigcup_{B\in\mathcal{G}} 5B$$

en donde $5B$ representa la bola de mismo centro que B pero de radio 5 veces mayor.

Este resultado nos indica que a partir de una colección cualquiera de bolas cerradas de radio finito, es posible escoger una familia numerable y disjunta de bolas que incluye la colección inicial. El precio a pagar para obtener estas dos propiedades (numerabilidad, bolas disjuntas) consiste en considerar bolas de radio 5 veces mayor.

Demostración. Notemos $R = \sup\limits_{B\in\mathscr{C}} \text{diam}(B)$ y para todo $j = 1, 2, ...$ consideremos el conjunto

$$\mathscr{C}_j = \left\{ B \in \mathscr{C} : \frac{R}{2^j} < diam(B) \leq \frac{R}{2^{j-1}} \right\}$$

y definimos ahora conjuntos $\mathcal{G}_j \subset \mathscr{C}_j$ de la siguiente manera:

- Para $j = 1$ utilizamos la relación de orden entre conjuntos dada por la inclusión y observamos que el conjunto formado por familias de bolas disjuntas de $\mathscr{C}_1$ es inductivo: en efecto, si A_n y A_m son dos familias de bolas disjuntas de $\mathscr{C}_1$, podemos ordenarlas usando la inclusión. En este punto usamos el Lema de Zorn para obtener la existencia de una familia maximal de bolas disjuntas de $\mathscr{C}_1$ y llamaremos esta familia $\mathcal{G}_1$.

- Para definir el resto de familias $\mathcal{G}_j$ con $j > 1$ procederemos por inducción. Supondremos que disponemos de las familias $\mathcal{G}_1, ..., \mathcal{G}_n$ y diremos a continuación cómo obtener la familia $\mathcal{G}_{n+1}$. Para empezar, consideramos la colección de bolas disjuntas B de $\mathscr{C}_{n+1}$ que verifican

$$B \cap B' = \emptyset, \text{ con } B' \in \bigcup_{i=1}^{n} \mathcal{G}_i,$$

[8]Guiseppe Vitali (1875-1932), matemático italiano.
[9]Abram Samoilovitch Besicovitch (1891-1970), matemático ruso.

vemos que este conjunto de bolas es disjunto de las familias $(\mathcal{G}_i)_{1\leq i\leq n}$ y vemos además que esta colección es inductiva para la inclusión, de manera que por el Lema de Zorn, existe un elemento maximal que notamos $\mathcal{G}_{n+1}$.

Una vez que disponemos de esta familia de conjuntos $(\mathcal{G}_j)_{j\geq 1}$ definimos el conjunto

$$\mathcal{G} = \bigcup_{j\geq 1} \mathcal{G}_j,$$

observamos que este conjunto $\mathcal{G}$ es numerable pues es por construcción una colección de bolas disjuntas de $\mathbb{R}^n$.

Para terminar, vamos a verificar que para toda bola $B \in \mathscr{C}$ existe una bola $B' \in \mathcal{G}$ tal que $B \cap B' \neq \emptyset$ y tal que $B \subset 5B'$. En efecto, sea $B \in \mathscr{C}$, existe por lo tanto un índice j tal que $B \in \mathscr{C}_j$ y entonces por la maximalidad de los conjuntos $\mathcal{G}_j$ existe una bola $B' \in \bigcup_{i=1}^{j} \mathcal{G}_i$ con $B \cap B' \neq \emptyset$. Pero como $\frac{R}{2^j} \leq \text{diam}(B) \leq \frac{R}{2^{j-1}}$ y $\frac{R}{2^j} \leq \text{diam}(B')$ se tiene que $\text{diam}(B) \leq 2\text{diam}(B')$ y por lo tanto se obtiene que $B \subset 5B'$. Con esto se termina la demostración del teorema. ■

Para continuar necesitaremos un concepto adicional.

Definición 1.3.6 (Recubrimiento fino) *Sea X un subconjunto de $\mathbb{R}^n$. Diremos que un recubrimiento por medio de bolas cerradas $\mathscr{C}$ del conjunto X es un* recubrimiento fino *si para todo punto $x \in X$ se tiene*

$$\inf_{B\ni x, B\in\mathscr{C}} diam(B) = 0.$$

Esta noción de recubrimiento fino nos indica que hay suficientes bolas en el recubrimiento que nos permiten acercarnos tanto como queremos a cada punto del conjunto X.

Tenemos así el siguiente corolario.

Corolario 1.3.1 *Sea X un subconjunto de $\mathbb{R}^n$ y sea $\mathscr{C}$ un recubrimiento fino de X por medio de bolas cerradas de diámetro finito, es decir $\sup\limits_{B\in\mathscr{C}} diam(B) < +\infty$. Entonces existe una familia numerable $\mathcal{G}$ de bolas disjuntas de $\mathscr{C}$ tal que para todo subconjunto finito de bolas $\{B_1, B_2, ..., B_n\} \subset \mathscr{C}$ se tenga*

$$X \setminus \bigcup_{k=1}^{n} B_k \subset \bigcup_{B\in\mathcal{G}\setminus\{B_1,B_2,...,B_n\}} 5B$$

Prueba. Dado el recubrimiento fino $\mathscr{C}$ podemos aplicar el lema de recubrimiento de Vitali para obtener una familia numerable $\mathcal{G}$. Sea ahora una colección de bolas $\{B_1, B_2, ..., B_n\} \subset \mathscr{C}$. Si se tiene $X \subset \bigcup_{k=1}^{n} B_k$ no hay nada que verificar, de modo que suponemos que existe $x \in X \setminus \bigcup_{k=1}^{n} B_k$. Dado que las bolas del recubrimiento $\mathscr{C}$ son cerradas y dado que es un recubrimiento fino, existe una bola $B \in \mathscr{C}$ tal que $x \in B$ y tal que $B \cap B_k = \emptyset$ para todo $1 \leq k \leq n$. Entonces, usando el mismo argumento utilizado en la demostración del teorema anterior, existe una bola $B' \in \mathcal{G}$ tal que $B \cap B' \neq \emptyset$ y $B \subset 5B'$. ■

Vamos ahora a aplicar estos resultados anteriores a la medida de conjuntos abiertos. Más precisamente, el siguiente corolario nos indica que es posible "rellenar" interiormente un conjunto abierto por medio de una colección de bolas disjuntas.

Corolario 1.3.2 *Sea $A \subset \mathbb{R}^n$ un conjunto abierto y sea $\delta > 0$ un real. Existe entonces una familia numerable $\mathcal{G}$ de bolas cerradas disjuntas contenidas en A, tales que $diam(B) < \delta$ para todo $B \in \mathcal{G}$, y tal que se tenga*

$$\lambda_n\left(A \setminus \bigcup_{B\in\mathcal{G}}\right) = 0.$$

Prueba. Empecemos suponiendo que $\lambda_n(A) < +\infty$. Existe entonces una sucesión finita $(B_i)_{1\leq i\leq n_1}$ de bolas cerradas disjuntas en A tales que $diam(B_i) < \delta$ y tales que

$$\lambda_n\left(A \setminus \bigcup_{i=1}^{n_1} B_i\right) \leq \theta\lambda_n(A), \tag{1.44}$$

con $1-\frac{1}{5^n} < \theta < 1$. En efecto sea $\mathscr{C}_1 = \{B \subset A : diam(B) < \delta\}$ entonces por el Teorema 1.3.2 sabemos que existe una familia numerable disjunta $\mathcal{G}_1 \subset \mathscr{C}_1$ tal que $A \subset \bigcup_{B\in\mathcal{G}_1} 5B$. Entonces podemos escribir

$$\lambda_n(A) \leq \sum_{B\in\mathcal{G}_1} \lambda_n(5B) = 5^n \sum_{B\in\mathcal{G}_1} \lambda_n(B) = 5^n\lambda_n\left(\bigcup_{B\in\mathcal{G}_1} B\right),$$

y por lo tanto se tiene $\frac{1}{5^n}\lambda_n(A) \leq \lambda_n\left(\bigcup_{B\in\mathcal{G}_1} B\right)$ de donde se deduce

$$\lambda_n\left(A \setminus \bigcup_{B\in\mathcal{G}_1} B\right) \leq \left(1-\frac{1}{5^n}\right)\lambda_n(A).$$

Ahora, como la familia $\mathcal{G}_1$ es numerable existen bolas $B_1, ..., B_n$ en $\mathcal{G}_1$ que verifican (1.44) con un cierto θ tal que $1-\frac{1}{5^n} < \theta < 1$.

Definimos ahora los conjuntos

$$A_2 = A \setminus \bigcup_{i=1}^{n_1} B_i \quad \text{y} \quad \mathscr{C}_2 = \{B \subset A_2 : diam(B) < \delta\}.$$

Procediendo de la misma manera, podemos encontrar una familia finita de bolas disjuntas $B_{n_1+1}, ..., B_{n_2}$ en $\mathscr{C}_2$ tales que

$$\lambda_n\left(A \setminus \bigcup_{i=1}^{n_2} B_i\right) = \lambda_n\left(A_2 \setminus \bigcup_{i=n_1+1}^{n_2} B_i\right) \leq \theta\lambda_n(A_2) \leq \theta^2\lambda_n(A).$$

Repitiendo este procedimiento se obtiene una familia numerable de bolas disjuntas tales que

$$\lambda_n\left(A \setminus \bigcup_{i=1}^{n_k} B_i\right) \leq \theta^k\lambda_n(A).$$

Pero dado que $\theta^k \underset{k\to+\infty}{\longrightarrow} 0$ se obtiene el corolario en el caso cuando $\lambda_n(A) < +\infty$.

En el caso cuando $\lambda_n(A) = +\infty$ es suficiente aplicar el mismo razonamiento a los conjuntos $A_n = \{x \in A : k < |x| < k+1\}$ para todo $k \in \mathbb{N}$. ■

Este teorema está muy bien adaptado a la medida de Lebesgue pues se usa de forma intensiva la siguiente propiedad de "duplicación" del volumen de las bolas de $\mathbb{R}^n$:

$$\lambda_n(B(x_0, 2r)) = 2^n\lambda_n(B(x_0, r)).$$

Sin embargo, cuando es necesario trabajar con medidas más generales es necesario utilizar un resultado diferente que está dado en el teorema a continuación.

Teorema 1.3.3 (Lema de recubrimiento de Besicovitch) *Para todo $n \geq 1$ existe una constante K_n tal que para toda colección $\mathscr{C}$ de bolas cerradas de $\mathbb{R}^n$, cuyos radios son estrictamente positivos y uniformemente acotados (es decir $\sup\limits_{B\in\mathscr{C}} diam(B) < +\infty$), es posible encontrar familias $\mathcal{G}_1, ..., \mathcal{G}_{K_n} \subset \mathscr{C}$ tal que cada familia $\mathcal{G}_i$ está conformada por bolas dos a dos disjuntas y tal que el conjunto $A = \{a_1, a_2..., \}$ formado por los centros de las bolas de la colección $\mathscr{C}$ puede ser recubierto por los elementos del conjunto $\mathcal{G}_1 \cup \cdots \cup \mathcal{G}_{K_n}$.*

Antes de pasar a la demostración de este resultado es conveniente explicar la diferencia fundamental que existe entre el Lema de recubrimiento de Vitali dado en el Teorema 1.3.2 y el Lema de Besicovitch enunciado en las líneas precedentes. En ambos casos se tiene como dato inicial una colección $\mathscr{C}$ de bolas cerradas de $\mathbb{R}^n$ de radio estrictamente positivo y finito. En el lema de Vitali se obtiene un recubrimiento disjunto de la colección $\mathscr{C}$ por medio de una colección numerable, pero de bolas de radio mayor al inicialmente dado. En el lema de Besicovitch se obtiene en cambio un recubrimiento, numerable por medio de familias disjuntas, del conjunto formado por los centros de las bolas de la $\mathscr{C}$ (es decir que no se recubre la colección de bolas, pero solo el conjunto de formado por sus centros) con la particularidad de usar las mismas bolas dadas inicialmente.

Demostración. Para empezar suponemos que el conjunto A formado por los centros de las bolas de la familia $\mathscr{C}$ es un conjunto acotado, es decir que $diam(A) < +\infty$. Sea $R = \sup\{r : \overline{B}(a, r) \in \mathscr{C}\}$, podemos entonces encontrar una bola $B_1 = \overline{B}(a_1, r_1) \in \mathscr{C}$ tal que $r_1 > \frac{3}{4}R$. Definimos ahora bolas B_j con $j > 1$ de forma recursiva de la siguiente manera: sea $A_j = A \setminus \bigcup_{i=1}^{j-1} B_i$; si el conjunto A_j es vacío, entonces damos por terminada nuestra construcción y, fijando $J = j - 1$, obtenemos una colección de J bolas $B_1, ..., B_J$.

Si A_j no es vacío, escogemos $B_j = \overline{B}(a_j, r_j) \in \mathscr{C}$ de la siguiente manera:

$$a_j \in A_j \quad \text{y} \quad r_j > \frac{3}{4}\sup\{r : \overline{B}(a, r) \in \mathscr{C},\ a \in A_j\}$$

y seguimos el proceso. En el caso de una sucesión numerable de bolas B_j, notaremos $J = +\infty$.

Al fijar de esta manera estas bolas, con sus respectivos centros y radios, tenemos las propiedades a continuación:

(a) si $j > i$, entonces $r_j \leq \frac{4}{3} r_i$,

(b) las bolas cerradas $\overline{B}(a_j, \frac{r_j}{3})$ son disjuntas, y si $J = +\infty$ se tiene que $r_j \to 0$,

(c) se tiene $A \subset \bigcup_{j=1}^{J} \overline{B}(a_j, r_j)$.

En efecto, el punto (a) se tiene por definición de los radios r_i y por la inclusión $a_j \in A_j \subset A_i$. El punto (b) se deduce del hecho que si $j > i$, entonces $a_j \notin B_i$ y por lo tanto, si comparamos la distancia entre los centros de a_i y a_j se tiene por el primer punto que

$$|a_j - a_j| > r_i = \frac{r_i}{3} + \frac{2r_i}{3} \geq \frac{r_i}{3} + \frac{2}{3} \times \frac{3r_j}{4} > \frac{r_i}{3} + \frac{r_j}{3},$$

de manera que las bolas $\overline{B}(a_i, \frac{r_i}{3})$ y $\overline{B}(a_j, \frac{r_j}{3})$ son disjuntas. Nótese que las bolas $\overline{B}(a_i, r_i)$ pueden *no* ser disjuntas, es por esta razón que se divide su radio por tres. Como el conjunto A es acotado, se obtiene que $r_j \to 0$ en el caso de una sucesión infinita.

Finalmente, si $J < +\infty$ el último punto (c) es inmediato por construcción. En el caso cuando $J = +\infty$ vemos que si a pertenece al conjunto A, entonces es un centro de una cierta bola y por lo tanto existe un radio $r > 0$ tal que la bola $\overline{B}(a, r)$ pertenece a la colección $\mathscr{C}$. Dado que $r_j \to 0$, existe r_j tal que $r_j < \frac{3}{4}r$, luego, por construcción de los conjuntos A_j, se tiene que $a \notin A_j$ de manera que $a \in \bigcup_{i=1}^{j-1} B_i$. Realizando esto para todo punto $a \in A$ se obtiene lo deseado.

Con este proceder obtenemos una familia numerable $\mathscr{B} = \{B_1, B_2, ...\}$ de bolas que contiene el conjunto A. Pero esta familia no es disjunta y son necesarias unas algunas etapas adicionales para obtener el resultado deseado.

Fijamos ahora $k > 1$ y nos interesamos en conocer el número máximo de bolas de la familia $\mathscr{B}$ que intersecta con la bola B_k. Para ello consideramos los conjuntos

$$I_k = \{j \in \mathbb{N} : j < k, B_j \cap B_k \neq \emptyset\}, \quad \text{y} \quad M_k = I_k \cap \{j \in \mathbb{N} : r_j \leq 3r_k\}.$$

Estudiaremos primero la cardinalidad del conjunto M_k para luego estudiar la cardinalidad del conjunto $I_k \setminus M_k$. Para el estudio de la cardinalidad del conjunto M_k vamos a ver que se tiene

$$\operatorname{card}(M_k) \leq 20^n.$$

En efecto, si $j \in M_k$ entonces por definición las bolas B_j y B_k tienen una intersección no vacía y se tiene $r_j \leq 3r_k$, de manera que podemos escoger un punto cualquiera $x \in \overline{B}(a_j, \frac{r_j}{3})$ y podemos escribir

$$|x - a_k| \leq |x - a_j| + |a_j - a_k| \leq \frac{r_j}{3} + r_j + r_k \leq 5r_k,$$

es decir, $\overline{B}(a_j, \frac{r_j}{3}) \subset \overline{B}(a_k, 5r_k)$. Ahora, por el hecho que las bolas $\overline{B}(a_j, \frac{r_j}{3})$ son disjuntas y por la propiedad (a) anterior, obtenemos

$$\lambda_n(\overline{B}(a_k, 5r_k)) \geq \sum_{j \in M_k} \lambda_n(\overline{B}(a_j, \frac{r_j}{3})) = \sum_{j \in M_k} v_n r_j^n 3^{-n} \geq \sum_{j \in M_k} v_n r_k^n 4^{-n} = \operatorname{card}(M_k) v_n r_k^n 4^{-n},$$

donde v_n es el volumen de la bola unidad en $\mathbb{R}^n$. A partir de estas estimaciones se obtiene que $5^n \geq \operatorname{card}(M_k) 4^{-n}$ que es la estimación buscada.

