Universidad Autónoma de Querétaro

Facultad de Ingeniería

Licenciatura en Matemáticas Aplicadas

Comparación entre la integral de Riemann y la integral de Lebesgue

Tesis

Que como parte de los requisitos para obtener el Grado de Licenciado en Matemáticas Aplicadas.

Presenta

Edgar Iván Hernández Reséndiz

Dirigido por:

Dr. Francisco Gerardo Jiménez López

Centro Universitario, Querétaro, Qro.
Agosto, 2021.
México

Universidad Autónoma de Querétaro

Facultad de Ingeniería

Licenciatura en Matemáticas Aplicadas

Comparación entre la integral de Riemann y la integral de Lebesgue

Tesis

Que como parte de los requisitos para obtener el Grado de Licenciado en Matemáticas Aplicadas.

Presenta

Edgar Iván Hernández Reséndiz

Dirigido por:

Dr. Francisco Gerardo Jiménez López

Dr. Francisco Gerardo Jiménez López

Presidente

Dr. Víctor Antonio Aguilar Arteaga

Secretario

Dra. Angélica Rosario Jiménez Sánchez

Vocal

Dr. Jesús Jerónimo Castro

Sinodal

Centro Universitario, Querétaro, Qro.
Agosto, 2021.
México

Agradecimientos

Me gustaría comenzar agradeciendo a la Universidad Autónoma de Querétaro, y siendo más específicos, a la Facultad de Ingeniería, pues en ella he encontrado no solo un lugar donde estudiar, aprender y prepararme, sino que gracias a todas las personas que forman parte de esta, me he sentido pleno en mi estancia en la universidad.

Le agradezco infinitamente a mis maestros, pues además de las lecciones aprendidas, muchos de ellos sembraron en mí el profundo aprecio que siento por las matemáticas. De manera especial, agradezco a la persona encargada de dirigir esta tesis, el Dr. Francisco Gerardo Jiménez López que, gracias a su ayuda y dirección, he podido disfrutar y aprender mucho al hacer este trabajo.

A mis compañeros, que fueron la parte más importante en mi vida universitaria. Gracias por crecer junto a mí, por las risas y por todo.

Por último, pero más importante, quiero agradecer a mi familia quienes me han brindado todo su cariño y apoyo. Agradezco a mis hermanos, pero agradezco a mi madre más que a cualquier otra persona, pues no hay muestra de amor más grande que todo lo que ella ha hecho por nosotros, razón por lo cual, le dedico a ella en especial todo lo que este trabajo representa.

Índice

1. Introducción

La integral es uno de los conceptos fundamentales del cálculo y el análisis matemático que formaliza una idea simple e intuitiva: el área. Dicho concepto es estudiado en las carreras de ingeniería debido a las aplicaciones físicas que de ésta emanan y tiene sus orígenes desde que el hombre se interesó en cálcular áreas o volúmenes de figuras regulares. No obstante, no fue sino hasta 1854 que en [**13**], Riemann expuso la primera definición formal de integral para una función acotada y definida sobre un intervalo y si bien, para muchas situaciones la integral de Riemann resulta bastante útil y suficiente, hay algunas otras donde este tipo de integral no puede ser aplicada, ya sea debido a que las funciones no son acotadas o tienen una cantidad "muy grande" de discontinuidades, o bien, el conjunto de definición de las funciones que queremos estudiar, no es el de los números reales. Por estos motivos, resulta interesante explorar la posibilidad de ampliar en algún sentido la definición de integral que es dada por Riemann. Es así como nacen teorías de integración más generales como lo son la integral de Riemann-Stieltjes o aún más general, la integral de Lebesgue, que permite definir el análogo de área bajo la curva a funciones real valuadas cuyo dominio es un espacio arbitrario. Es precisamene aquí, entre la integral de Riemann y la integral de Lebesgue, donde yace la ocupación de este trabajo. Hacer una comparación entre estas dos teorías desarrolladas para el mismo concepto.