Estimamos ahora la cardinalidad del conjunto $I_k \setminus M_k$. Para ellos consideramos dos elementos distintos $i, j \in I_k \setminus M_k$, tenemos entonces $1 \leq i, j < k$, $B_i \cap B_k \neq \emptyset$, $B_j \cap B_k \neq \emptyset$ y $r_i > 3r_k$, $r_j > 3r_k$.

Podemos suponer, sin pérdida de generalidad, que $a_k = 0$. Sea ahora $0 \leq \theta \leq \pi$ el ángulo formado entre los vectores a_i y a_j y buscamos estimar el coseno de este ángulo, en particular queremos verificar que se tiene

$$\cos(\theta) \leq \frac{61}{64}. \tag{1.45}$$

Dado que $1 \leq i, j < k$, por el proceso de selección de los centros a_k, se tiene que $0 = a_k \notin B_i \cup B_j$ y entonces tenemos $r_i < |a_i|$, $r_j < |a_j|$. Podemos además suponer que $|a_i| \leq |a_j|$ y entonces

$$3r_k < r_i < |a_i| \leq r_i + r_k, \quad 3r_k < r_j < |a_j| \leq r_j + r_k.$$

Observamos ahora que si $\cos(\theta) > \frac{5}{6}$, entonces $a_i \in B_j$. En efecto, si se tiene $|a_i - a_j| \geq |a_j|$ (esto implica que $a_i \notin B_j$), entonces, por la Ley de los Cosenos se tiene

$$\cos(\theta) = \frac{|a_i|^2 + |a_j|^2 - |a_i - a_j|^2}{2|a_i||a_j|} \leq \frac{|a_i|}{2|a_j|} \leq \frac{1}{2} < \frac{5}{6}.$$

Ahora, si se tiene en cambio $|a_i - a_j| \leq |a_j|$ pero $a_i \notin B_j$, entonces $r_j < |a_i - a_j|$ y entonces

$$\cos(\theta) = \frac{|a_i|^2 + |a_j|^2 - |a_i - a_j|^2}{2|a_i||a_j|} \leq \frac{|a_i|}{2|a_j|} + \frac{(|a_j| - |a_i - a_j|)(|a_j| + |a_i - a_j|)}{2|a_i||a_j|}$$

usando la desigualdad $|a_j| + |a_i - a_j| \leq 2|a_j|$ tenemos

$$\cos(\theta) \leq \frac{1}{2} + \frac{r_j + r_k - r_j}{r_i} \leq \frac{5}{6}.$$

Como vemos, en todos los casos se tiene, bajo las hipótesis hechas, que si $\cos(\theta) > \frac{5}{6}$, entonces $a_i \in B_j$.

Necesitamos ahora verificar la siguiente desigualdad si $a_i \in B_j$:

$$0 \leq |a_i - a_j| + |a_i| - |a_j| \leq \frac{8}{3}(1 - \cos(\theta))|a_j|. \tag{1.46}$$

Dado que $a_i \in B_j$, por selección de los centros se tiene $i < j$ y tenemos $a_j \notin B_i$ de manera que $|a_i - a_j| > r_i$, entonces, recordando que habíamos fijado $|a_i| \leq |a_j|$, podemos escribir

$$\begin{aligned}
0 \quad &\leq \quad \frac{|a_i - a_j| + |a_i| - |a_j|}{|a_j|} \leq \frac{|a_i - a_j| + |a_i| - |a_j|}{|a_j|}\frac{|a_i - a_j| - |a_i| + |a_j|}{|a_i - a_j|} \\
&\leq \quad \frac{|a_i - a_j|^2 - (|a_j| - |a_i|)^2}{|a_j||a_i - a_j|} = \frac{2|a_i|(1 - \cos(\theta))}{|a_i - a_j|} \leq \frac{2(r_i + r_k)(1 - \cos(\theta))}{r_i} \\
&\leq \quad \frac{2(1 + \frac{1}{3})r_i(1 - \cos(\theta))}{r_i} = \frac{8}{3}(1 - \cos(\theta)).
\end{aligned}$$

Podemos ahora verificar la desigualdad (1.45). En efecto, si $\cos(\theta) < \frac{5}{6}$ entonces $\cos(\theta) < \frac{61}{64}$, por otro lado, si $\cos(\theta) > \frac{5}{6}$, entonces por lo que hemos verificado se tiene que $a_i \in B_j$, entonces $i < j$ y por lo tanto $a_j \notin B_i$ de donde se deduce que $r_i < |a_i - a_j| \leq r_j$, adicionalmente se tiene $r_j \leq \frac{4}{3}r_i$ entonces, gracias a la estimación $r_j > 3r_k$ se obtiene

$$|a_i - a_j| + |a_i| - |a_j| \geq r_i + r_i - r_j - r_k \geq \frac{r_j}{2} - r_k \geq \frac{1}{6}r_j = \frac{1}{6}(\frac{3}{4})(r_j + \frac{1}{3}r_j) \geq \frac{1}{8}(r_j + r_k) \geq \frac{1}{8}|a_j|$$

lo que nos da $\frac{|a_j|}{8} \leq 8(1-\cos(\theta))\frac{|a_j|}{3}$ por (1.46) y a partir de esta desigualdad se obtiene que $\cos(\theta) \leq \frac{61}{64}$.

Gracias a estas observaciones geométricas, vamos a ver que es posible obtener la existencia de un número $L_n \in \mathbb{N}$, que depende únicamente de n, tal que

$$\operatorname{card}(I_k \setminus M_k) \leq L_n.$$

En efecto, sea $r_0 > 0$ tal que, si x es un vector tal que $|x| = 1$ y si $y, z \in \overline{B}(x, r_0)$, entonces el ángulo entre el punto y y el punto z es menor que $\theta_0 = \arccos(61/64)$. Sea ahora $L_n \in \mathbb{N}$ tal que la esfera unidad $\partial\overline{B}(0, 1)$ puede ser recubierta por L_n bolas de radio r_0 con centros en la esfera $\partial\overline{B}(0, 1)$ pero tal que esta esfera no pueda ser recubierta por $L_n - 1$ bolas de radio r_0 con centros en la esfera.

Vemos entonces, por dilatación, que la esfera ∂B_k puede ser recubierta por L_n bolas de radio $r_0 r_k$ con centros en ∂B_k. Usando (1.45), para todo $i, j \in I_k \setminus M_k$, el ángulo entre a_i y a_j es mayor que θ_0, de donde se deduce que los radios generados por los vectores $a_i - a_k$ y $a_j - a_k$ no pueden cruzar una misma bola de radio r_0 y de centro en ∂B_k. En particular, no pueden cruzar una misma bola del recubrimiento formado por L_n bolas y esto muestra que $\operatorname{card}(I_k \setminus M_k) \leq L_n$.

Sea ahora $N_n = 20^n + L_n + 1$ entonces, gracias a los resultados anterior podemos escribir

$$\operatorname{card}(I_k) = \operatorname{card}(M_k) + \operatorname{card}(I_k \setminus M_k) \leq 20^n + L_n < N_n,$$

y con esto hemos logrado dar una estimación del cardinal del conjunto I_k.

Con todos estos antecedentes, podemos proceder a la construcción de las familias $\mathcal{G}_i$. Para ello definimos una aplicación

$$\sigma : \{1, 2, ...\} \longrightarrow \{1, ..., N_n\}$$

de la siguiente manera: si $1 \leq k \leq N_n$ entonces fijamos $\sigma(k) = k$. En cambio si $k \geq N_n$ es necesario definir $\sigma(k+1)$ de otra forma y para ello observamos, por los cálculos anteriores, que tenemos

$$\text{card}(\{j \in \mathbb{N} : 1 \leq j \leq k, B_j \cap B_{k+1} \neq \emptyset\}) < N_n,$$

es decir que existe un entero $\ell \in \{1, ..., N_n\}$ con $B_j \cap B_{k+1} = \emptyset$ para todo $j \in \{1, ..., k\}$ tal que $\sigma(j) = \ell$ y definimos entonces $\sigma(k+1) = \ell$.

Finalmente definimos

$$\mathcal{G}_j = \{B_i \in \mathscr{C} : \sigma(i) = j\}, \quad 1 \leq j \leq N_n.$$

Tenemos por construcción de la aplicación σ que la colección $\mathcal{G}_j$ está formada por bolas disjuntas. En particular, se observa que cada bola B_i pertenece a alguna colección $\mathcal{G}_j$ y entonces tenemos

$$A \subset \bigcup_{j=1}^{J} B_j = \bigcup_{j=1}^{N_n} \bigcup_{B \in \mathcal{G}_j} B,$$

y de esta manera hemos obtenido el resultado deseado cuando el conjunto formado por los centros A es acotado.

Para terminar, veamos lo que sucede en el caso cuando el conjunto A no es acotado. Definimos entonces el conjunto $A_\ell = A \cap \{x \in \mathbb{R}^n : 6R(\ell - 1) \leq |x| < 6R\ell\}$, y notamos $\mathscr{C}^\ell$ la familia de bolas de $\mathscr{C}$ con centros en el conjunto A_ℓ. Como acabamos de demostrar, para cada ℓ existe una colección, a lo mucho numerable, $\mathcal{G}_j^\ell$ con $1 \leq j \leq N_n$, de bolas disjuntas tal que su unión recubre A_ℓ:

$$A_\ell \subset \bigcup_{j=1}^{N_n} \bigcup_{B \in \mathcal{G}_j^\ell} B.$$

Dado que los radios de las bolas son menores que R, vemos que ninguna bola de la colección $\mathscr{C}^\ell$ se intersecta con las bolas de la colección $\mathscr{C}^{\ell+2}$. En este punto, basta tomar, para todo $j \leq N_n$, la colección $\mathcal{G}_j = \bigcup_{\ell=1}^{+\infty} \mathcal{G}_j^{2\ell-1}$ y la colección $\mathcal{G}_{j+N_n} = \bigcup_{\ell=1}^{+\infty} \mathcal{G}_j^{2\ell}$. Al fijar $K_n = 2N_n$ y al considerar la unión de todas estas familias $\mathcal{G}_j$ terminamos la demostración de este teorema. ■

Una vez que disponemos de este resultado, vemos que es posible dar un resultado similar al enunciado en el Corolario 1.3.2 usando el Lema de recubrimiento de Besicovitch.

Corolario 1.3.3 *Sea μ^* una medida exterior definida sobre $\mathbb{R}^n$ y tal que $\mathcal{B}or(\mathbb{R}^n) \subset \mathscr{M}_{\mu^*}$. Sea $\mathscr{C}$ una colección de bolas cerradas de radio estrictamente positivo y notemos A el conjunto formado por sus centros. Supongamos que se tiene $\mu^*(A) < +\infty$ y que para todo centro $a \in A$ y para todo $\varepsilon > 0$, la colección $\mathscr{C}$ contiene una bola $\overline{B}(a, r)$ con $r < \varepsilon$. Entonces, para todo conjunto abierto no vacío $U \subset \mathbb{R}^n$, existe una familia numerable de bolas disjuntas $B_j \in \mathscr{C}$ tales que*

$$\bigcup_{j=1}^{+\infty} B_j \subset U \quad y \quad \mu^*\left((A \cap U) \setminus \bigcup_{j=1}^{+\infty} B_j\right) = 0.$$

Prueba. Suponemos evidentemente que $A \cap U \neq \emptyset$ pues de lo contrario no hay nada que demostrar. Sea K_n la constante obtenida en el anterior teorema y sea $\alpha \in]1 - \frac{1}{K_n}, 1[$. Mostremos que la colección $\mathscr{C}$ contiene una familia finita de bolas disjuntas $B_1, ..., B_{k_1}$ con la propiedad

$$\bigcup_{j=1}^{k_1} B_j \subset U, \quad \mu^* \left((A \cap U) \setminus \bigcup_{j=1}^{k_1} B_j \right) \leq \alpha \mu^*(A \cap U). \tag{1.47}$$

Para ello notamos $\mathscr{C}^1$ el subconjunto de $\mathscr{C}$ formado por las bolas de radio menor a 1 contenidas en U. Por el teorema de Besicovitch, existe una colección $\mathcal{G}_1^1, ..., \mathcal{G}_{K_n}^1$ formada por familias de bolas disjuntas que pertenecen todas ellas a la familia $\mathscr{C}^1$ tales que $A \cap U \subset \bigcup_{j=1}^{K_n} \bigcup_{B \in \mathcal{G}_j^1} B$. Entonces se tiene $\mu^*(A \cap U) \leq \sum_{j=1}^{K_n} \mu^* \left((A \cap U) \cap \bigcup_{B \in \mathcal{G}_j^1} B \right)$ de manera que existe $j \in \{1, ..., K_n\}$ tal que $\mu^* \left((A \cap U) \cap \bigcup_{B \in \mathcal{G}_j^1} B \right) \geq \frac{1}{K_n} \mu^*(A \cap U)$ y por lo tanto existe una familia finita de bolas $B_1, ..., B_{k_1} \in \mathcal{G}_j^1$ tal que $\mu^* \left((A \cap U) \cap \bigcup_{B \in \mathcal{G}_j^1} B \right) \geq (1 - \alpha) \mu^*(A \cap U)$ de manera que se obtiene (1.47) por la μ^*-medibilidad de los conjuntos B_i: en efecto se tiene $\mu^*(A \cap U) = \mu^* \left((A \cap U) \cap \bigcup_{i=1}^{k_1} Bi \right) + \mu^* \left((A \cap U) \setminus \bigcup_{i=1}^{k_1} Bi \right)$. Ahora, definimos el conjunto $U_2 = U \setminus \bigcup_{j=1}^{k_1} B_j$ y consideramos la familia $\mathscr{C}^2$ formada por todas las bolas de $\mathscr{C}$ contenidas en U_2 de radio menor a 1. Se tiene que el conjunto U_2 es abierto y podemos entonces encontrar una familia finita de bolas disjuntas $B_{k_1+1}, ..., B_{k_2}$ de $\mathscr{C}^2$ con la propiedad $\bigcup_{j=k_1+1}^{k_2} B_j \subset U_2$ y tal que $\mu^* \left((A \cap U) \setminus \bigcup_{j=1}^{k_2} B_j \right) = \mu^* \left((A \cap U_2) \setminus \bigcup_{j=k_1+1}^{k_2} B_j \right) \leq \alpha \mu^*(A \cap U_2) \leq \alpha^2 \mu^*(A \cap U)$. Por inducción, obtenemos una sucesión de bolas disjuntas $B_j \in \mathscr{C}$ tal que

$$\mu^* \left((A \cap U) \setminus \bigcup_{j=1}^{k_p} B_j \right) \leq \alpha^p \mu^*(A \cap U).$$

Dado que $\mu^*(A) < +\infty$ y que $\alpha^p \underset{p \to +\infty}{\longrightarrow} 0$ se obtiene el resultado deseado. ■

B) Diferenciación de Medidas

Una vez que disponemos de estos resultados de recubrimiento, vamos a ponerlos en acción al estudiar la diferenciación de medidas. Damos a continuación la definición de derivadas superior e inferior de una medida con respecto a otra medida.

Definición 1.3.7 (Derivadas de Medidas, densidad) *Sean μ y ν dos medidas de Radon (positivas) definidas sobre $(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n))$.*

1) *Definimos la* derivada superior *de la medida ν con respecto a la medida μ en el punto $x \in \mathbb{R}^n$ por*

$$\overline{D}_\mu(\nu)(x) = \begin{cases} \limsup\limits_{r \to 0} \frac{\nu(\overline{B}(x,r))}{\mu(\overline{B}(x,r))} & si\ \mu(\overline{B}(x,r)) > 0 \ para\ todo\ r > 0. \\ +\infty & si\ \mu(\overline{B}(x,r)) = 0 \ para\ algún\ r > 0. \end{cases} \tag{1.48}$$

2) *Definimos la* derivada inferior *de la medida ν con respecto a la medida μ en el punto $x \in \mathbb{R}^n$ por*

$$\underline{D}_\mu(\nu)(x) = \begin{cases} \liminf\limits_{r \to 0} \frac{\nu(\overline{B}(x,r))}{\mu(\overline{B}(x,r))} & si\ \mu(\overline{B}(x,r)) > 0 \ para\ todo\ r > 0. \\ +\infty & si\ \mu(\overline{B}(x,r)) = 0 \ para\ algún\ r > 0. \end{cases} \tag{1.49}$$

Las derivadas superior $\overline{D}_\mu(\nu)$ e inferior $\underline{D}_\mu(\nu)$ son funciones a valores en $\overline{\mathbb{R}}_+$ cuyos valores en el punto x están dadas por (1.48) y (1.49).

3) Diremos que la medida ν es diferenciable en el punto x con respecto a la medida μ *si las cantidades $\overline{D}_\mu(\nu)(x)$ y $\underline{D}_\mu(\nu)(x)$ existen y son iguales, diremos entonces que la* derivada de la medida ν con respecto a la medida μ en el punto x *está dada por la cantidad*

$$D_\mu(\nu)(x) = \overline{D}_\mu(\nu)(x) = \underline{D}_\mu(\nu)(x). \tag{1.50}$$

La derivada de la medida ν con respecto a la medida μ *es entonces la función $D_\mu(\nu)$ que está definida por la expresión (1.50) en cada punto x en donde ν es diferenciable. Notaremos también*

$$D_\mu(\nu) = \frac{d\nu}{d\mu}$$

Se suele también designar $D_\mu(\nu)$ como la densidad *de la medida ν con respecto a la medida μ.*

Observación 1.10

1) Notamos, de la misma manera que en la Definición 1.2.4, página 30, que la derivada de una medida que se obtiene por medio de este procedimiento no es una medida, sino una función.

2) En los puntos x en los cuales la derivada $D_\mu(\nu)(x)$ existe, el comportamiento de la cantidad $\nu(\overline{B}(x,r))$ debe estar controlado por comportamiento de la cantidad $\mu(\overline{B}(x,r))$ cuando $r \to 0$ y esta situación debe relacionarse a la continuidad absoluta requerida en la Definición 1.2.4.

Demos un ejemplo muy simple. Consideremos la medida ν sobre $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ definida para todo $A \in \mathcal{B}or(\mathbb{R})$ por

$$\nu(A) = \int_A \mathbb{1}_{]0,1[}(x)dx$$

y calculemos su derivada con respecto a la medida de Lebesgue λ. Observamos primero que si $x \notin]0,1[$ entonces se tiene inmediatamente que $\overline{D}_\lambda(\nu)(x) = \underline{D}_\lambda(\nu)(x) = 0$, de manera que es suficiente estudiar lo que sucede si $x \in]0,1[$. Sea $\overline{B}$ una bola de radio ε que contiene el punto x (en este caso, la bola es un intervalo cerrado), si ε es suficientemente pequeño se tiene que la bola $\overline{B}$ está completamente contenida en el conjunto $]0,1[$ y entonces $\overline{D}_\lambda(\nu)(x) = \underline{D}_\lambda(\nu)(x) = 1$. Finalmente podemos escribir que $D_\lambda(\nu)(x) = 1$ si $x \in]0,1[$ y que $D_\lambda(\nu)(x) = 0$ sino y de esta manera hemos calculado la derivada de la medida ν con respecto a la medida de Lebesgue.

Nos proponemos ahora estudiar en qué sentido la derivada $D_\mu(\nu)$ existe y cuáles son sus propiedades más inmediatas y para ello necesitaremos el siguiente lema.

Lema 1.3.6 *Sean μ, ν dos medidas de Radon definidas sobre $\mathbb{R}^n$ dotado de su estructura de espacio medible natural. Sea $0 < \rho < +\infty$ un real fijo. Entonces se tienen los puntos siguientes:*

1) Si A es un conjunto tal que $A \subset \{x \in \mathbb{R}^n : \underline{D}_\mu(\nu)(x) \leq \rho\}$ entonces $\nu(A) \leq \rho\mu(A)$.

2) Si A es un conjunto tal que $A \subset \{x \in \mathbb{R}^n : \overline{D}_\mu(\nu)(x) \geq \rho\}$ entonces $\nu(A) \geq \rho\mu(A)$.

Prueba. Empezamos con el primer punto. Podemos suponer que $\mu(\mathbb{R}^n) < +\infty$ y que $\nu(\mathbb{R}^n) < +\infty$ pues de otra manera podemos restringirnos a subconjuntos compactos de $\mathbb{R}^n$.

Fijemos $\varepsilon > 0$ y sea U un abierto tal que $A \subset U$. Definimos ahora la colección de bolas cerradas:

$$\mathscr{C} = \{B : B = \overline{B}(a,r), a \in A, B \subset U, \nu(B) \leq (\rho + \varepsilon)\mu(B)\}.$$

Tenemos entonces que $\inf\{r : \overline{B}(a,r) \in \mathscr{C}\} = 0$ de manera que podemos aplicar el lema de recubrimiento de Besicovitch por medio del Corolario 1.3.3 para obtener una familia numerable $\mathcal{G}$ de bolas disjuntas de $\mathscr{C}$ tal que

$$\nu\left(A \setminus \bigcup_{B\in\mathcal{G}} B\right) = 0.$$

Entonces

$$\nu(A) \leq \sum_{B\in\mathcal{G}} \nu(B) \leq (\rho+\varepsilon) \sum_{B\in\mathcal{G}} \mu(B) \leq (\rho+\varepsilon)\mu(U).$$

Esta desigualdad es válida para todo conjunto abierto U que contiene A, de manera que, dado que las medidas utilizadas son medidas de Radon (es decir que se puede aproximar la medida de un conjunto por medio de abiertos), se obtiene

$$\nu(A) \leq (\rho+\varepsilon)\mu(A).$$

Como el real ε era abritrario se obtiene el primer punto. El segundo punto se verifica de forma totalmente similar de manera que los detalles son dejados al lector. ■

Con este resultado pasamos a estudiar cuándo la derivada de una medida con respecto a otra existe. Veremos pues que la noción de medida de Radon es el concepto necesario para asegurar la existencia de este tipo de derivada y obtendremos además que la función resultante es una función medible.

Teorema 1.3.4 (Existencia de Derivadas) *Sean μ y ν dos medidas de Radon definidas sobre $\mathbb{R}^n$. Entonces la derivada $D_\mu(\nu)$ existe y es finita en μ-casi todas partes. Se tiene además que la función derivada $D_\mu(\nu)$ es una función medible.*

En este enunciado se indica claramente que la medida μ con respecto a la cual se deriva imprime su huella en el resultado, pues los objetos obtenidos existen en casi todas partes con respecto a esta medida μ.

Demostración. Podemos suponer sin pérdida de generalidad que $\nu(\mathbb{R}^n) < +\infty$ y $\mu(\mathbb{R}^n) < +\infty$. Mostremos para empezar que $D_\mu(\nu)$ existe y es finita en μ-casi todas partes.
Sea pues

$$I = \{x \in \mathbb{R}^n : \overline{D}_\mu(\nu)(x) = +\infty\}$$

y para todo $0 < a < b$ definimos el conjunto

$$R(a,b) = \{x \in \mathbb{R}^n : \underline{D}_\mu(\nu)(x) < a < b < \overline{D}_\mu(\nu)(x) < +\infty\}.$$

Observamos que para todo $\rho > 0$, se tiene la inclusión $I \subset \{x \in \mathbb{R}^n : \overline{D}_\mu(\nu)(x) \geq \rho\}$ de donde se obtiene por el Lema 1.3.6 que

$$\mu(I) \leq \frac{1}{\rho}\nu(I).$$

Como la cantidad $\nu(I)$ es finita pues hemos supuesto que $\nu(\mathbb{R}^n) < +\infty$, haciendo tender $\rho \to +\infty$ se deduce que $\mu(I) = 0$ y por lo tanto se obtiene que $\overline{D}_\mu(\nu)$ is finita en μ-casi todas partes.

Aplicando una vez más el Lema 1.3.6 podemos ver que

$$b\mu(R(a,b)) \leq \nu(R(a,b)) \leq a\mu(R(a,b)),$$

lo que implica que $\mu(R(a,b)) = 0$ pues habíamos fijamos que $b > a$. Vemos además que

$$\{x \in \mathbb{R}^n : \underline{D}_\mu(\nu)(x) < \overline{D}_\mu(\nu)(x) < +\infty\} \subset \bigcup_{\substack{0<a<b \\ a,b \text{ racionales}}} R(a,b)$$

de donde se obtiene que los valores de la función $D_\mu(\nu)$ existen y son finitos en μ-casi todas partes.

Para terminar la demostración del teorema debemos verificar que la función $D_\mu(\nu)$ es medible y para ello procedemos como sigue. Empezamos verificando que, para todo $x \in \mathbb{R}^n$ y todo $r > 0$, se tiene que

$$\limsup_{y \to x} \mu(\overline{B}(y,r)) \leq \mu(\overline{B}(x,r)) \tag{1.51}$$

y que se tiene la misma relación para la medida ν.

En efecto, sea $(y_n)_{n \in \mathbb{N}} \in \mathbb{R}^n$ una sucesión tal que $y_k \to x$. Definimos entonces las funciones $f_k = \mathbb{1}_{\overline{B}(y_k,r)}$ y $f = \mathbb{1}_{\overline{B}(x,r)}$ y vemos que se tiene

$$\limsup_{k \to +\infty} f_k \leq f \quad \text{y} \quad \liminf_{k \to +\infty} (1 - f_k) \geq (1 - f).$$

En este punto podemos aplicar el Lema de Fatou para obtener

$$\int_{\overline{B}(x,2r)} (1-f) d\mu \leq \int_{\overline{B}(x,2r)} \liminf_{k \to +\infty} (1 - f_k) d\mu \leq \liminf_{k \to +\infty} \int_{\overline{B}(x,2r)} (1 - f_k) d\mu$$

es decir:

$$\mu(\overline{B}(x,2r)) - \mu(\overline{B}(x,r)) \leq \liminf_{k \to +\infty} \Big(\mu(\overline{B}(x,2r)) - \mu(\overline{B}(y_k,r)) \Big).$$

Pero dado que la medida μ es una medida de Radon, se tiene que $\mu(\overline{B}(x,2r)) < +\infty$ de donde se obtiene (1.51).

Gracias a la fórmula (1.51) para las medidas μ y ν, vemos que para todo $r > 0$ se tiene que las funciones $x \longmapsto \mu(\overline{B}(x,r))$ y $x \longmapsto \nu(\overline{B}(x,r))$ son funciones semi-continuas superiormente y por lo tanto son funciones medibles con respecto a la σ-álgebra boreliana de $\mathbb{R}^n$. Tenemos entonces que, para todo $r > 0$, la función definida por

$$f_r(x) = \begin{cases} \frac{\nu(\overline{B}(x,r))}{\mu(\overline{B}(x,r))} & \text{si } \mu(\overline{B}(x,r)) > 0 \text{ para todo } r > 0. \\ +\infty & \text{si } \mu(\overline{B}(x,r)) = 0. \end{cases}$$

es una función medible. Pero vemos que se tiene $D_\mu(\nu) = \lim_{r \to 0} f_r = \lim_{k \to 0} f_{\frac{1}{k}}$ en μ-casi todas partes, de manera que se obtiene que la función derivada $D_\mu(\nu)$ es medible. ■

Una vez que se dispone de esta derivada entre medidas, pasamos al problema fundamental que consiste en relacionar el proceso de derivación con el proceso de integración. Veremos entonces, con los resultados a continuación, en qué sentido se puede recuperar una medida al integrar su derivada y recíprocamente, cómo se puede obtener una medida al derivar su integral.

Teorema 1.3.5 (Teorema Fundamental del Cálculo para medidas de Radon) *Sean μ y ν dos medidas de Radon definidas sobre el espacio medible $(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n))$. Si $\nu \ll \mu$ entonces se tiene la relación*

$$\nu(A) = \int_A D_\mu(\nu)(x) d\mu(x) \tag{1.52}$$

para todo conjunto boreliano A.

Demostración. Para empezar definimos los siguientes conjuntos

$$Z = \{x \in \mathbb{R}^n : D_\mu(\nu)(x) = 0\}, \quad I = \{x \in \mathbb{R}^n : D_\mu(\nu)(x) = +\infty\}.$$

Por los resultados anteriores sabemos que estos dos conjuntos son medibles y en la demostración del Teorema 1.3.4 hemos visto que $\mu(I) = 0$ y que entonces se tiene que $\nu(I) = 0$ puesto que por hipótesis se tiene $\nu \ll \mu$. El Lema 1.3.6 impica que $\nu(Z) \leq \rho\mu(Z)$ para todo $\rho > 0$, de donde se deduce que $\nu(Z) = 0$ y entonces podemos escribir

$$\nu(Z) = 0 = \int_Z D\mu(\nu)d\mu \quad \text{y} \quad \nu(I) = 0 = \int_I D\mu(\nu)d\mu.$$

Una vez que hemos tratado estos dos conjuntos, pasemos al caso general.

Sea pues A un conjunto boreliano y sea $1 < t < +\infty$ un real. Definimos entonces los conjuntos A_n, con $n \in \mathbb{Z}$, como

$$A_n = A \cap \{x \in \mathbb{R}^n : t^n \leq D_\mu(\nu)(x) < t^{n+1}\},$$

y, dado que la función $D_\mu(\nu)$ es medible, se tiene que estos conjuntos son también medibles. Tenemos entonces

$$A \setminus \bigcup_{n\in\mathbb{Z}} A_n \subset Z \cup I \cup \{x \in \mathbb{R}^n : \overline{D}_\mu(\nu)(x) \neq \underline{D}_\mu(\nu)(x)\},$$

por lo tanto

$$\mu\left(A \setminus \bigcup_{n\in\mathbb{Z}} A_n\right) = \nu\left(A \setminus \bigcup_{n\in\mathbb{Z}} A_n\right) = 0.$$

Podemos entonces escribir, usando el Lema 1.3.6

$$\nu(A) = \sum_{n\in\mathbb{Z}} \nu(A_n) \leq \sum_{n\in\mathbb{Z}} t^{n+1}\mu(A_n)$$

Es decir

$$\nu(A) \leq t\sum_{n\in\mathbb{Z}} t^n\mu(A_n) \leq t\sum_{n\in\mathbb{Z}} \int_{A_n} D_\mu(\nu)d\mu = t\int_A D_\mu(\nu)d\mu.$$

De forma simétrica usando el mismo Lema 1.3.6 podemos también escribir:

$$\nu(A) = \sum_{n\in\mathbb{Z}} \nu(A_n) \geq \sum_{n\in\mathbb{Z}} t^m\mu(A_n) = \frac{1}{t}\sum_{n\in\mathbb{Z}} t^{n+1}\mu(A_n) \geq \frac{1}{t}\sum_{n\in\mathbb{Z}} \int_{A_n} D_\mu(\nu)d\mu = \frac{1}{t}\int_A D_\mu(\nu)d\mu$$

De esta manera tenemos las acotaciones siguientes para todo $1 < t < +\infty$

$$\frac{1}{t}\int_A D_\mu(\nu)d\mu \leq \nu(A) \leq t\int_A D_\mu(\nu)d\mu,$$

de modo que para terminar la demostración es suficiente hacer tender $t \to 1^+$. ■

Sabemos desde los primeros cursos de cálculo, que la derivada de funciones constantes es igual a cero. ¿Cuál es el equivalente de esta idea en el marco de las derivadas de medidas? Para estudiar esta noción utilizaremos el Teorema de descomposición de Lebesgue 1.2.4 dado en la página 37.

Proposición 1.3.5 *Sea μ y ν dos medidas de Radon definidas sobre el espacio medible* $(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n))$. *Sea*

$$\nu = \nu_a + \nu_s$$

la descomposición de Lebesgue de la medida ν con respecto a la medida μ. Entonces se tiene

$$D_\mu(\nu) = D_\mu(\nu_a) \quad \text{y} \quad D_\mu(\nu_s) = 0, \quad \text{en } \mu\text{-casi todas partes.}$$

Y por lo tanto podemos escribir

$$\nu(A) = \int_A D_\mu(x) d\mu(x) + \nu_s(A)$$

para todo conjunto boreliano A.

La diferencia entre este resultado y la expresión (1.52) reside en la hipótesis hecha en el Teorema 1.3.5: ahí se supone que $\nu \ll \mu$ de manera que la parte singular de la medida ν no aparece.

Prueba. Podemos suponer sin pérdida de generalidad que las medidas μ y ν son finitas, pues de lo contrario bastaría restringirse a los conjuntos compactos. Definimos luego el conjunto E por

$$E = \{A \subset \mathbb{R}^n : A \in \mathcal{B}or(\mathbb{R}^n),\ \mu(\mathbb{R}^n \setminus A) = 0\}$$

y consideramos los conjuntos $B_k \in E$, con $k \geq 1$ tales que $\nu(B_k) = \inf\limits_{A \in E} \nu(A) + \frac{1}{k}$. Si notamos $B = \bigcap_{k \geq 1} B_k$, dado que

$$\nu(\mathbb{R}^n \setminus B) \leq \sum_{k \geq 1} \nu(\mathbb{R}^n \setminus B_k) = 0$$

se obtiene que $B \in E$ y además que $\nu(B) = \inf\limits_{A \in E} \nu(A)$. Una vez que disponemos de este conjunto B, definimos para $\rho > 0$ el conjunto $C = \{x \in B : D_\mu(\nu_s)(x) \geq \rho\}$. Entonces, por el Lema 1.3.6 tenemos

$$\rho\mu(C) \leq \nu_s(C) = 0$$

de donde se deduce que $D_\mu(\nu_s) = 0$ en μ-casi todas partes, lo que nos permite escribir

$$\nu(A) = \int_A D_\mu(x) d\mu(x) + \nu_s(A)$$

■

Con este resultado hemos terminado nuestra exposición sobre la derivada de medidas.

1.3.3. Funciones de variación acotada

Las funciones de variación acotada juegan un rol fundamental en varias ramas de las matemáticas y poseen conecciones muy particulares con los temas que han sido estudiados en este capítulo. En esta subsección vamos a empezar construyendo una biyección entre el conjunto de todas las medidas positivas finitas definidas sobre $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y el conjunto de funciones que verifican ciertas propiedades. Luego generalizaremos este resultado para dar una caracterización de las medidas con signo finitas sobre $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ que son absolutamente continuas con respecto a la medida de Lebesgue.

Empecemos pues por un resultado que nos proporciona algunas propiedades que serán constantemente utilizadas en lo que sigue.

Proposición 1.3.6 *Sea μ una medida finita definida sobre el espacio medible $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y sea una función $F_\mu : \mathbb{R} \longrightarrow \mathbb{R}$ definida por*

$$F_\mu(x) = \mu(]-\infty, x]) = \int_{-\infty}^{x} d\mu. \tag{1.53}$$

Entonces F_μ es una función acotada, creciente, continua por la derecha y verifica $\lim\limits_{x \to -\infty} F_\mu(x) = 0$.