Aún cuando la integral de Lebesgue representa una generalización útil de la integral de Riemann, resulta difícil para una persona sin mucha formación en cálculo y teoría de conjuntos llegar a entender con precisión los contenidos de un curso de teoría de la medida. Sin embargo, es posible brindar, de manera amigable y sin necesidad de extender demasiado la teoría más allá de un curso de cálculo integral, resultados análogos a los teoremas más repesentativos de la integral de Lebesgue, a saber, el teorema de la convergencia monótona (TCM) y el teorema de la convergencia dominada (TCD), pero adecuados a la integral de Riemann. Es por ello que uno de los puntos importantes a tratar en este trabajo es hacer ese acercamiento a los alumnos de primeros semestres de la carrera de matemáticas aplicadas o afines hacia estos resultados que no son incluidos en los cursos convencionales de cálculo. Esto con la intención de desarrollar la intuición del alumno para que logre asimilar de mejor manera los conceptos que se le presenten en cursos más avanzados. Antes de mostrar dicho acercamiento a las personas interesadas, en el capítulo 2 exhibimos algunos resultados que aligeran los contenidos de los capítulos 3 y 4 y además, definimos los conceptos de sucesiones de funciones y convergencia puntual. Siendo el objetivo de esto último, introducir dichas ideas a las personas que no se encuentren del todo familiarizados con estos conceptos.

Comenzamos el capítulo 3 desarrollando la teoría básica de la integral de Riemann y la integral de Darboux, para posteriormente exponer en la sección 3.4 los equivalentes al teorema de la convergencia monótona y teorema de la convergencia dominada para la integral de Riemann y finalmente, en la sección 3.5, extendemos estos resultados a la integral impropia de Riemann. El trabajo presentado en esta parte del documento, está fuertemente basado en lo presentado por Brian S. Thomson y Wilhelmus Luxemburg en [**17**] y [**9**], respectivamente. Sin embargo, aquí no se reproducen los argumentos de estos autores, sino que se proponen algunas variantes con la intención de aprovechar lo mejor posible las ideas presentadas en sus trabajos, pero sin dejar de lado la accesibilidad que se pretende tenga esta parte de la tesis.

A lo largo de la sección 3.4 se habla más detalladamente de este enfoque diferente.

Continuando con la comparación entre ambas integrales, en el capítulo 4 partimos presentando la teoría básica y necesaria de la integral de Lebesgue con la intención de demostrar los teoremas más representativos de esta integral; el TCM, TCD y el lema de Fatou para posteriormente, en la sección 4.6, exponer un único resultado. Dicho resultado no solo dicta la relación que se da entre la integral de Lebesgue y la de Riemann, sino que además, brinda una caracterización sumamente interesante de las funciones Riemann integrables en función de la medida del conjunto de puntos de discontinuidades de la función. Finalmente, de la misma manera que en la primera parte del trabajo nos preguntamos por los equivalentes a los resultados más representativos de la integral de Lebesgue, ahora en la sección 4.7, nos dedicamos a presentar los análgos al primer y segundo teorema fundamental del cálculo pero adecuados para la integral de Lebsegue. Usualmente, este tipo de resultados no son incluidos en un primer curso de teoría de la media, por lo que puede que el alumno no llegue siquiera a enterarse de la existencia de dichos resultados, provocando que su formación en esta área quede inconclusa. Es aquí donde radica el segundo punto princpal a tratar en esta tesis. Con los contenidos que se presentan en este capítulo, las personas que ya han tomado un curso de teoría de la medida tendrán una visión mucho más amplia y rica de las integrales de Lebesgue y Riemann.

En el último capítulo de este documento se muestra una prueba alternativa a la brindada en la sección 3.4 del lema de Cousin propuesta por el autor de este trabajo.

2. Conceptos preliminares

En la primera sección de este capítulo se hará un repaso rápido de las propiedades más importantes del supremo e ínfimo. En la segunda, se introducirá el concepto de sucesiones de funciones y las nociones más basicas de estas.

2.1. Propiedades del ínfimo y supremo

Antes de empezar a definir lo que significa la integral de una función, se tiene que hablar de la propiedad más importante de los números reales, la propiedad de la cota superior mínima.

A coninuación, se define lo que significa una cota superior.

Definición 1. Sea A un subconjunto de los números reales. Si existe un número real x que cumpla con que $a \leqslant x$ para todo $a \in A$, entonces se dice que A está acotado superiormente y a x se le llama cota superior del conjunto A.

Está claro que al existir una cota superior x de un conjunto A, automáticamente existen una infinidad de cotas superiores de A. Sin embargo, entre todas esas cotas superiores, existe una que es especial, a la cual se le denomina cota superior mínima.

Definición 2. Sea A un subconjunto de los números reales acotado superiormente. Se le denomina cota superior mínima de A, al número real α que cumpla con que

i) Es una cota superior de A.

ii) Si y es una cota superio de A, entonces $\alpha \leqslant y$.

Es común el uso de la palabra supremo para referirse a la cota superior mínima de un conjunto y utilizar el símbolo $\sup A$ para denotarlo.

Es prácticamente inmediato de la definición que, en caso de existir el supremo de un conjunto, este será único. Lo interesante en realidad, es la cuestión de la existencia de dicho supremo en sí. Con respecto a esto, tenemos el siguiente axíoma de los números reales.

Axioma (de la cota superior mínima). Todo subconjunto de los números reales, no vacío y acotado superiormente, tiene una cota superior mínima.

La no vacuidad del conjunto es necesario, pues al no tener elementos, todo número real es cota superior del vacío, haciendo que la cota superior mínima no exista.

De manera similar a como se ha hecho con el supremo, se dice que un número real x es cota inferior de $A \subseteq \mathbb{R}$, si $x \leqslant a$ para todo a en A y diremos que β es la cota inferior máxima de A, si cumple que

i) Es una cota inferior de A.

ii) Si y es una cota inferior de A, entonces $y \leqslant \beta$.

Generalmente se le denomina ínfimo a esta cota inferior máxima y se denota por ínf A.

Existe una estrecha relación entre el ínfimo y el supremo, razón por la cual no existe como tal una propiedad de la cota inferior máxima, si no que, más bien, esta puede ser demostrada a partir de la propiedad de la cota superior mínima, como veremos más adelante en IV) del teorema .

Estos conceptos de ínfimo y supremo, juegan un papel importantísimo en el desarrollo del cálculo y en particular, en la integral de Darboux que veremos más adelante, razón por la cual, es necesesario conocer algunas de sus propiedades. Con este motivo, definamos los siguientes conjuntos.

Definición 3. Sean $A_1, A_2, \ldots, A_n$ subconjuntos de los números reales y k un número real, entonces definimos los siguientes conjuntos

$$A_1 + A_2 + \ldots + A_n := \{a_1 + a_2 + \ldots + a_n \ : \ a_1 \in A_1,\, a_2 \in A_2,\, \ldots,\, a_n \in A_n\}$$
$$kA := \{ka \ : \ a \in A\}$$

***Observación* 1.** *Resulta fácil probar que* $k\,(A_1 + \ldots + A_n) = kA_1 + \ldots + kA_n$.

Teorema 1. *Sean A, B, C y D subconjuntos de $\mathbb{R}$ no vacios y k un número real positivo. Supongamos que A y B están acotados superiormente y C y D están acotados inferiormente, entonces*

I) *kA está acotado superiormente y* $\sup(kA) = k \sup A$.

II) *kC está acotado inferiormente e* $\inf(kC) = k \inf C$.

III) *Si $a \leqslant b$ para todo a en A y b en B, entonces* $\sup A \leqslant \inf B$.

IV) *$-C$ está acotado superiormente e* $\inf C = -\sup(-C)$.