Prueba. Dado que la medida μ es finita, se tiene sin problema que la función F_μ es acotada pues se tiene $0 \leq \mu(]-,\infty,x]) \leq \mu(\mathbb{R}) < +\infty$. Además se obtiene, para todo $x \leq y$ que $\mu(]-\infty,x]) \leq \mu(]-\infty,y])$, de donde se deduce el crecimiento de la función F_μ. Sea ahora $x \in \mathbb{R}$ y sea $(x_n)_{n\in\mathbb{N}}$ la sucesión definida por $x_n = x + \frac{1}{n}$. Tenemos entonces que $]-\infty,x] = \bigcap_{n\in\mathbb{N}}]-\infty,x_n]$, de manera que la continuidad de las medidas implica que se tiene $F_\mu(x) = \lim\limits_{n\to+\infty} F_\mu(x_n)$ y a partir de esto se deduce la continuidad por la derecha de la función F_μ. Razonando de la misma manera se obtiene que $\lim\limits_{x\to-\infty} F_\mu(x) = 0$. ∎

Este resultado nos proporciona una serie de informaciones sobre la función F_μ definida por medio de la expresión (1.53). Veamos ahora otras propiedades de esta función: dado que el conjunto $]a,b]$ se escribe como la diferencia de los intervalos $]-\infty,b]$ y $]-\infty,a]$, se tiene con esta proposición que

$$\mu(]a,b]) = F_\mu(b) - F_\mu(a). \tag{1.54}$$

Esta expresión será muy utilizada en las páginas siguientes y conviene estudiarla con un poco más de detalle. Dado que F_μ es acotada y creciente, el límite $\lim\limits_{t\to x^-} F_\mu(t)$ existe para todo $x \in \mathbb{R}$. Vemos entonces que este límite es igual a $\sup\{F_\mu(t) : t < x\}$ y será notado por $F_\mu(x^-)$. Ahora, si $(a_n)_{n\in\mathbb{N}}$ es una sucesión de creciente y que converge hacia un número real b, si aplicamos la fórmula (1.54) a cada intervalo de la forma $]a_n,b]$, entonces se tiene

$$\mu(\{b\}) = F_\mu(b) - F_\mu(b^-). \tag{1.55}$$

A partir de esto se deduce que la función F_μ es continua en el punto b si $\mu(\{b\}) = 0$ y es discontinua en este punto b, con una discontinuidad de tamaño $\mu(\{b\})$ si $\mu(\{b\}) \neq 0$. De esta manera se obtiene que la medida μ verifica $\mu(\{x\}) = 0$ para todo $x \in \mathbb{R}^n$ si y solo si la función F_μ es continua.

Con el teorema a continuación estudiamos la recíproca de la Proposición 1.3.6:

Teorema 1.3.6 *Para cada función acotada, creciente y continua por la derecha $F : \mathbb{R} \longrightarrow \mathbb{R}$ que verifica $\lim\limits_{x\to-\infty} F(x) = 0$, existe una única medida positiva μ definida sobre el espacio medible $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ tal que se tiene $F(x) = \mu(]-\infty,x])$ para todo $x \in \mathbb{R}$.*

Demostración. Sea $F : \mathbb{R} \longrightarrow \mathbb{R}$ una función acotada, creciente y continua por la derecha que verifica $\lim\limits_{x\to-\infty} F(x) = 0$. Empezamos construyendo la medida buscada y para ello consideramos la aplicación $\overline{\mu} : \mathcal{P}(\mathbb{R}) \longrightarrow \mathbb{R}$ exigiendo que $\overline{\mu}(A)$ sea el ínfimo de las sumas $\sum_{n=1}^{+\infty}(F(b_n) - F(a_n))$ donde a_n y b_n son los extremos de los intervalos $]a_n,b_n]$ que recubren A en el sentido que $A \subset \bigcup_{n\geq1}]a_n,b_n]$. Por la teoría desarrollada en el primer volumen se puede ver sin mayor problema que la aplicación $\overline{\mu}$ es una medida exterior.

Verifiquemos ahora que se tiene $\overline{\mu}(]-\infty,x]) = F(x)$ para todo $x \in \mathbb{R}$. Notemos que se tiene la desigualdad $\overline{\mu}(]-\infty,x]) \leq F(x)$ puesto que el intervalo $]-\infty,x]$ puede ser recubierto por la sucesión de intervalos $(]x-n,x-n+1])_{n\in\mathbb{N}}$ para los cuales se tiene

$$\sum_{n=1}^{+\infty} F(x-n+1) - F(x-n) = F(x).$$

Debemos pues concentrarnos en la desigualdad recíproca. Sea $\varepsilon > 0$ un real y sea $(]a_n,b_n])_{n\in\mathbb{N}}$ una sucesión de intervalos tal que $]-\infty,x] \subset \bigcup_{n\in\mathbb{N}}]a_n,b_n]$. Usamos el hecho que $\lim\limits_{x\to-\infty} F(x) = 0$ para escoger un real t tal que $t < x$ y tal que $F(t) < \varepsilon$. Ahora, para todo $n \in \mathbb{N}$, usando la continuidad por la derecha de F podemos fijar un real positivo δ_n tal que $F(b_n + \delta_n) < F(b_n) + \frac{\varepsilon}{2^n}$. Entonces el intervalo $[t,x]$ es compacto, cada intervalo $]a_n, b_n + \delta_n[$ es abierto y se tiene $[t,x] \subset \bigcup_{n\in\mathbb{N}}]a_n, b_n+\delta_n[$ y $\sum_{n\in\mathbb{N}} F(b_n+\delta_n) - F(a_n) \leq \sum_{n\in\mathbb{N}} F(b_n - F(a_n) + \varepsilon$. La compacidad de $[t,x]$ implica que existe un entero N tal que $[t,x] \subset \bigcup_{n=1}^{N}]a_n, b_n + \delta_n[$ y se obtiene que $]t,x]$ es la unión de una familia finita de intervalos disjuntos $]c_j, d_j]$ cada uno de los cuales se encuentra contenido en algún intervalo $]a_n, b_n+\delta_n]$. Se deduce de esto que

$$F(x) - F(t) = \sum_j F(d_j) - F(c_j) \leq \sum_{n\in\mathbb{N}} F(b_n + \delta_n) - F(a_n)$$

y por lo tanto

$$F(x) - \varepsilon \leq \sum_{n\in\mathbb{N}} F(b_n) - F(a_n) + \varepsilon.$$

Pero dado que el real ε y los conjuntos $]a_n, b_n]$ son cualquiera, se obtiene la desigualdad $F(x) \leq \overline{\mu}(]-\infty, x])$. De donde se deduce la identidad $F(x) = \overline{\mu}(]-\infty, x])$.

Sea ahora μ la restricción de $\overline{\mu}$ a los borelianos de $\mathbb{R}$. Tenemos entonces que μ es una medida que verifica $\mu(]-\infty, x]) = F(x)$ para todo $x \in \mathbb{R}$. Dado que F es una función acotada y que se tiene $\mu(\mathbb{R}) = \lim\limits_{n\to+\infty} \mu(]-\infty, n]) = \lim\limits_{n\to+\infty} F(n)$, de obtiene que la medida μ es finita.

Pasemos a verificar que esta medida es única. Si μ y ν son dos medidas finitas tales que se tiene $\mu(\mathbb{R}) = \nu(\mathbb{R})$ y tales que se tiene $\mu(]-\infty, a]) = \nu(]-\infty, a])$ para todo $a \in \mathbb{R}$, dado que $\mathbb{R}$ es la unión creciente de una sucesión de intervalos de la forma $]-\infty, a]$, se tiene tiene que esta familia de intervalos es un π-sistema que genera $\mathcal{B}or(\mathbb{R})$, de manera que se obtiene la igualdad de medidas $\mu = \nu$ y por lo tanto la unicidad de este tipo de medidas. ∎

Con estos resultados hemos construido una biyección entre el conjunto de todas las medidas positivas finitas definidas sobre el espacio medible $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y el conjunto de funciones acotadas, crecientes y continuas por la derecha que tienden hacia 0 en $-\infty$. Este es un primer paso y ahora deseamos pasar a estudiar una caracterización similar para el conjunto de medidas finitas con signo que son absolutamente continuas con respecto a la medida de Lebesgue.

Para ello necesitaremos el siguiente concepto.

Definición 1.3.8 (Variación de una función, funciones de variación acotada) *Sea una función $f: \mathbb{R} \longrightarrow \mathbb{R}$, sea $[a,b]$ un intervalo y sea $\mathcal{S}$ la colección de todas sucesiones finitas $(t_i)_{0\leq i\leq n}$ tales que $a \leq t_0 < t_1 < \cdots < t_n \leq b$.*

1) *La* variación de la función f sobre el intervalo $[a,b]$ *será notada por $Var_f[a,b]$ y está definida por la expresión*

$$Var_f[a,b] = \sup\left\{\sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| : (t_i)_{0\leq i\leq n} \in \mathcal{S}\right\}. \tag{1.56}$$

La variación de f sobre los intervalos de tipo $]-\infty, a]$ o $]-\infty, +\infty[$ se definen de forma totalmente similar.

La función f será dicha de variación acotada *sobre el intervalo $[a,b]$ si $Var_f[a,b] < +\infty$. Por simplicidad notaremos $Var_f = Var_f]-\infty, +\infty[$.*

2) *Si $f: \mathbb{R} \longrightarrow \mathbb{R}$ es una función de variación acotada sobre $\mathbb{R}$, podemos definir la* variación *de f como la función $V_f: \mathbb{R} \longrightarrow \mathbb{R}$ determinada por $V_f(x) = Var_f]-\infty, x]$.*

Esta noción de variación que acabamos de presentar permite capturar los "saltos" de una función demos unos ejemplos.

(i) Vemos sin problema que la función $\varphi = \mathbb{1}_{[0,1]}$ es de variación acotada y que se tiene $Var_\varphi = 2$.

(ii) Veamos otro ejemplo importante: sea μ una medida finita con signo definida sobre el espacio medible $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y consideremos la función $F_\mu(x) = \mu(]-\infty, x])$, entonces esta función es una función de variación acotada. En efecto, si $(t_i)_{0\leq i\leq n}$ es una sucesión creciente de reales, entonces

$$\sum_{i=1}^{n} |F_\mu(t_i) - F_\mu(t_{i-1})| = \sum_{i=1}^{n} |\mu(]t_{i-1}, t_i])| \leq |\mu|(\mathbb{R}) = \|\mu\|_{VT} < +\infty$$

puesto que la medida con signo μ es finita y por lo tanto su variación total $\|\mu\|_{VT}$ es finita. A partir de esto se deduce que $Var_{F_\mu} \leq \|\mu\|_{VT} < +\infty$ y se obtiene que la función F_μ es de variación acotada.

(iii) Por el contrario, si consideramos la función f definida en el intervalo $]-1,1[$ por $f(x) = x^2\cos^2(\pi/x^2)$ si $x \neq 0$ y por $f(0) = 0$, entonces vemos que esta función no es de variación acotada puesto que las oscilaciones de la función coseno hacen que no sea posible controlar las variaciones de esta función. Ver más detalles en los ejercicios 1.5 y 1.6.

Presentamos ahora algunas propiedades de las funciones de variación acotada que serán de utilidad en lo que sigue y que se deducen directamente de su definición.

Proposición 1.3.7 *Sea $f : \mathbb{R} \longrightarrow \mathbb{R}$ una función de variación acotada, entonces*

1) f es una función acotada,

2) si $-\infty < a < b < +\infty$ entonces $Var_f]-\infty, b] = Var_f]-\infty, a] + Var_f[a, b]$,

3) si $b \in \mathbb{R}$, entonces $Var_f]-\infty, b] = \lim\limits_{a\to-\infty} Var_f]a, b]$.

Prueba. El primer punto es inmediato pues si la función no es acotada, entonces la variación de f tendería a $+\infty$ por construcción. El segundo punto es igual de inmediato pues refleja la linealidad en la definición de la variación de una función. Para el último punto consideramos $\varepsilon > 0$ un real y escogemos una sucesión creciente $(t_i)_{0\leq i\leq n}$ de números reales que pertenecen al intervalo $]-\infty, b]$ y que verifican

$$\sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| > Var_f]-\infty, b] - \varepsilon.$$

Se tienen entonces las desigualdades $Var_f]-\infty, b] - \varepsilon < Var_f[a, b] \leq Var_f]-\infty, b]$ para todo $a \leq t_0$; de donde se obtiene el resultado. ■

A partir de esta proposición se deducen los siguientes puntos sobre la función de variación.

Proposición 1.3.8 *Sea f una función de variación acotada. Entonces*

1) la función V_f es una función acotada y creciente,

2) la función V_f se anula en $-\infty$,

3) si f es continua por la derecha, entonces la función V_f es continua por la derecha.

Prueba. El primer punto es inmediato por definición. Los puntos 2) y 3) de la proposición anterior justifican las identidades siguientes para $b > x$:

$$\lim_{x\to-\infty} V_f(x) = \lim_{x\to-\infty} Var_f]-\infty, x] = \lim_{x\to-\infty}\big(Var_f]-\infty, b]-Var_f[x, b]\big) = Var_f]-\infty, b]-Var_f]-\infty, b] = 0,$$

de manera que se obtiene el punto 2). Razonando de manera similar se obtiene el punto 3): en efecto, si la función f es continua por la derecha, entonces si $a < c$ se tiene $Var_f[a, c] = \lim_{b\to a+} Var_f[b, c]$. ■

Proposición 1.3.9 *Sea $f : \mathbb{R} \longrightarrow \mathbb{R}$ una función de variación acotada. Entonces existen funciones crecientes f_1 y f_2 tales que $f = f_1 - f_2$. Además, si f es continua por la derecha, entonces las funciones f_1 y f_2 son continuas por la derecha y si f se anula en $-\infty$, entonces f_1 y f_2 se anulan en $-\infty$.*

Prueba. Basta considerar $f_1 = (V_f + f)/2$ y $f_2 = (V_f - f)/2$. ■

Podemos ahora construir la biyección anunciada.

Teorema 1.3.7 *Sea μ una medida con signo finita definida sobre $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y sea $F_\mu : \mathbb{R} \longrightarrow \mathbb{R}$ una función definida por la relación*

$$F_\mu(x) = \mu(]-\infty, x]).$$

Entonces esta relación define una biyección $\mu \longmapsto F_u$ entre el conjunto de medidas con signo finitas definidas sobre $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y el conjunto de funciones continuas por la derecha que son de variación acotada y que se anulan en $-\infty$.

Demostración. Sabemos por los resultados anteriores (ver el ejemplo (ii) de la página anterior) que la función F_μ es una función continua por la derecha de variación finita que se anula en $-\infty$.

Debemos verificar entonces que si F es una función con estas características, entonces existe una medida μ tal que $F(x) = \mu(]-\infty, x])$. Por la Proposición 1.3.9 sabemos que la función F puede descomponerse como $F = F_1 - F_2$ donde tanto F_1 como F_2 son funciones crecientes, acotadas, continuas por la derecha y tales que se anulan en $-\infty$. En este punto invocamos el Teorema 1.3.6 para obtener la existencia de dos medidas positivas μ_1 y μ_2 tales que $F_1(x) = \mu_1(]-\infty, x])$ y $F_2(x) = \mu_2(]-\infty, x])$. Tenemos pues que $F(x) = F_1(x) - F_2(x) = \mu_1(]-\infty, x]) - \mu_2(]-\infty, x]) = \mu(]-\infty, x])$ donde $\mu = \mu_1 - \mu_2$. Finalmente, debemos verificar la inyectividad de la aplicación $\mu \longmapsto F_u$ y para ello vemos que si μ y ν son medidas con signo finitas tales que $F_\mu = F_\nu$, y si $\mu = \mu^+ - \mu^-$ y $\nu = \nu^+ - \nu^-$ son sus descomposiciones de Jordan, entonces $F_{\mu^+} - F_{\mu^-} = F_{\nu^+} - F_{\nu^-}$ y esto implica $F_{\mu^+} + F_{\nu^-} = F_{\nu^+} + F_{\mu^-}$. Aplicando el Teorema 1.3.6 se obtiene que $\mu^+ + \nu^- = \nu^+ + \mu^-$ y por lo tanto que $\mu = \nu$. ■

Este importante resultado nos indica que *todas* las medidas con signo finitas pueden expresarse por medio de la expresión $F_\mu(x) = \mu(]-\infty, x])$ en donde F es una función de variación acotada, lo que permite dar una caracterización importante de tales medidas.

Para terminar esta sección presentamos algunas propiedades interesantes en donde intervienen los conceptos de funciones de variación acotada y de continuidad absoluta de medidas. En particular, con la definición a continuación, vamos a ver cuál es el análogo de la noción de continuidad absoluta para las funciones y veremos cómo se relaciona esta idea con los resultados anteriores.

Definición 1.3.9 (Funciones absolutamente continuas) *Una función $f : \mathbb{R} \longrightarrow \mathbb{R}$ es* absolutamente continua *si para todo $\varepsilon > 0$ existe un número $\delta > 0$ tal que se tenga la estimación*

$$\sum_{i=1}^{n} |f(t_i) - f(s_i)| < \varepsilon$$

para toda sucesión de intervalos abiertos disjuntos $(]s_i, t_i[)_{1\leq i\leq n}$ *tales que* $\sum_{i=1}^{n}(t_i - s_i) < \delta$.

La noción de continuidad absoluta es un concepto más fuerte que la continuidad y se puede ver sin mayor problema que un ejemplo de función absolutamente continua es la aplicación $x \longmapsto \sqrt{x}$ definida sobre $[0, +\infty[$. En particular tenemos la siguiente proposición.