V) *$A + B$ está acotado superiormente y* $\sup(A + B) = \sup A + \sup B$.

VI) *$C + D$ está acotado inferiormente e* $\inf(C + D) = \inf C + \inf D$.

VII) *Si $A \subseteq B$, entonces* $\sup A \leqslant \sup B$.

VIII) *Si $C \subseteq D$, entonces* $\inf D \leqslant \inf C$.

Demostración. Únicamente demostraremos V).

V) Comencemos demostrando el caso cuando $A = \{a\}$. Nótese que en este caso $\sup A = a$. Además, está claro que $A + B$ es acotado superiormente, por ejemplo, $a + \sup B$ es una cota superior de este conjunto. Así pues, llamémosle α al supremo de $A + B$, entonces se cumple que $a + b \leqslant \alpha$ para todo b en B, equivalentemente, $b \leqslant \alpha - a$ para todo b en B. Esto quiere decir que $\alpha - a$ es una cota superior del conjunto B. Vamos a demostrar que es la mínima cota superior. Supongamos que no es así y que existe un número γ tal que $b \leqslant \gamma < \alpha - a$ para todo b en B. Esto implica que $b + a \leqslant \gamma + a < \alpha$ para todo $b \in B$, lo cual es una contradicción, pues se supone que α es la cota superior mínima del conjunto $A + B$ y aquí hemos llegado a que $\gamma + a$ es una cota superior más pequeña que α. Con esto concluimos que $\sup B = \sup(A + B) - a$, que es lo que queríamos demostrar.

Ahora demostremos el caso general. Comencemos notando que $A+B$ está acotado. Esto es inmediato, ya que $a+b \leqslant \sup A + \sup B$, por lo que $\sup A + \sup B$ es una cota superior del conjunto $A+B$. Pero el supremo es la cota superior más pequeña, esto implica que $\sup(A+B) \leqslant \sup A + \sup B$.

Ahora, llamémosle α al $\sup(A+B)$ y demostremos que $\sup A + \sup B \leqslant \alpha$, con lo cual terminaríamos la demostración.

Ya que α es el supremo de $A+B$, entonces se cumple que $a+b \leqslant \alpha$ y por lo tanto $a \leqslant \alpha - b$ para todo $a \in A$ y $b \in B$.

Gracias a III), podemos decir que $\sup A \leqslant \text{ínf}\left(\{\alpha - b : b \in B\}\right)$, o bien que $\sup A \leqslant \text{ínf}\left(\{\alpha\} + (-B)\right)$. Además, gracias a IV) y a la observación 1, sabemos que

$$\text{ínf}\left(\{\alpha\} + (-B)\right) = -\sup\left(-\{\alpha\} + B\right).$$

Pero este es el caso que demostramos al inicio, cuando uno de los conjuntos tiene un solo elemento, por lo que

$$-\sup\left(-\{\alpha\} + B\right) = -\left(\sup\left(-\{\alpha\}\right) + \sup\left(B\right)\right) = \alpha - \sup B.$$

De esta manera llegamos a que $\sup A \leqslant \alpha - \sup B$, terminando así la demostración.

■

Algo muy importante que debemos notar es que, a pesa de que solo demostramos V) y VI) para dos conjuntos, en realidad podemos extender esto a n conjuntos no vacios y acotados mediante un argumento inductivo. Con lo cual, podemos concluir que, si $A_1, A_2, \ldots, A_n$ son conjuntos no vacíos y acotados superiormente, entonces $A_1 + A_2 + \ldots + A_n$ está acotado superiormente y

$$\sup\left(A_1 + A_2 + \ldots + A_n\right) = \sup A_1 + \sup A_2 + \ldots + \sup A_n.$$

De manera anáolga, si $B_1, B_2, \ldots, B_n$ son conjuntos no vacíos y acotados inferiormente, entonces $B_1 + B_2 + \ldots + B_n$ está acotado inferiormente y

$$\text{ínf}\left(B_1 + B_2 + \ldots + B_n\right) = \text{ínf}\, B_1 + \text{ínf}\, B_2 + \ldots + \text{ínf}\, B_n.$$

Lo cual, junto con I) y II) del teorema 1, nos permite dar el siguiente corolario.

Corolario 1. *Sean $A_1, \ldots, A_n$ y $B_1, \ldots, B_n$ subconjuntos de $\mathbb{R}$ no vacios, de tal manera que A_i es acotado superiormente y B_i es acotado inferiormente para toda $1 \leqslant i \leqslant n$. Sean $k_1, \ldots k_n$ constantes positivas. Entonces, se cumple que*

I) $\sup\left(k_1A_1 + \ldots + k_nA_n\right) = k_1 \sup A_1 + \ldots + k_n \sup A_n$.

II) $\text{ínf}\left(k_1A_1 + \ldots + k_nA_n\right) = k_1\, \text{ínf}\, A_1 + \ldots + k_n\, \text{ínf}\, A_n$.

Continuemos ahora, con una propiedad bien conocida y muy útil del supremo y el ínfimo.

Teorema 2. *Sea A un conjunto no vacío y acotado superiormente y B un conjunto no vacío y acotado inferiormente, entonces*

I) *Para todo* $\varepsilon > 0$ *existe un* $a \in A$ *tal que* $\sup A - \varepsilon < a$.

II) *Para todo* $\varepsilon > 0$ *existe un* $b \in B$ *tal que* $b < \text{ínf}\, B + \varepsilon$.

Demostración.

I) Demostrémoslo por contradicción. Supongamos que existe un $\varepsilon > 0$ tal que para todo $a \in A$ se cumple que $a \leqslant \sup A - \varepsilon$. Pero esto es una contradicción, pues $\sup A - \varepsilon$ es cota superior de A y además es una cota superior menor a $\sup A$, contradiciendo el hecho de que $\sup A$ es la cota superior mínima.

II) la demostración es análoga a la de I). ■

El inciso I) del teorema que acabamos de ver, básicamente nos dice que, si al supremo del conjunto le restamos un poco, podremos encontrar un elemento del conjunto que sea mayor a esta resta, haciendo que el supremo siempre esté muy "pegado" por arriba al conjunto.

En cambio, el inciso II), nos dice que el ínfimo siempre está muy "pegado" por abajo al conjunto.

Con esto concluimos lo que hay que saber del supremo e ínfimo, para pasar a una proposición sencilla de demostrar, pero que será utilizada en repetidos argumentos a lo largo de los próximos capítulos. A su vez, este breve argumento se basa en una propiedad más conocida del valor absoluto, si p y q son número reales tales que $p > 0$, entonces $-p \leqslant q \leqslant p$ si y solo si $\mid q \mid \leqslant p$.

Proposición 1. *Sean* x, y, a *y* b *números reales tales que* $x \leqslant a \leqslant y$ *y* $x \leqslant b \leqslant y$, *entonces*

$$\mid a - b \mid \leqslant y - x$$

Demostración. Ya que $a \leqslant y$, esto implica que $a - b \leqslant y - b$. Pero $y - b \leqslant y - x$ y por lo tanto $a - b \leqslant y - x$.

De manera similar, ya que $b \leqslant y$, esto implica que $b - a \leqslant y - x$, o equivalentemente $x - y \leqslant a - b$.

Por la propiedad del valor absoluto mencionada antes de este teorema , podemos concluir que $\mid a - b \mid \leqslant y - x$ ■

2.2. Sucesiones de funciones

En esta breve sección, se introduciran los conceptos básicos de sucesiones de funciones, los suficientes como para comprender los resultado principales del capítulo de la integral de Riemann, el teorema de la convergencia monótona y el teorema de la convergencia acotada. En las referencias **[15]** y **[1]** se ofrece una versión más completa de la teoría de sucesiones de funciones.

Para introducir el concepto de sucesiones de funciones, es buena idea partir del concepto de sucesiones de números reales.