Proposición 1.3.10 *Sea* $f : [a, b] \longrightarrow \mathbb{R}$ *una función absolutamente continua, entonces* f *es una función uniformemente continua sobre el intervalo* $[a, b]$ *y además la función* f *es de variación acotada sobre este mismo intervalo.*

Prueba. Empecemos con la continuidad uniforme. Sea f una función absolutamente continua, sea $\varepsilon > 0$ un número real y sea $\delta > 0$ el número asociado a ε en la Definición 1.3.9. Si $x, y \in [a, b]$ verifican $|x - y| < \delta$, entonces por definición de continuidad absoluta se tiene que $|f(x) - f(y)| < \varepsilon$ de donde se obtiene la continuidad uniforme de la función f sobre el intervalo $[a, b]$.

Para verificar que la función es de variación acotada, fijemos $\varepsilon > 0$ y $\delta > 0$ de la misma manera y consideremos $N \in \mathbb{N}$ un entero suficientemente grande tal que $\frac{1}{N}(b - a) < \delta$. Sea $\mathcal{S}_0$ la partición del intervalo $[a, b]$ por medio de N intervalos $I_1, ..., I_N$ de igual longitud $\frac{1}{N}(b-a)$ y sea ahora $\mathcal{S}$ una partición cualquiera del intervalo $[a, b]$. Consideramos la partición $\mathcal{S} \cup \mathcal{S}_0$ que divide cada intervalo I_k en varios subintervalos $I_{k,1}, ..., I_{k,n(k)}$ con $n(k) \in \mathbb{N}$ para todo $k = 1, ..., N$, es decir que consideramos intervalos del tipo $I_{k,j} = [a_{k,j}, b_{k,j}]$ con $j = 1, ..., n(k)$ y $k = 1, ..., N$. Ahora, dado que $\sum_{j=1}^{n(k)} |b_{k,j} - a_{k,j}| = \frac{1}{N}(b - a) < \delta$ implica $\sum_{j=1}^{n(k)} |f(b_{k,j}) - f(a_{k,j})| < \varepsilon$ por definición de función absolutamente continua, tenemos entonces que

$$\sum_{k=1}^{N}\sum_{j=1}^{n(k)} |f(b_{k,j}) - f(a_{k,j})| \leq N\varepsilon < +\infty$$

y esta estimación es uniforme para toda partición $\mathcal{S}$ de donde se deduce que $Var_f[a, b] \leq N\varepsilon < +\infty$ y se obtiene que la función f es de variación acotada. ■

Indiquemos que la continuidad absoluta es un concepto más restrictivo que la noción de variación acotada, en el sentido que no se tiene la recíproca de este resultado. Ver más detalles en el Ejercicio 1.5.

Lema 1.3.7 *Si* $f : \mathbb{R} \longrightarrow \mathbb{R}$ *es una función absolutamente continua entonces la función* V_f *es absolutamente continua.*

Prueba. Sea $\varepsilon > 0$ un número real, usando la continuidad absoluta de f podemos fijar un número real $\delta > 0$ correspondiente. Si $(]s_i, t_i[)_{1\leq i\leq n}$ es una sucesión finita de intervalos abiertos tales que $\sum_{i=1}^{n}(t_i - s_i) < \delta$, entonces toda sucesión finita $(]u_j, v_j[)_{1\leq j\leq n}$ de intervalos abiertos contenidos en $\bigcup_{i=1}^{n}]t_i, s_i[$ verifica $\sum_{j=1}^{n}(v_j - u_j) < \delta$ y por lo tanto verifica $\sum_{j=1}^{n} |f(v_j) - f(u_j)| < \varepsilon$. Dado que la sucesión $(]u_j, v_j[)_{1\leq j\leq n}$ puede ser escogida de tal manera que la cantidad $\sum_{j=1}^{n} |f(v_j) - f(u_j)| < \varepsilon$ sea arbitrariamente cercana a $\sum_{j=1}^{n} Var_f[s_i, t_i]$, se obtiene

$$\sum_{i=1}^{n} |V_f(t_i) - V_f(s_i)| = \sum_{i=1}^{n} Var_f[s_i, t_i] \leq \varepsilon$$

y a partir de esto se obtiene la continuidad absoluta de V_f. ■

El primer resultado que presentamos que relaciona las nociones de continuidad absoluta de las medidas con la continuidad absoluta de las funciones es el siguiente.

Proposición 1.3.11 *Sea μ una medida con signo finita definida sobre $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y sea $F_\mu : \mathbb{R} \longrightarrow \mathbb{R}$ definida por $F_\mu(x) = \mu(]-\infty, x])$. Entonces F_μ es absolutamente continua si y solo si la medida μ es absolutamente continua con respecto a la medida de Lebesgue.*

Prueba. Supongamos primero que la medida μ es absolutamente continua con respecto a la medida de Lebesgue. Sea $\varepsilon > 0$ un real positivo, usamos la Proposición 1.2.2 dada en la página 26 para obtener un número real $\delta > 0$ tal que se tenga $|\mu|(A) < \varepsilon$ siempre y cuando A es un conjunto boreliano que verifica $\lambda(A) < \delta$. Si $(]s_i, t_i[)_{1\leq i\leq n}$ es una sucesión finita de intervalos abiertos disjuntos tales que $\sum_i (t_i - s_i) < \delta$, entonces $\lambda(\bigcup_i]s_i, t_i]) < \delta$ y entonces

$$\sum_{i=1}^{n} |F_\mu(t_i) - F_\mu(s_i)| = \sum_i |\mu(]s_i, t_i])| \leq |\mu|(\bigcup_i]s_i, t_i]) < \varepsilon.$$

A partir de esto se deduce que F_μ es absolutamente continua.

Supongamos ahora que F_μ es absolutamente continua. Entonces V_{F_μ} es absolutamente continua y entonces las funciones $F_1 = (V_{F_\mu} + F_\mu)/2$ y $F_2 = (V_{F_\mu} - F_\mu)/2$ son absolutamente continuas. Sean μ_1 y μ_2 las medidas positivas finitas definidas sobre $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ que corresponden a las funciones F_1 y F_2. Entonces se tiene que $\mu = \mu_1 - \mu_2$ y lo único que debemos mostrar es que $\mu \ll \lambda$. Sean pues ε, δ dos números reales positivos tales que se tiene

$$\sum_i |F_1(t_i) - F_1(s_i)| < \varepsilon$$

siempre y cuando $(]s_i, t_i[)_{1\leq i\leq n}$ es una sucesión finita de intervalos abiertos disjuntos tales que $\sum_i (t_i - s_i) < \delta$. Si A es un subconjunto boreliano de $\mathbb{R}$ tal que $\lambda(A) < \delta$, y utilizando la regularidad de la medida de Lebesgue podemos escoger un conjunto abierto U que contiene A y tal que $\lambda(U) < \delta$. Entonces U es la unión de una sucesión $(]s_i, t_i[)_{1\leq i\leq n}$ de intervalos abiertos disjuntos y a raíz de esto se obtiene que

$$\mu_1\left(\bigcup_{i=1}^{n}]s_i, t_i[\right) = \sum_{i=1}^{n} (F_1(t_i) - F_1(s_i)) < \varepsilon$$

para todo n. Entonces se tiene $\mu_1(U) = \mu_1(\bigcup_{i=1}^{+\infty}]s_i, t_i[) \leq \varepsilon$ y por lo tanto $\mu_1(A) \leq \varepsilon$. A partir de esto se obtiene la continuidad absoluta de la medida μ_1. El caso de la medida μ_2 es totalmente similar.■

Finalmente, la proposición a continuación nos proporciona un resultado de orden general que relaciona todos los conceptos usados anteriormente.

Proposición 1.3.12 *Las funciones $F : \mathbb{R} \longrightarrow \mathbb{R}$ que puede ser definidas por medio de la expresión*

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

con $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}or(\mathbb{R}), \lambda, \mathbb{R})$ son exactamente las funciones que son de variación acotada que son absolutamente continuas y que se anulan en $-\infty$.

Prueba. Supongamos que $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}or(\mathbb{R}), \lambda, \mathbb{R})$ y que la función F se construye a partir de f de la manera indicada. La medida con signo μ definida por $\mu(A) = \displaystyle\int_A f d\lambda$ es absolutamente continua con respecto a la medida de Lebesgue y se tiene $F = F_\mu$. Se obtiene entonces a partir del Teorema 1.3.7 y de la Proposición 1.3.11 que F es de variación acotada, absolutamente continua y que se anula en $-\infty$.

Supongamos ahora que una función $F:\mathbb{R}\longrightarrow\mathbb{R}$ es de variación acotada, absolutamente continua y se anula en $-\infty$. El Teorema 1.3.7 implica que existe una medida finita con signo μ tal que $F=F_\mu$ y la Proposición 1.3.11 implica que $\mu\ll\lambda$. Si $f=\frac{d\mu}{d\lambda}$, entonces se obtiene entonces la definición de F para todo $x\in\mathbb{R}$. ■

Para terminar esta subsección indicamos que la fórmula (1.54) permite estudiar la integral de Stieltjes, pero esto será tratado posteriormente en la subSección 1.4.3.

1.3.4. Diferenciación de funciones

Después de haber estudiado el proceso de diferenciación de medidas, nos interesamos ahora en dar un primer estudio de las propiedades de las derivadas de las funciones cuando se las estudia desde el punto de vista de la teoría de la medida; en particular estaremos en capacidad de presentar desde otra prespectiva a los teoremas fundamentales de cálculo.

Uno de los resultados más importante de esta subsección es el Teorema de Diferenciación de Lebesgue que será tratado únicamente en dimensión 1. Este teorema nos indica cómo recuperar funciones a partir de una integral por medio de un límite muy particular. Dado que la generalización de este hecho a varias dimensiones requiere otro tipo de herramientas, dejamos por el momento este punto abierto. Este aspecto y muchos otros temas relacionados serán tratados en los capítulos siguientes.

Empecemos con un resultado que utiliza ideas expuestas en la sección anterior.

Lema 1.3.8 *Sea μ una medida boreliana definida sobre $\mathbb{R}$ y sea $F:\mathbb{R}\longrightarrow\mathbb{R}$ una función definida por $F(x)=\mu(]-\infty,x])$. Si μ es diferenciable con respecto a la medida de Lebesgue λ en el punto x_0, entonces F es diferenciable en el punto x_0 y se tiene $F'(x_0)=D_\lambda(\mu)(x_0)$.*

Prueba. La diferenciabilidad de μ en el punto x_0 implica que $\mu(\{x_0\})=0$ y por lo tanto que F es continua en el punto x_0. Entonces $\frac{F(x)-F(x_0)}{x-x_0}$ es igual a $\frac{\mu([x_0,x])}{\lambda([x_0,x])}$ si $x_0<x$ y a $\frac{\mu([x,x_0])}{\lambda([x,x_0])}$ si $x<x_0$. Basta en este punto aplicar la definición de $D_\lambda(\mu)(x_0)$ dada en la página 59 para obtener el resultado.■

Lema 1.3.9 *Sea $F:\mathbb{R}\longrightarrow\mathbb{R}$ una función creciente, entonces:*

1) los límites $F(x^-)$ y $F(x^+)$ existen para todo $x\in\mathbb{R}$,

2) el conjunto de puntos en los cuales F no es continua es un conjunto de medida nula,

3) la función $G:\mathbb{R}\longrightarrow\mathbb{R}$ definida por $G(x)=F(x^+)$ es creciente y continua por la derecha y coincide con F en todo punto en donde F es continua.

Prueba. Como F es creciente, los límites $F(x^-)$ y $F(x^+)$ existen y están dados por $F(x^-)=\sup\{F(t): t<x\}$ y $F(x^+)=\inf\{F(t):t>x\}$. Estos límites verifican $F(x^-)\leq F(x)\leq F(x^+)$, de manera que F es continua en x si y solo si $F(x^+)=F(x^-)$. Sea ahora N el conjunto de puntos en los cuales F no es continua y para todo $x\in N$ fijamos un número racional r_x tal que $F(x^-)<r_x<F(x^+)$. Entonces r_x y r_y son diferentes cada vez que x y y son elementos distintos de N, de manera que el conjunto N es de cardinal numerable y por lo tanto es de medida nula.

Supongamos ahora que definimos una función G por medio de la expresión $G(x)=F(x^+)$. La relación $G(x)=\inf\{F(t):t>x\}$ implica que G es creciente y continua por la derecha. Dado que se tiene $F(x)=F(x^+)$ si F es continua en x, se obtiene el lema. ■

Podemos dar ahora un resultado de orden general que será aplicado a las funciones de variación acotada.

Teorema 1.3.8 *Sea $F : \mathbb{R} \longrightarrow \mathbb{R}$ una función creciente. Entonces F es diferenciable en λ-casi todas partes.*

Antes de pasar a la demostración, observamos que a partir de este enunciado es posible considerar funciones monótonas, ya sean éstas crecientes o decrecientes: en efecto, en el caso de que la función F es decreciente, bastaría aplicar el teorema a la función $-F$.

Este teorema, debido a Lebesgue, es bastante sorprendente. En efecto, en base a las intuiciones desarrolladas en los primeros cursos de cálculo, el hecho que una función sea creciente no implica que la función sea diferenciable en un punto y, en el caso de que la función sea diferenciable en un punto, no se tiene necesariamente que esta función es continua en todos sus puntos. Por otro lado, el hecho que existan funciones continuas que no son derivables en ningún punto muestra que no basta ser continua en todo punto para poder ser diferenciable en al menos un punto.

Demostración. Empezamos suponiendo que F es una función creciente acotada, continua por la derecha y que se anula en $-\infty$. Entonces existe una medida boreliana finita μ tal que se tiene para todo $x \in \mathbb{R}$ la relación $F(x) = \mu(]-\infty, x])$. Esto implica entonces por el Teorema 1.3.4 y el Lema 1.3.8 que F es diferenciable en casi todas partes.

Supongamos ahora que F no es continua por la derecha. Entonces definimos $G : \mathbb{R} \longrightarrow \mathbb{R}$ por $G(x) = F(x^+)$ y se obtiene que G es continua por la derecha, de manera que, por lo recién demostrado, se deduce que G es diferenciable en casi todas partes. Observando que se tiene $F(x^-) = G(x^-)$ en casi todas partes, se obtiene que F y G coinciden en cada punto en tonde G es continua y si $F(x_0) = G(x_0)$ entonces

$$\frac{G(x) - G(x_0)}{x - x_0} \leq \frac{F(x) - F(x_0)}{x - x_0} \leq \frac{G(x^-) - G(x_0)}{x - x_0},$$

de manera que si G es diferenciable en x_0, entonces F es diferenciable en x_0 y se tiene $F'(x_0) = G'(x_0)$, de donde se deduce la diferenciabilidad en casi todas partes de F.

Para terminar suponemos que F es una función creciente cualquiera. Es entonces suficiente verificar que F es diferenciable en casi todas partes sobre un intervalo abierto $]a, b[$ puesto que es posible considerar la función

$$x \longmapsto \begin{cases} 0 \text{ si } x \leq a, \\ F(x) - F(a) \text{ si } a < x < b, \\ F(b) - F(a) \text{ si } b \leq x. \end{cases}$$

Con esto se termina la demostración. ■

Corolario 1.3.4 *Sea $F : \mathbb{R} \longrightarrow \mathbb{R}$ una función de variación acotada. Entonces F es diferenciable en λ-casi todas partes*

Prueba. Esto es una consecuencia inmediata del teorema anterior y de la descomposición dada en la Proposición 1.3.9. ■

Veamos ahora un resultado donde las funciones consideradas son definidas por medio de series convergentes.

Proposición 1.3.13 *Sea $F_n : \mathbb{R} \longrightarrow \mathbb{R}$ una sucesión de funciones tales que la serie $\sum_{n\in\mathbb{N}} F_n(x)$ converge para todo punto $x \in \mathbb{R}$. Si definimos $F(x) = \sum_{n\in\mathbb{N}} F_n(x)$, entonces se tiene $F'(x) = \sum_{n\in\mathbb{N}} F_n'(x)$ en casi todas partes.*

Prueba. Supongamos para empezar que las funciones F_n son acotadas, crecientes y continuas por la derecha y que se anulan en $-\infty$ y que la función F definida como la suma de estas funciones es acotada. Sea μ_n las medidas borelianas asociadas a las funciones F_n y definimos la medida μ por medio de la expresión $\mu(A) = \sum_{n\in\mathbb{N}} \mu_n(A)$. Dado que F es acotada y como se tiene $\mu(]-\infty, x]) = F(x)$, se deduce que la medida μ es finita.

Para cada n, consideramos la descomposición de Lebesgue $\mu_n = \mu_{n,a} + \mu_{n,s}$ de la medida μ_n y sea f_n la derivada de Radon-Nikodym de $\mu_{n,a}$ con respecto a la medida de Lebesgue λ y sea N_n un conjunto boreliano de medida nula sobre el cual la medida $\mu_{n,s}$ está concentrada. Vemos entonces que $\sum_{n\in\mathbb{N}} \mu_{n,s}$ está concentrada en $\bigcup_{n\in\mathbb{N}} N_n$ y que se tiene para todo $A \in \mathcal{B}or(\mathbb{R})$ la relación

$$\sum_{n\in\mathbb{N}} \mu_{n,a}(A) = \int_A \sum_{n\in\mathbb{N}} f_n d\lambda.$$

Entonces la descomposición de Lebesgue de μ está dada por $\mu = \sum_{n\in\mathbb{N}} \mu_{n,a} + \sum_{n\in\mathbb{N}} \mu_{n,s}$ y $\sum_{n\in\mathbb{N}} f_n$ es la derivada de Radon-Nikodym de $\sum_{n\in\mathbb{N}} \mu_{n,a}$ con respecto a λ. Se tiene entonces a partir del Teorema 1.3.4 y del Lema 1.3.8, para casi todo $x \in \mathbb{R}$ que

$$\sum_{n\in\mathbb{N}} F_n'(x) = \sum_{n\in\mathbb{N}} D_\lambda(\mu_n)(x) = \sum_{n\in\mathbb{N}} f_n(x) = D_\lambda(\mu)(x) = F'(x).$$

A partir de esto es posible relajar las hipótesis utilizadas para obtener el resultado deseado. ■

Con estos resultados podemos dar otro punto de vista de los teoremas fundamentales del cálculo.