Definición 4. Se define como sucesión de números reales, a cualquier función $\Phi : \mathbb{N} \longrightarrow \mathbb{R}$. Generalmente, al momento de trabajar con sucesiones de números reales simplemente hacemos referencia a la imagen de la función Φ, de manera que, si $\phi_n = \Phi(n)$, en lugar de referirnos diréctamente a la función Φ, denotamos a la sucesión simplemente como (ϕ_n).

Así mismo, diremos que una sucesión de números reales (ϕ_n) converge a un número real L, si para todo $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que si $n \geqslant N$, entonces

$$|\phi_n - L| < \varepsilon.$$

y lo denotamos como

$$\lim_{n\to\infty} \phi_n = L.$$

Con respecto a las sucesiones convergentes de números reales, tenemos el siguiente teorema.

Teorema 3. *Si (x_n) es una sucesión de números reales tal que*

$$\lim_{n\to\infty} x_n = x.$$

Entonces, la sucesión (s_n) definida por $s_n = \sup\{x_n, x_{n+1}, x_{n+2}, \ldots\}$, es decreciente y además

$$\lim_{n\to\infty} s_n = x.$$

Demostración. Que (s_n) sea decreciente es una consecuencia inmediata del inciso VIII) del teorema 1.

Ahora, sea $\varepsilon > 0$ arbitrario y sea $N \in \mathbb{N}$ tal que si $n \geqslant N$, entonces

$$|x_n - x| < \varepsilon$$

Esto implica que

$$-\varepsilon + x < x_n < \varepsilon + x$$

para todo $n \geqslant N$. Por lo tanto

$$-\varepsilon + x < \sup\{x_n, x_{n+1}, x_{n+2}, \ldots\} \leqslant \varepsilon + x.$$

para todo $n \geqslant N$. Esto quiere decir que, si $n \geqslant N$ entonces

$$|s_n - x| \leqslant \varepsilon.$$

■

Con la motivación de haber visto como se definen las sucesiones de números reales, ¿Qué pasaría si la función Φ no tuviera como codominio a los números reales?, ¿Qué pasaría si en su lugar, tuviera como codominio el conjunto de funciones que van de los números reales a los números reales?, es decir

$$\Phi : \mathbb{N} \longrightarrow \{f : \mathbb{R} \longrightarrow \mathbb{R}\}.$$

Entonces, se estaría hablando de una sucesión de funciones.

Definición 5. Sea A un conjuntos cualquiera. Una sucesión de funciones definidas en A, es una función Φ tal que

$$\Phi : \mathbb{N} \longrightarrow \{f : \mathbb{A} \longrightarrow \mathbb{R}\}.$$

Tal y como se hizo con las sucesiones de números reales, para referirnos a una sucesión de funciones, únicamente se hará alusión a la imagen de la función Φ, es decir, si $f_n = \Phi(n)$, denotamos la sucesión por (f_n). En este caso, diremos que (f_n) es una sucesión de funciones definidas en A.

El dominio de la sucesión de funciones es algo que será especificado antes de trabajar con alguna sucesión. Este puede ser un intervalo acotado $[a, b]$ o un intervalo de la forma $[a, \infty)$. También se harán las distinciones respectivas para referirnos, ya sea, a una sucesión de números reales o a una sucesión de funciones.

Para visualizar mejor este concepto, se tiene el siguiente ejemplo.