Teorema 1.3.9 *Sea $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}or(\mathbb{R}), \lambda, \mathbb{R})$ y sea $F : \mathbb{R} \longrightarrow \mathbb{R}$ definida por*

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

Entonces F es diferenciable y su derivada está dada por $F'(x) = f(x)$ para λ-casi todo punto $x \in \mathbb{R}$.

Demostración. Supongamos que f es positiva, definimos una medida boreliana positiva μ sobre $\mathbb{R}$ por $\mu(A) = \displaystyle\int_A f dx$. Sea ahora f_0 una función boreliana que coincide con f en λ-casi todas partes. Entonces, por el Teorema 1.3.4 y el Lema 1.3.8 se tiene $F'(x) = D_\lambda(\mu)(x) = f_0(x) = f(x)$ en λ-casi todas partes, de manera que se tiene la demostración en el caso en que f sea positiva. Para una función arbitraria f, basta usar la descomposición $f = f^+ - f^-$ y repetir los mismos argumentos. ■

Teorema 1.3.10 *Una función $F : [a, b] \longrightarrow \mathbb{R}$ es absolutamente continua si y solo si es diferenciable λ-casi todas partes y puede ser reconstruida a partir de su derivada a partir de la fórmula:*

$$F(x) = F(a) + \int_a^x F'(t)dt.$$

Demostración. Supongamos que F es absolutamente continua, entonces también es de variación acotada y por lo tanto también es diferenciable en λ-casi todas partes. De manera que aplicando la Proposición 1.3.12 a la función

$$x \longmapsto \begin{cases} 0 \text{ si } x < a, \\ F(x) - F(a) \text{ si } a \leq x \leq b, \\ F(b) - F(a) \text{ si } b < x. \end{cases}$$

nos proporciona una función $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}or(\mathbb{R})$ para la cual se tiene

$$F(x) = F(a) + \int_a^x f(t)dt.$$

para todo $x \in [a, b]$. El Teorema 1.3.9 implica entonces que F es diferenciable con derivada igual a $F'(x) = f(x)$ en casi todas partes.

Para demostrar la recíproca, la Proposición 1.3.12 implica que cada F definida de esta manera es absolutamente continua. ∎

Corolario 1.3.5 *Sean F y G dos funciones absolutamente continuas definidas en un intervalo $[a, b]$. Entonces se tiene la siguiente fórmula de integración por partes:*

$$F(b)G(b) - F(a)G(a) = \int_a^b F(t)G'(t)dt + \int_a^b G(t)F'(t)dt.$$

Prueba. Empecemos mostrando que la función FG es absolutamente continua. Dado que las funciones F y G son continuas y que el intervalo $[a, b]$ es compacto, entonces existen números positivos m y n tales que se tiene $|F(x)| \leq m$ y $|G(x)| \leq n$ para todo $x \in [a, b]$. Supongamos que $(]s_i, t_i[)_{1 \leq i \leq N}$ es una familia finita de subconjuntos abiertos disjuntos de $[a, b]$, entonces, para cada i se tiene

$$|F(t_i)G(t_i) - F(s_i)G(s_i)| \leq |F(t_i) - F(s_i)||G(t_i)| + |G(t_i) - G(s_i)||F(s_i)| \leq n|F(t_i) - F(s_i)| + m|G(t_i) - G(s_i)|$$

de manera que se obtiene

$$\sum_{i=1}^{N} |F(t_i)G(t_i) - F(s_i)G(s_i)| \leq n \sum_{i=1}^{n} |F(t_i) - F(s_i)| + m \sum_{i=1}^{N} |G(t_i) - G(s_i)|$$

y como las funciones F y G son absolutamente continuas, se puede hacer que las sumas de la derecha de la expresión anterior sean tan pequeñas como se desea, de manera que se deduce la continuidad absoluta de FG.

Aplicamos ahora el Teorema 1.3.10 a la función FG, y, dado que se tiene la identidad $(FG)'(x) = F'(x)G(x) + G'(x)F(x)$ para casi todo punto $x \in [a, b]$, se obtiene el resultado. ∎

Pasemos al resultado más relevante de esta subsección.

Teorema 1.3.11 (Teorema de Diferenciación de Lebesgue) *Sea $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}or(\mathbb{R}), \lambda, \mathbb{R})$ una función, entonces se tiene la relación*

$$\lim_{I \ni x} \frac{1}{|I|} \int_I |f(t) - f(x)|dt = 0 \tag{1.57}$$

en λ-casi todas partes. Aquí I es un intervalo cerrado que contiene x y el límite se toma cuando la longitud del intervalo I tiende a cero.

Demostración. Supongamos primero que la función es boreliana. Basta en este caso fijar un conjunto boreliano acotado y abierto $]a, b[$ y mostrar que se tiene (1.57) para casi todo punto $x \in]a, b[$.

En este sentido, para todo número racional r se puede definir una medida boreliana finita μ:

$$\mu_r(A) = \int_A |f(t) - r|\mathbb{1}_{]a,b[}(t)dt.$$

El Teorema 1.3.4 implica que existe un conjunto N_r de medida de Lebesgue nula tal que

$$D_\lambda(\mu_r)(x) = |f(x) - r|$$

para todo $x \in]a,b[\setminus N_r$. Definamos ahora el conjunto $N = \bigcup_{r\in\mathbb{Q}} N_r$ y supongamos que x pertenece a $]a,b[\setminus N$, que I es un subintervalo cerrado de $]a,b[$ que contiene al punto x y que r es un número racional. Entonces

$$\int_I |f(t) - f(x)|dt \leq \int_I |f(t) - r|dt + \int_I |r - f(x)|dt,$$

y si dividimos cada uno de estos términos por $\lambda(I)$ y hacemos que la longitud de I tienda hacia cero tenemos

$$\limsup_{I\ni x} \frac{1}{|I|}\int_I |f(t) - f(x)|dt \leq D_\lambda(\mu_r)(x) + |r - f(x)| = 2|f(x) - r|.$$

Dado que $|f(x) - r|$ puede ser arbitrariamente pequeño escogiendo correctamente el número racional r se obtiene la identidad (1.57). ■

1.4. Medidas en espacios localmente compactos

Para terminar este capítulo vamos a presentar el teorema de representación de Riesz[10] que permite hacer una conección muy particular entre la teoría de la medida y las formas lineales y este vínculo especial es el que nos permitirá estudiar, gracias al material desarrollado en el capítulo anterior, de manera más detallada las propiedades funcionales de los espacios de Lebesgue.

1.4.1. Introducción

Para enunciar el teorema de representación de Riesz, es necesario recordar algunos aspectos y en esta subsección daremos unos resultados y definiciones que permiten fijar el marco de trabajo.

Sea pues X un espacio separado localmente compacto, si $f : X \longrightarrow \mathbb{R}$ es una función su *soporte* está dado por el conjunto $sop(f) = \overline{\{x \in X : f(x) \neq 0\}}$ y notaremos $\mathcal{C}_c(X,\mathbb{R})$ el conjunto de funciones $f : X \longrightarrow \mathbb{R}$ continuas a soporte compacto.

Lema 1.4.1 *Sea X un espacio separado localmente compacto, sea K un subconjunto compacto de X y sea U un subconjunto abierto de X tal que $K \subset U$. Entonces existe una función $f \in \mathcal{C}_c(X,\mathbb{R})$ tal que $\mathbb{1}_K(x) \leq f(x) \leq \mathbb{1}_U(x)$ y tal que $sop(f) \subset U$.*

Prueba. Usamos la Proposición 1.2.10 del primer volumen para obtener un conjunto abierto V tal que $K \subset V \subset \overline{V} \subset U$. Utilizando el Lema de Urysohn, existe una función continua $g : \overline{V} \longrightarrow [0,1]$ tal que $g(x) = 1$ para todo $x \in K$ y $g(x) = 0$ para todo $\overline{V} \setminus V$. Definimos entonces la función f exigiendo que $f = g$ sobre $\overline{V}$ y tal que $f = 0$ afuera de $\overline{V}$. Por construcción esta función es continua y su soporte está incluido en $\overline{V}$ y por lo tanto es compacto y está incluido en U. ■

Proposición 1.4.1 *Sea X un espacio separado localmente compacto y sea $f \in \mathcal{C}_c(X,\mathbb{R})$. Si $U_1, ..., U_n$ son subconjuntos abiertos de X tales que $sop(f) \subset \bigcup_{k=1}^n U_k$, entonces existen funciones $f_1, ..., f_n \in \mathcal{C}_c(X,\mathbb{R})$ tales que se tenga $f = f_1 + \cdots + f_n$ y tal que para todo $1 \leq k \leq n$, se tiene que el soporte de f_k está contenido en U_k. Además, si la función f es positiva, se puede escoger las funciones f_k de tal forma que sean todas positivas.*

[10] Frigyes Riesz (1880-1956), matemático húngaro.

Prueba. Supongamos para empezar que $n = 2$, es decir que solo hay dos conjuntos U_1 y U_2. Vemos sin mayor problema que es posibel construir dos conjuntos compactos K_1 y K_2 tales que $K_1 \subset U_1$ y $K_2 \subset U_2$ y tales que $sop(f) \subset K_1 \cup K_2$. Utilizamos entonces el Lema 1.4.1 para obtener dos funciones h_1 y h_2 que pertenecen al conjunto $\mathcal{C}_c(X, \mathbb{R})$ y tales que $\mathbb{1}_{K_{1,2}} \leq h_{1,2} \leq \mathbb{1}_{U_{1,2}}$ y $sop(h_{1,2}) \subset U_{1,2}$. Definimos a partir de estas funciones dos funciones adicionales g_1 y g_2 tales que $g_1 = h_1$ y $g_2 = h_2 - \text{ínf}(h_1, h_2)$, tenemos entonces que g_1 y g_2 son funciones positivas, que sus soportes están contenidos en U_1 y U_2 respectivamente y que se tiene $g_1(x) + g_2(x) = \sup(h_1, h_2)(x) = 1$ en cada punto $x \in sop(f)$. Para terminar definimos $f_1 = fg_1$ y $f_2 = fg_2$. Para estudiar el caso cuando $n > 2$ basta proceder por inducción. ■

No es complicado ver que cada una de las funciones de $\mathcal{C}_c(X, \mathbb{R})$ es integrable con respecto a toda medida boreliana regular (es decir es una medida de Radon) definida sobre X: en efecto, cada una de estas funciones continuas es medible con respecto a la σ-álgebra de los borelianos y dado que son funciones acotadas que se anulan fuera de un compacto se deduce que son funciones integrables por la regularidad exigida a las medidas de Radon.

Esta particularidad muestra que el espacio de funciones $\mathcal{C}_c(X, \mathbb{R})$ posee buenas propiedades de integrabilidad con respecto a las medidas de Radon y este hecho será utilizado en los resultados expuestos a continuación.

Necesitaremos para continuar nuestra exposición la siguiente definición.

Definición 1.4.1 (Aplicación lineal positiva) *Una aplicación lineal T definida sobre $\mathcal{C}_c(X, \mathbb{R})$ es* positiva *si para toda función $f \geq 0$ sobre $\mathcal{C}_c(X, \mathbb{R})$, se tiene que $T(f) \geq 0$.*
Esta noción de positividad se extiende a las formas lineales reales definidas sobre $\mathcal{C}_c(X, \mathbb{R})$.

Demos un ejemplo: si μ es una medida boreliana regular positiva sobre X, entonces la aplicación lineal $T(f) = \displaystyle\int_X f(x)d\mu(x)$ es una aplicación lineal positiva. Notamos en particular que una aplicación lineal positiva sobre $\mathcal{C}_c(X, \mathbb{R})$ preserva el orden en el sentido que si $f, g \in \mathcal{C}_c(X, \mathbb{R})$ son dos funciones tales que se tiene $f(x) \leq g(x)$ para todo $x \in X$, entonces se tiene $T(f) \leq T(g)$.

Sabíamos que la integral era un ejemplo de aplicación lineal de manera que lo que es necesario recalcar con la definición anterior es la noción de *positividad*.

1.4.2. Teorema de Representación de Riesz

Presentamos ahora el resultado más relevante de esta subsección en donde se juntan las herramientas estudiadas en este capítulo con las nociones dades en el capítulo anterior. En efecto, vamos a ver que toda forma lineal definida sobre el espacio de funciones continuas a soporte compacto puede ser *representada* por medio de una integral con respecto a una medida de Radon:

Teorema 1.4.1 (de representación de Riesz) *Sea X un espacio separado localmente compacto y sea I una forma lineal positiva definida sobre $\mathcal{C}_c(X, \mathbb{R})$. Entonces existe una única medida de Radon μ definida sobre X tal que*

$$I(f) = \int_X f d\mu$$

para todo $f \in \mathcal{C}_c(X, \mathbb{R})$.

Antes de pasar a la demostración de este teorema necesitaremos una serie de resultados preliminares. El primero de ellos nos explica cómo la medida de un conjunto abierto U puede obtenerse usando las funciones que pertenecen al conjunto $\mathcal{C}_c(X, \mathbb{R})$.

Lema 1.4.2 *Sea X un espacio separado localmente compacto y sea μ una medida boreliana regular definida sobre X. Si U es un conjunto abierto de X se tiene*

$$\mu(U) = \sup\left\{\int_X f(x)d\mu(x) : f \in \mathcal{C}_c(X,\mathbb{R}), 0 \leq f \leq \mathbb{1}_U\right\}$$

Prueba. En el supremo podemos considerar las funciones cuyo soporte está contenido en U. Dado que se tiene que $\mu(U) = \displaystyle\int_X \mathbb{1}_U d\mu$ y que esta cantidad es al menos mayor o igual al supremo considerado, únicamente debemos verificar que se tiene

$$\mu(U) \leq \sup\left\{\int_X f(x)dx : f \in \mathcal{C}_c(X,\mathbb{R}), 0 \leq f \leq \mathbb{1}_U\right\}$$

Para ello consideramos un real α tal que $\alpha < \mu(U)$ y usamos la regularidad de la medida μ para poder escoger un subconjunto compacto K de U tal que $\alpha < \mu(K)$. El Lema 1.4.1 nos proporciona una función $f \in \mathcal{C}_c(X,\mathbb{R})$ que verifica $\mathbb{1}_K \leq f$, $sop(f) \subset U$ y $0 \leq f \leq \mathbb{1}_U$. Tenemos entonces que $\alpha \leq \displaystyle\int_X f d\mu$ y por lo tanto

$$\alpha < \sup\left\{\int_X f(x)d\mu(x) : f \in \mathcal{C}_c(X,\mathbb{R}), 0 \leq f \leq \mathbb{1}_U\right\}$$

Dado que el real α era arbitrario y menor que $\mu(U)$ se obtiene el resultado. ■

El segundo resultado que necesitamos es muy similar al lema anterior, pero en vez de usar una integral utilizaremos una forma lineal positiva y lo que obtendremos no será exactamente una medida, sino una medida exterior. Más precisamente tenemos:

Proposición 1.4.2 *Sea X un espacio separado localmente compacto y sea I una forma lineal positiva definida sobre el conjunto $\mathcal{C}_c(X,\mathbb{R})$. Definimos una función μ^* sobre los conjuntos abiertos por medio de la expresión*

$$\mu^*(U) = \sup\{I(f) : f \in \mathcal{C}_c(X,\mathbb{R}), 0 \leq f \leq \mathbb{1}_U, sop(f) \subset U\} \tag{1.58}$$

y la extendemos a todos los subconjuntos de X escribiendo

$$\mu^*(A) = \inf\{\mu^*(U) : U \text{ abierto y } A \subset U\}. \tag{1.59}$$

Entonces, la aplicación μ^* es una medida exterior sobre el conjunto X y todo conjunto boreliano de X es μ^*-medible.

Prueba. La relación $\mu^*(\emptyset) = 0$ y la propiedad de crecimiento de μ^* se obtienen directamente a partir de la definición de μ^*, de manera que debemos mostrar la σ-aditividad de μ^*. Supongamos primero que $(U_n)_{n\in\mathbb{N}}$ es una sucesión de subconjuntos abiertos de X y mostremos que se tiene

$$\mu^*\left(\bigcup_{n\in\mathbb{N}} U_n\right) \leq \sum_{n\in\mathbb{N}} \mu^*(U_n).$$

Sea pues f una función que pertenece a $\mathcal{C}_c(X,\mathbb{R})$ y que verifica $0 \leq f \leq \mathbb{1}_{\cup_{n\in\mathbb{N}} U_n}$ y tal que $sop(f) \subset \bigcup_{n\in\mathbb{N}} U_n$. Entonces se tiene que $sop(f)$ es, por definición, un subconjunto compacto de $\bigcup_{n\in\mathbb{N}} U_n$ de manera que existe un entero N tal que $sop(f) \subset \bigcup_{n=0}^{N} U_n$. Aplicamos ahora la Proposición 1.4.1 para obtener que la función f puede representarse como la suma de funciones $f_0, ..., f_N$ que pertenecen a $\mathcal{C}_c(X,\mathbb{R})$ y que verifican $0 \leq f_k \leq \mathbb{1}_{U_k}$ y $sop(f_k) \subset U_k$ para todo $0 \leq k \leq N$; es decir que se tiene

$$f = \sum_{k=0}^{N} f_k$$

de manera que, por la linealidad de la funcional I se obtiene $I(f) = \sum_{k=0}^{N} I(f_k)$ y como $I(f_k) \leq \mu^*(U_k)$ para todo $0 \leq k \leq N$, deducimos

$$I(f) \leq \sum_{n=0}^{N} \mu^*(U_n) \leq \sum_{n\in\mathbb{N}} \mu^*(U_n)$$

de donde se obtiene la σ-aditividad de la función μ^* para los conjuntos abiertos.

Ahora supongamos que $(A_n)_{n\in\mathbb{N}}$ es una sucesión cualquiera de subconjuntos de X y podemos suponer sin pérdida de generalidad que $\sum_{n\in\mathbb{N}} \mu^*(A_n) < +\infty$ pues sino no hay nada que demostrar. Sea entonces $\varepsilon > 0$ un real, para todo n usamos (1.59) para escoger un conjunto abierto U_n que contiene A_n y que verifica $\mu^*(U_n) \leq \mu^*(A_n) + \varepsilon/2^n$. Entonces, usando la σ-aditividad sobre los conjuntos abiertos tenemos

$$\mu^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \mu^*\left(\bigcup_{n\in\mathbb{N}} U_n\right) \leq \sum_{n\in\mathbb{N}} \mu^*(U_n) \leq \sum_{n\in\mathbb{N}} \mu^*(A_n) + \varepsilon,$$

pero como ε es arbitrario se deduce la relación de σ-aditividad buscada. De esta manera obtenemos que la función μ^* es una medida exterior sobre X.

Debemos ahora mostrar que todo conjunto boreliano de X es μ^*-medible y para ello bastará considerar los conjuntos abiertos de X. Ahora, para mostrar que un subconjunto abierto U de X es μ^*-medible es suficiente verificar que se tiene la desigualdad

$$\mu^*(A) \geq \mu^*(A \cap U) + \mu^*(A \cap U^c)$$

para todo subconjunto A de X tal que $\mu^*(A) < +\infty$ (ver la Sección 2.3.1 del primer volumen para más detalles relativos a las medidas exteriores). Sea pues A un tal conjunto y sea $\varepsilon > 0$ un real. Por la definición de μ^* dada en la fórmula (1.59) podemos escoger un conjunto abierto V que contiene A y tal que $\mu^*(A) \leq \mu^*(V) \leq \mu^*(A) + \varepsilon$. En este punto bastaría mostrar que se tiene la desigualdad

$$\mu^*(V) \geq \mu^*(V \cap U) + \mu^*(V \cap U^c) - 2\varepsilon \tag{1.60}$$

pues a partir de esto se puede deducir que

$$\mu^*(A) + \varepsilon \geq \mu^*(A \cap U) + \mu^*(A \cap U^c) - 2\varepsilon$$

y se obtendría la desigualdad buscada pues el real $\varepsilon > 0$ era arbitrario.

Demostremos pues la fórmula (1.60) y para ello fijemos una función f_1 en $\mathcal{C}_c(X, \mathbb{R})$ tal que $0 \leq f_1 \leq \mathbb{1}_{V\cap U}$, $sop(f_1) \subset V \cap U$ y tal que $I(f_1) \geq \mu^*(V \cap U) - \varepsilon$, esto es posible por definición de la aplicación μ^*. Definimos ahora $K = sop(f_1)$, entonces $V \cap K^c$ es un conjunto abierto que contiene $V \cap U^c$ y por lo tanto existe una función f_2 en $\mathcal{C}_c(X, \mathbb{R})$ tal que $0 \leq f_2 \leq \mathbb{1}_{V\cap K^c}$, $sop(f_2) \subset V \cap K^c$ y tal que $I(f_2) \geq \mu^*(V \cap K^c) - \varepsilon$. Dado que se tiene que la suma $f_1 + f_2$ verifica $0 \leq f_1 + f_2 \leq \mathbb{1}_V$, $sop(f_1 + f_2) \subset V$, se obtiene por definición $I(f_1 + f_2) \leq \mu^*(V)$ y esto muestra la desigualdad (1.60), de donde se obtiene que U es un conjunto μ^*-medible. ■

El lema a continuación nos proporciona más informaciones sobre la medida exterior que acabamos de construir.

Lema 1.4.3 *Sea μ^* la medida exterior definida por (1.58) y (1.59). Supongamos que A es un subconjunto de X y que $f \in \mathcal{C}_c(X, \mathbb{R})$. Se tienen los puntos siguientes*

1) si se tiene que $\mathbb{1}_A \leq f$, entonces $\mu^*(A) \leq I(f)$, *2) si* $0 \leq f \leq \mathbb{1}_A$ *y si* A *es compacto, entonces* $I(f) \leq \mu^*(A)$.

Prueba. Para el primer punto consideramos $0 < \varepsilon < 1$ un real y definamos U_ε por $U_\varepsilon = \{x \in X : f(x) > 1 - \varepsilon\}$. Vemos entonces que U_ε es un conjunto abierto y cada función $g \in \mathcal{C}_c(X, \mathbb{R})$ que verifica $g \leq \mathbb{1}_{U_\varepsilon}$ también verifica $g \leq \frac{1}{1-\varepsilon} f$ de donde, utilizando (1.58), se obtiene que $\mu^*(U_\varepsilon) \leq \frac{1}{1-\varepsilon} I(f)$. Dado que $A \subset U_\varepsilon$ y como ε puede acercarse arbitrariamiente a 0, se obtiene que $\mu^*(A) \leq I(f)$.

Para el segundo punto, tomamos un conjunto abierto U que contiene A, entonces $0 \leq f \leq \mathbb{1}_U$ y $sop(f) \subset U$ de manera que (1.58) implica que $I(f) \leq \mu^*(U)$. Como U era un abierto arbitrario que contiene A, utilizando (1.59) se tiene que $I(f) \leq \mu^*(A)$. ■

Con estos resultados podemos empezar con toda comodidad la demostración del Teorema 1.4.1.

Demostración del Teorema de Representación de Riesz. Empezamos verificando la unicidad de la medida μ. Supongamos entonces que μ y ν son dos medidas borelianas regulares positivas sobre X tales que se tiene, para toda función $f \in \mathcal{C}_c(X, \mathbb{R})$, las identidades

$$\int_X f d\mu = \int_X f d\nu = I(f).$$

Se deduce a partir del Lema 1.4.2 que se tiene la relación $\mu(U) = \nu(U)$ para todo subconjunto abierto U de X. A partir de esto, usando la regularidad exterior de μ y de ν se deduce que $\mu(A) = \nu(A)$ para todo conjunto boreliano A de X y se obtiene entonces que las medidas μ y ν coinciden, de donde se deduce la unicidad buscada.

Debemos ahora demostrar que a cada forma lineal positiva I definida sobre el espacio $\mathcal{C}_c(X, \mathbb{R})$, existe una medida boreliana regular positiva tal que se tenga la expresión:

$$I(f) = \int_X f d\mu$$

para todo $f \in \mathcal{C}_c(X, \mathbb{R})$.

Sea pues μ^* la medida exterior definida por (1.58) y (1.59), sea μ la restricción a la σ-álgebra de los borelianos $\mathcal{B}or(X)$ de la medida exterior μ^* y sea $\overline{\mu}$ la restricción de μ^* a la σ-álgebra $\mathscr{M}_{\mu^*}$ formada por los conjuntos μ^*-medibles.

Vamos a ver que las medidas μ y $\overline{\mu}$ son medidas regulares y que se tiene

$$\int_X f(x) d\mu(x) = \int_X f(x) d\overline{\mu}(x) = I(f)$$

para todo $f \in \mathcal{C}_c(X, \mathbb{R})$, obteniendo así el resultado deseado.

Sabemos por la teoría desarrollada en el primer volumen que $\overline{\mu}$ es una medida sobre $\mathscr{M}_{\mu^*}$ y, dado que $\mathcal{B}or(X) \subset \mathscr{M}_{\mu^*}$ por la Proposición 1.4.2, se tiene que μ es una medida sobre la σ-álgebra $\mathcal{B}or(X)$. Ahora, utilizando el Lema 1.4.1, tenemos que para todo conjunto compacto K de X existe una función $f \in \mathcal{C}_c(X, \mathbb{R})$ que verifica $\mathbb{1}_K \leq f$; se deduce entonces por el Lema 1.4.3 que las medidas μ y $\overline{\mu}$ son finitas sobre los compactos.

Verifiquemos la regularidad interior y exterior de estas medidas. Vemos directamente que la regularidad exterior se deduce de la expresión (1.59). La regularidad interior requiere unas pocas líneas adicionales. Queremos ver que se tiene para todo conjunto abierto U la fórmula $\mu^*(U) = \sup\{\mu(K) : K \subset U, \text{ con } K \text{ compacto}\}$. Por el crecimiento de las medidas tenemos siempre que

$\sup\{\mu(K) : K \subset U, \text{ con } K \text{ compacto}\} \leq \mu^*(U)$ de manera que nos concentramos en la desigualdad recíproca. Pero, por la fórmula (1.58) tenemos

$$\mu^*(U) = \sup\{I(f) : f \in \mathcal{C}_c(X, \mathbb{R}), 0 \leq f \leq \mathbb{1}_U, \ sop(f) \subset U\}$$

y por otro lado, utilizando el Lema 1.4.3 obtenemos

$$\mu^*(U) \leq \sup\{\mu^*(sop(f)) : f \in \mathcal{C}_c(X, \mathbb{R}), 0 \leq f \leq \mathbb{1}_U, \ sop(f) \subset U\}$$

y con esto tenemos que estas dos medidas son regulares interiormente.

Ahora estudiemos la identidad dada en el enunciado de la proposición. Dado que cada función de $\mathcal{C}_c(X, \mathbb{R})$ puede expresarse como la diferencia de dos funciones positivas de $\mathcal{C}_c(X, \mathbb{R})$, podemos restringirnos a las funciones positivas de $\mathcal{C}_c(X, \mathbb{R})$. Sea pues f una tal función y sea $\varepsilon > 0$ un real. Definimos para todo $n \geq 1$ las funciones f_n por

$$f_n(x) = \begin{cases} 0 & \text{si } f(x) \leq (n-1)\varepsilon \\ f(x) - (n-1)\varepsilon & \text{si } (n-1)\varepsilon < f(x) \leq n\varepsilon \\ \varepsilon & \text{si } n\varepsilon < f(x) \end{cases}$$

Entonces se tiene que $f = \sum_{n \geq 1} f_n$ y cada función f_n pertenece a $\mathcal{C}_c(X, \mathbb{R})$, además existe un entero N tal que $f_n = 0$ si $n > N$. Sea ahora $K_0 = sop(f)$ y para todo $n \geq 1$ definimos $K_n = \{x \in X : f(x) \geq n\varepsilon\}$. Se tiene entonces $\varepsilon \mathbb{1}_{K_n} \leq f_n \leq \varepsilon \mathbb{1}_{K_{n-1}}$ para todo n y entonces el Lema 1.4.3 así como las propiedades de la integral implican que se tienen las desigualdades $\varepsilon\mu(K_n) \leq I(f_n) \leq \varepsilon\mu(K_{n-1})$ y $\varepsilon\mu(K_n) \leq \displaystyle\int_X f_n d\mu \leq \varepsilon\mu(K_{n-1})$ para todo $n \geq 1$. Dado que $f = \sum_{n \geq 1} f_n$, podemos escribir

$$\sum_{n=1}^{N} \varepsilon\mu(K_n) \leq I(f) \leq \sum_{n=0}^{N-1} \varepsilon\mu(K_n) \quad \text{y} \quad \sum_{n=1}^{N} \varepsilon\mu(K_n) \leq \int_X f d\mu \leq \sum_{n=0}^{N-1} \varepsilon\mu(K_n).$$

Entonces $I(f)$ y $\displaystyle\int_X f d\mu$ se encuentran en un intervalo de longitud $\sum_{n=0}^{N-1} \varepsilon\mu(K_n) - \sum_{n=1}^{N} \varepsilon\mu(K_n) \leq \varepsilon\mu(sop(f))$, pero dado que el real $\varepsilon > 0$ era arbitrario se obtiene que las cantidades $I(f)$ y $\displaystyle\int_X f d\mu$ deben coincidir. Dado que $\displaystyle\int_X f d\mu = \int_X f d\overline{\mu}$ hemos terminado la demostración del teorema de representación de Riesz. ■

Es importante notar que en la demostración de este teorema hemos construido dos medidas, μ y $\overline{\mu}$, que verifican el enunciado. La primera medida μ está definida sobre la σ-álgebra de los borelianos mientras que la segunda medida $\overline{\mu}$ está definida sobre la σ-álgebra $\mathscr{M}_{\mu^*}$, y, si bien estas dos medidas coinciden sobre la σ-algebra boreliana, puede darse el caso que estas medidas difieran. Esta diferencia debe comprenderse desde el punto de vista del problema de prolongación, de extensión y de completación de medidas estudiado en la Sección 2.3.3 del primer capítulo.

1.4.3. Medidas de Stieltjes sobre $\mathbb{R}$

Vamos a ver de qué manera es posible aplicar el Teorema de Representación de Riesz al estudio de las medidas de Stieltjes[11] sobre $\mathbb{R}$ y su relación con la integral de Stieltjes.

[11]Thomas-Joannes Stieltjes (1856-1894), matemático neerlandés.

Empecemos recordando la noción de integral[12] de Stieltjes. Sea pues $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ una función creciente, dado que toda función escalonada $f : \mathbb{R} \longrightarrow \mathbb{R}$ continua por la derecha y que se anula afuera de un conjunto compacto puede escribirse (de una infinidad de formas) como

$$f(x) = \sum_{k=1}^{n} \alpha_k \mathbb{1}_{[a_{k-1}, a_k[}(x) \qquad (\alpha_k \in \mathbb{R})$$

donde $(a_k)_{0\leq k\leq n}$ es una sucesión creciente de reales, podemos definir la cantidad $S_\varphi(f)$ de la siguiente manera

$$S_\varphi(f) = \sum_{k=1}^{n} \alpha_k \left(\varphi(a_k^-) - \varphi(a_{k-1}^-)\right).$$

Vemos en particular que esta cantidad depende únicamente de la función escalonada f y no de la descomposición que utiliza la sucesión $(a_k)_{0\leq k\leq n}$. Ahora, si definimos $a_0 = a$ y $a_n = b$ tenemos

$$|S_\varphi(f)| \leq \|f\|_\infty \left(\varphi(b^-) - \varphi(a^-)\right). \tag{1.61}$$

Sea ahora f un elemento de $\mathcal{C}_c(\mathbb{R}, \mathbb{R})$ y sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones escalonadas, continuas por la derecha, nulas fuera de un intervalo determinado $[a, b]$ que convergen uniformemente hacia f. Gracias a la desigualdad (1.61), vemos sin mayor problema que la sucesión $(S_\varphi(f_n))_{n\in\mathbb{N}}$ converge hacia un número que depende únicamente de la función f. A este resultado se lo conoce como la *integral de Stieltjes* de f con respecto a la función φ y se lo nota por

$$\int_{\mathbb{R}} f(x) d\varphi(x).$$

Vamos ahora a relacionar esta integral con resultados presentados anteriormente en la Sección 1.3.3 por medio del teorema de representación de Riesz.

Teorema 1.4.2 *Sea una función creciente $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$. Entonces existe una única medida boreliana μ definida sobre $\mathbb{R}$ tal que se tenga la relación*

$$\int_{\mathbb{R}} f(x) d\mu(x) = \int_{\mathbb{R}} f(x) d\varphi(x),$$

para todo elemento $f \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$. Además esta medida μ está determinada sobre todo intervalo de la forma $]a, b]$ por la relación

$$\mu(]a, b]) = \varphi(b^-) - \varphi(a^-). \tag{1.62}$$

Diremos entonces que la medida μ es la medida de Stieltjes asociada a la función φ.

Demostración. Vemos sin mayor problema que la aplicación $I : f \longmapsto \displaystyle\int f(x) d\varphi(x)$ es una forma lineal positiva definida sobre el conjunto $\mathcal{C}_c(\mathbb{R}, \mathbb{R})$. Aplicamos en este punto el Teorema de Representación de Riesz para obtener la existencia de una única medida boreliana regular positiva μ definida sobre $\mathbb{R}$ que verifica

$$I(f) = \int_{\mathbb{R}} f(x) d\mu(x)$$

para todo elemento $f \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$.

Sean ahora a, b dos puntos de $\mathbb{R}$ tales que $a < b$ y sea $n > \frac{1}{b-a}$ un entero. Definimos una función f_n de la siguiente manera: f_n es nula en los intervalos $]-\infty, a - 1/n]$ y $]b, +\infty[$, vale 1 si $a < x \leq b - 1/n$

[12] También conocida como la Integral de Riemann-Stieltjes

y es lineal en los intervalos $]a-1/n, a]$ y $]b-1/n, b]$. Vemos sin problema que esta función es continua y que converge en casi todas partes hacia la función indicatriz del intervalo $]a, b]$. Podemos entonces aplicar sin problema el Teorema de Convergencia Dominada de Lebesgue para ver que la integral $\displaystyle\int f_n d\mu$ converge hacia $\mu(]a, b])$ cuando $n \to +\infty$. Como se tienen las desigualdades

$$\mathbb{1}_{]a-1/n,a]} \leq f_n \leq \mathbb{1}_{]a-1/n,b]},$$

se tiene por definición de la forma lineal $I(f_n) = \displaystyle\int f_n(x)d\varphi(x)$ que

$$\varphi((b-1/n)^-) - \varphi(a^-) \leq \int_{\mathbb{R}} f d\mu \leq \varphi(b^-) - \varphi((a-1/n)^-)$$

de donde se obtiene el resultado deseado al hacer tender $n \to +\infty$. ■

Observación 1.11 El lector notará que este vínculo con la integral de Stieltjes se basa en la posibilidad de representar medidas por medio del Teorema de Representación de Riesz, otra forma de ver este resultado consiste en adoptar el punto de vista expuesto en la Sección 1.3.3.