Ejemplo 1. Sea f_n una sucesión de funciones definidas en el intervalo $[0, 1]$ de la siguiente manera

$$f_n(x) = n.$$

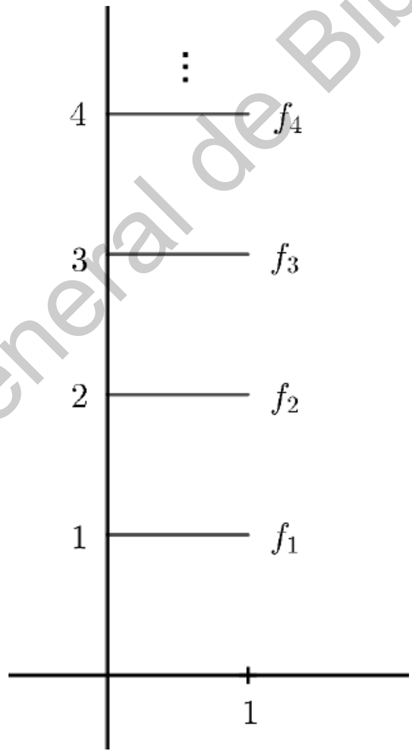

Figura 1: Sucesión del ejemplo 1.

Una manera visual de representar a las sucesiones de funciones es graficando algunos términos en el plano cartesiano como se muestra en la figura 1 .

Ejemplo 2. Un ejemplo un poco más interesante es el siguiente.

Sea (f_n) una sucesión de funciones definidas en $[0, 1]$ de la siguiente manera

$$f_n(x) = x^n.$$

En este ejemplo las funciones ya no son constantes, si no que f_1 es la función identidad, f_2 es una función cuadrática, etcétera. En la figura 2 se aprecia una representación de este ejemplo.

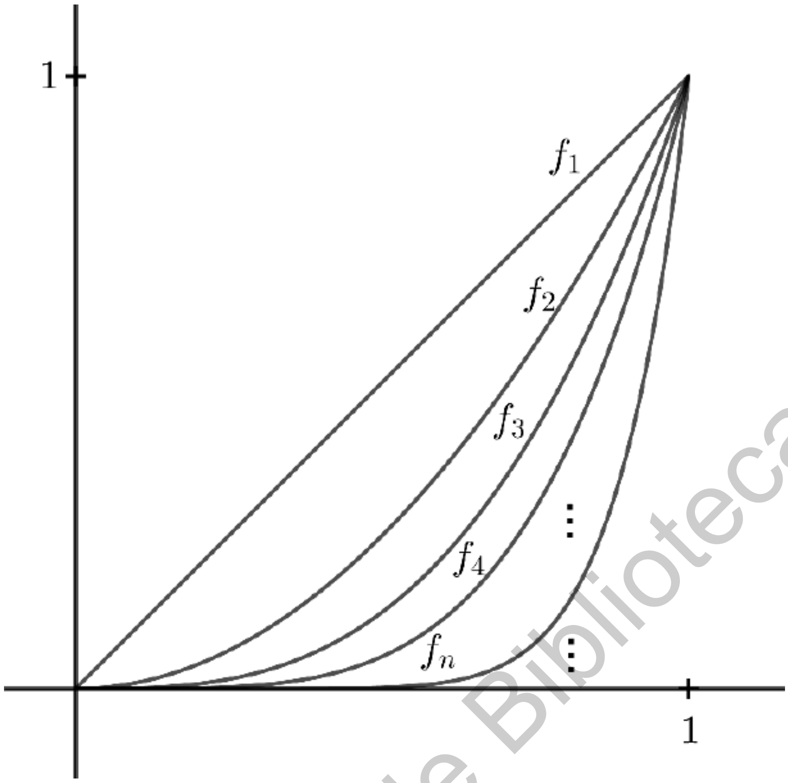

Figura 2: Sucesión del ejemplo 2.

Es importante resaltar una diferencia muy importante que ocurre entre los dos ejemplos que acabamos de ver. Mientras que en el segundo ejemplo, a medida que avanzamos a traves de los f_n, pareciera que la sucesión se "aproxima" a una función, en el primer ejemplo no ocurre esto, pues a conforme el índice aumenta, los términos f_n de la sucesión aumentan y aumentan. En realidad, estamos percibiendo el concepto de límite de una sucesión.