En efecto, en el enunciado del teorema anterior hemos supuesto que la función φ era creciente y esto puede relacionarse con la noción de funciones de variación acotada via el Teorema 1.3.8 presentado en la página 72 y la Proposición 1.3.9 dada en la página 68.

Finalmente, es interesante notar que, cuando la función φ es igual a la función identidad sobre $\mathbb{R}$ (es decir que $\varphi(x) = x$), entonces la medida obtenida por medio de la expresión (1.62) no es más que la medida de Lebesgue y este resultado proporciona otra verificación de la existencia y unicidad de esta medida.

Resumen

En este capítulo hemos desarrollado algunos aspectos de la teoría de la medida y de la integración que no habían sido expuestos en el primer volumen y que era indispensable presentar para estudiar con comodidad las diversas propiedades de los espacios de Lebesgue y de Lorentz así como las aplicaciones de estas propiedades en el análisis matemático. En estas pocas líneas resaltamos las ideas más importantes:

- El hecho de considerar medidas con signo nos ha llevado a desarrollar varios teoremas de descomposición (de Hahn y de Jordan) que permiten, junto con la noción de variación total de una medida, dar una estructura de espacio de Banach al espacio de medidas cuya norma de variación total es finita.

- Gracias a la posibilidad de estudiar medidas con signo, hemos presentado el teorema de Radon-Nikodym, resultado que reposa sobre la noción de continuidad absoluta y que permite expresar medidas como integrales de funciones. Este tipo de caracterización es fundamental en el estudio de la teoría de probabilidades y es el punto de partida de muchos conceptos, como la noción de derivada de Radon-Nikodym de una medida.

- El teorema de Radon-Nikodym mencionado anteriormente permite "integrar" una medida, pero para "derivarlas" es necesario estudiar el comportamiento de las medidas con respecto a un cierto límite muy particular. Para ello presentamos unos lemas de recubrimiento que son la base de muchos resultados en el análisis armónico como tendremos la oportunidad de verlo posteriormente.

- Gracias a estos resultados presentamos, rápidamente y en una sola dimensión, el teorema de diferenciación de Lebesgue cuya generalización a varias dimensiones requerirá introducir nuevas herramientas muy potentes para resolver diversos problemas del análisis matemático.
- Finalmente, hemos estudiado la relación que existe entre formas lineales positivas y las medidas de Radon. Ya sabíamos que la integral de una función es una forma lineal, pero lo expuesto en este capítulo muestra de qué manera cada forma lineal positiva puede escribirse como una integral y esta forma de *representar* las formas lineales se conoce como el teorema de representación de Riesz. Para obtener esta representación hemos trabajado sobre el espacio de funciones continuas a soporte compacto definidas sobre un espacio separado localmente compacto y veremos posteriormente cómo ampliar este marco de estudio.

1.5. Ejercicios

Ejercicio 1.1 *Sea μ una medida positiva y sea $f : \mathbb{R} \longrightarrow \mathbb{R}$ una función tal que $f \in L^1(X, \mathscr{A}, \mu, \mathbb{R})$. Mostrar que la aplicación ν definida por*

$$\nu(A) = \int_A f(x) d\mu(x)$$

para todo $A \in \mathscr{A}$ es una medida con signo, para ello seguir los puntos siguientes:

1. *Verificar que $\nu(\emptyset) = 0$.*
2. *Verificar que la medida $\nu(A)$ de todo conjunto $A \in \mathscr{A}$ puede escribirse como*

$$\nu(A) = \int_A f^+(x) d\mu(x) - \int_A f^-(x) d\mu(x)$$

   con $f^+ = \text{máx}(f, 0)$ y $f^- = \text{máx}(-f, 0)$.

3. *Mostrar que si $A, B \in \mathscr{A}$ son dos conjuntos disjuntos, entonces $\nu(A \cup B) = \nu(A) + \nu(B)$.*
4. *Usando el hecho que $f \in L^1(X, \mathscr{A}, \mu, \mathbb{R})$ y el teorema de convergencia dominada, obtener la propiedad de σ-aditividad numerable*

$$\nu\left(\bigcup_{i=0}^{+\infty} A_i\right) = \sum_{i=0}^{+\infty} \nu(A_i)$$

   donde $(A_i)_{i\in\mathbb{N}}$ es una sucesión de conjuntos disjuntos $\mathscr{A}$-medibles.

Ejercicio 1.2 *Sea $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y sea una medida μ definida por $\mu = -\delta_{-1} + \delta_1$. Dar dos descomposiciones de Hahn distintas del espacio $\mathbb{R}$ asociadas a la medida μ.*

Ejercicio 1.3 *Sea $(X, \mathscr{A})$ un espacio medible y sean μ, ν, ν_1 y ν_2 medidas definidas sobre $\mathscr{A}$. Demostrar los siguientes puntos.*

1. *Si la medida μ está concentrada en un conjunto E, entonces la medida $|\mu|$ también está concentrada sobre E.*
2. *Si $\mu \perp \nu$ entonces $|\mu| \perp |\nu|$.*
3. *Si $\nu_1 \perp \mu$ y si $\nu_2 \perp \mu$ entonces $(\nu_1 + \nu_2) \perp \mu$.*
4. *Si $\nu \ll \mu$ entonces $|\nu| \perp \mu$.*

5. *Si* $\nu_1 \ll \mu$ *y si* $\nu_2 \perp \mu$*, entonces* $\nu_1 \perp \nu_2$*.*

Ejercicio 1.4 *Sean* μ*,* ν *dos medidas positivas definidas sobre el mismo espacio medible* $(X, \mathscr{A})$*. Suponiendo que* $\nu \ll \mu$ *y que* $\frac{d\nu}{d\mu}$ *existe.*

1. *Mostrar que si se tiene* $\frac{d\nu}{d\mu} > 0$ *en* μ*-casi todas partes entonces* $\mu \ll \nu$ *y obtener que estas dos medidas son equivalentes.*

2. *Mostrar que si se tiene* $\frac{d\nu}{d\mu} > 0$ *en* μ*-casi todas partes entonces y si las medidas* μ *y* ν *son* σ*-finitas, entonces* $\frac{d\mu}{d\nu}$ *existe y se tiene la relación*

$$\frac{d\nu}{d\mu} = \left(\frac{d\mu}{d\nu}\right)^{-1}$$

en μ*- y* ν*-casi todas partes.*

Ejercicio 1.5 *El objetivo de este ejercicio es evidenciar que las nociones de continuidad uniforme y continuidad absoluta son diferentes. Sea* f *una función definida por:*

$$\begin{aligned} f : [0,1] &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) = \begin{cases} 0 & \text{si } x = 0, \\ x\sin(\frac{\pi}{x}) & \text{si } x \in]0,1] \end{cases} \end{aligned}$$

1. *Verificar que esta función es uniformemente continua sobre* $[0,1]$ *(usar el hecho que el intervalo* $[0,1]$ *es un compacto y la Proposición 1.2.8 del primer volumen).*

2. *Sean* $(x_n)_{n\geq 0}$ *y* $(y_n)_{n\geq 1}$ *dos sucesiones de términos generales dados por*

$$x_n = \frac{2}{2n+1} \qquad y_n = \frac{2}{2n-1}.$$

Verificar que para todo $n \geq 2$ *se tienen las desigualdades:* $0 < x_n < x_{n-1} = y_n \leq 1$*.*

3. *Sea* $\eta > 0$ *y sea* $k \geq 2$ *tal que* $y_n < \eta$ *para todo entero* $n \geq k$*. Considerando la familia*

$$(x_{2k}, x_{2k-1}, \cdots, x_k, y_{2k}, y_{2k-1}, \cdots, y_k)$$

de puntos del intervalo $[0,1]$ *mostrar que se tiene*

$$\sum_{j=0}^{k} (y_{k+j} - x_{k+j}) = y_k - x_{2k} < y_k < \eta.$$

4. *Mostrar que* $|f(y_n) - f(x_n)| > \frac{2}{n}$*.*

5. *Mostrar que se tiene*

$$\sum_{j=0}^{k} |f(y_{k+j}) - f(x_{k+j})| \geq 1.$$

6. *Deducir que la función no es absolutamente continua.*

Ejercicio 1.6 *En este ejercicio estudiamos las relaciones existentes entre funciones absolutamente continuas y funciones de variación acotada.*

1. *Mostrar que toda función $f : \mathbb{R} \longrightarrow \mathbb{R}$ absolutamente continua es de variación acotada.*

2. *Considerar la función de Lebesgue, también llamada la escalera de Cantor, (ver el Teorema 3.2.2 del primero volumen) para verificar que una función de variación acotada no es necesariamente absolutamente continua.*

3. *Verificar que la función f definida en el intervalo $]-1,1[$ por $f(x) = x^2 \cos^2(\pi/x^2)$ si $x \neq 0$ y por $f(0) = 0$ es continua.*

4. *Mostrar que esta función no es de variación acotada.*

Ejercicio 1.7 *Consideramos la estructura de espacio medido natural de $\mathbb{R}$ y sea $f : [a,b] \longrightarrow \mathbb{R}$ una función. Mostrar que las tres aserciones siguientes son equivalentes:*

- *f es una función absolutamente continua.*
- *Existe una función $g \in \mathcal{L}^1([a,b], dx)$ tal que para todo $x \in [a,b]$ se tenga*

$$f(x) = f(a) + \int_a^x g(t)dt$$

- *f es una función diferenciable en casi todo punto de $]a,b[$, $f' \in \mathcal{L}^1([a,b], dx)$ y*

$$f(x) = f(a) + \int_a^x f'(t)dt$$

para todo $x \in [a,b]$.

Ejercicio 1.8 *Sea $(X, \mathscr{A})$ un espacio medible y sea μ una medida positiva definida sobre él. Si $f : X \longrightarrow \mathbb{R}$ es una función μ-integrable mostrar que $(f\mu)^+ = f^+\mu$ y que $(f\mu)^- = f^-\mu$. Dar la descomposición de Jordan de la medida $(f\mu)$.*

Ejercicio 1.9 *Sea la medida de Lebesgue λ definida sobre el espacio medido $(\mathbb{R}, \mathcal{B}or(\mathbb{R}))$ y sea ν la medida definida por $\nu(A) = \lambda(A \cap]-\infty, 0]) + 2\lambda(A \cap]0, +\infty[)$ para todo $A \in \mathcal{B}or(\mathbb{R})$.*

1. *Si $f : \mathbb{R} \longrightarrow \mathbb{R}$ es una función tal que $f(x) = x^2$ si $x > 0$ y $f(x) = x^4$ si $x \leq 0$, verificar que $\nu([-\varepsilon, \varepsilon]) = 3\varepsilon$ para todo $\varepsilon > 0$.*

2. *Calcular $(f\nu)([-\varepsilon, \varepsilon])$ y $(f\lambda)([-\varepsilon, \varepsilon])$.*

3. *Determinar $\frac{d(f\nu)}{d(f\mu)}$. ¿Se tiene la identidad $\frac{d(f\nu)}{d(f\mu)} = \frac{d\nu}{d\mu}$?*

4. *¿Bajo qué condiciones sobre la función f se tiene la relación $\frac{d(f\nu)}{d(f\mu)} = \frac{d\nu}{d\mu}$?*

Ejercicio 1.10 *Sea el espacio medido $(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n), dx)$ y sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ una función localmente integrable[13]. Para todo $A \in \mathcal{B}or(\mathbb{R}^n)$ definimos*

$$\nu(A) = \int_A f(x)dx.$$

1. *Mostrar que ν es finita sobre los compactos.*

2. *¿Se tiene en toda generalidad $\nu(\mathbb{R}^n) < +\infty$? A modo de ejemplo considerar la función $x \longmapsto |x|^{-(n-1/4)}$.*

[13] Ver la Definición 4.4.1 del Volumen 1.

3. *¿Es la aplicación ν una medida? ¿Es una medida de Radon?*

Ejercicio 1.11 *Construir un ejemplo donde el Teorema de Convergencia Monótona de Beppo Levi falla al considerar funciones que pueden tomar valores negativos.*

Ejercicio 1.12 *Construir un ejemplo donde el Lema de Fatou falla al considerar funciones que pueden tomar valores negativos.*

Ejercicio 1.13 *Sea X un espacio separado localmente compacto y sea I una forma lineal positiva definida sobre el espacio $\mathcal{C}_c(X,\mathbb{R})$. Mostrar que si la forma lineal I es continua, entonces se tiene*

$$\|I\|_{\mathcal{C}_c(X,\mathbb{R})\to\mathbb{R}} = \mu(X) < +\infty,$$

donde μ es la medida de Radon asociada a la forma lineal I por medio del Teorema de Riesz.

Ejercicio 1.14 *Sea $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ la función de Heaviside*[14] *determinada por*

$$\varphi(x) = \begin{cases} 0 & \textit{si } x < 0 \\ 1 & \textit{si } 0 \le x \end{cases}$$

¿Cuál es la medida de Stieltjes asociada a la función φ?

Ejercicio 1.15 *Sea X un espacio separado localmente compacto y sea μ una medida de Radon definida sobre el espacio medible $(X, \mathcal{B}or(X))$. Sea f una función $\mathcal{B}or(X)$-medible a valores en $\mathbb{R}$. Mostrar que para todo $\varepsilon > 0$ existe un subconjunto boreliano $A \subset X$ tal que $\mu(X \setminus A) < \varepsilon$ y tal que la restricción de la función f al conjunto A es continua.*

A este resultado se lo conoce como el Teorema de Lusin[15].

[14] Oliver Heaviside (1850-1925), físico inglés.
[15] Nikolaï Nikolaïevitch Lusin (1883-1950), matemático ruso.

Índice alfabético

Absoluta continuidad
de funciones, 68
de medidas complejas, 29
de medidas con signo, 29
de medidas positivas, 26
Aplicación lineal
positiva, 76
Area de la esfera unidad, 51

Completitud
del espacio de Medidas, 19
Conjunto
despreciable con respecto a una medida con signo, 9
negativo, 11
positivo, 11
Continuidad
absoluta de funciones, 68
absoluta de medidas complejas, 29
absoluta de medidas con signo, 29
absoluta de medidas positivas, 26
de las medidas con signo, 8
Convergencia normal
de medidas compleja, 10
de medidas con signo, 6
Coordenadas
esféricas en $\mathbb{R}^3$, 48
esféricas en $\mathbb{R}^n$, 49
polares en $\mathbb{R}^2$, 48

Densidad de una medida, 59
Derivada
de Radon-Nikodym, 30
de una medida, 59
inferior de una medida, 59
superior de una medida, 59
Descomposición de medidas, 11
de Hahn, 12
de Jordan, 14
Determinante, 42

Equivalencia de medidas, 26
Espacio
de funciones integrables con respecto a una medida con signo o compleja, 22
de Lebesgue
con respecto a una medida compleja, 23
con respecto a una medida con signo, 23
de Medidas, 19
Esperanza Condicional, 36

Forma lineal
positiva, 76
Función
absolutamente continua, 68
de Heaviside, 86
de variación acotada, 66
diferenciable, 43
finitamente aditiva, 4
numerablemente aditiva, 4
soporte de una, 75

Hahn
Descomposición de, 12

Imagen de una Medida, 39
Integral
de Stieltjes, 80
con respecto a una medida compleja, 21
con respecto a una medida con signo, 21
de una función radial, 50

Jacobiano, 44
Jordan
Descomposición de, 14

Lebesgue
teorema de descomposición, 37
teorema de diferenciación, 74
Lema
de recubrimiento de Besicovitch, 55
de recubrimiento de Vitali, 52
Lusin

Teorema de, 86

Matriz Jacobiana, 44
Medida
invariante sobre la esfera, 49
compleja, 9
con signo, 4
con signo finita, 5
de Radon, 51
de Stieltjes, 80
descomposición, 11
de Hahn, 12
de Jordan, 14
de Lebesgue, 37
Imagen, 39
inducida por una función, 6
variación, 15
variación compleja, 16
variación total, 15

Norma
de la Variación Total, 19
de una aplicación lineal, 44

Producto de una medida por una función, 24

Recubrimiento fino, 53
Regla de la Cadena
para funciones, 44
para medidas, 33
Riesz
Teorema de Representación, 76

Soporte
de una función, 75
Stieltjes
Integral de, 80
Sub-σ-álgebra, 34

Teorema
de cambio de variable, 45
de convergencia dominada
para medidas complejas, 22
para medidas con signo, 22
de descomposición
de Hahn, 12
de Lebesgue, 37
de Diferenciación de Lebesgue, 74
de Lusin, 86
de Radon - Nikodym, 27
para medidas complejas, 30
para medidas con signo, 30
de Representación de Riesz, 76
del valor medio, 44
fundamental del cálculo
para medidas de Radon, 62

Variación
de una función, 66
de una medida, 15
de una medida compleja, 16
total de medidas, 15