Definición 6. Sea (f_n) una sucesión de funciones definidas en $A \subseteq \mathbb{R}$. Diremos que la sucesión (f_n) converge puntualmente a la función $f : A \longrightarrow \mathbb{R}$, si para cada $x \in A$

$$\lim_{n\to\infty} f_n(x) = f(x)$$

en cuyo caso, a f se le denomina función límite o límite de la sucesión (f_n) y se denota por

$$\lim_{n\to\infty} f_n = f.$$

En otras palabras, la sucesión (f_n) converge a f, si para cada $x \in A$ y para todo $\varepsilon > 0$, existe un $N(x) \in \mathbb{N}$ tal que si $n \geqslant N(x)$, esto implique que

$$|f_n(x) - f(x)| < \varepsilon.$$

La razón para usar el símbolo $N(x)$ en lugar de simplemente N, es para hacer énfasis en que, dependiendo del x elegido, el número natural $N(x)$ puede variar.

El porque de llamarla convergencia puntual en lugar de simplemente decirle convergencia, es debido a dos motivos; de esta manera se hace énfasis en que la sucesión de números $(f_n(x))$ converge para todo punto en el conjunto de definición A. El segundo motivo es debido a que, en el caso de las sucesiones de funciones, existen otros tipos de convergencia.

Como ya se había comentado, esta sección solo dará la noción básica de las sucesiones de funciones, por lo que otros tipos de convergencia no formarán parte de este trabajo. Sin embargo, en **[15]** y **[1]** se puede encontrar una teoría más desarrollada de este tema.

Ahora, un ejemplo de convergencia puntual.

Ejemplo 3. Proponer una función límite para la sucesión del ejemplo 2 .

Se puede partir de la visualización que es brindada por la figura 2, donde se aprecia que, a medida que n aumenta, las f_n van siendo cada vez más cercanas a cero, a excepción de $x = 1$, punto en el cual $f_n(x) = 1$ para todo n.

Con esta motivación, proponemos a $f : [0,1] \longrightarrow \mathbb{R}$ definida por

$$f(x) = \begin{cases} 0 & \text{si } 0 \leqslant x < 1 \\ 1 & \text{si } x = 1 \end{cases}$$

como la función límite de la sucesión (f_n). Aunque aquí no se dará la demostración, este hecho no es difícil de probar.

A continuación, se definirá cuándo una sucesión de funciones se considera creciente o decreciente.

Definición 7. Sea (f_n) una sucesión de funciones definidas en $A \subseteq \mathbb{R}$. Diremos que la sucesión es

I) Creciente, si $f_n(x) \leqslant f_{n+1}(x)$ para todo $n \in \mathbb{N}$ y para todo $x \in A$.

II) Decreciente, si $f_{n+1}(x) \leqslant f_n(x)$ para todo $n \in \mathbb{N}$ y para todo $x \in A$.

III) Monótona, en caso de que la sucesión sea creciente o decreciente.

Hasta el momento, de los dos ejemplos de sucesiónes de funciones que se han presentado, el primero corresponde a una sucesión creciente de funciones y el segundo corresponde a una sucesión decreciente de funciones.

Un ejemplo de una sucesión que no es monótona es la sucesion dada por

$$f_n(x) = (-1)^n$$

definida para $x \in [0,1]$.

En el caso de sucesiones monótonas de funciones, tenemos la siguiente terminología.

Definición 8. Sea (f_n) una sucesion de funciones definidas en $A \subseteq \mathbb{R}$ que convergen puntualmente a una función límite f. Entonces diremos que

I) La sucesión (f_n) crece a f, si la sucesión es creciente.

II) La sucesión (f_n) decrece a f, si la sucesión es decreciente.

Por lo tanto, cuando digamos que una sucesión de funciones (f_n) decrece a una función f, estaremos diciendo implícitamente que, la sucesión de funciones es decreciente y además, la sucesión converge puntualmente a f.

Es inmediato de la definición ver que, si (f_n) es una sucesión de funciones definidas en A que crece a una función f, entonces $f_n \leqslant f$ para todo n.

De manera similar, si la sucesión de funciones decrece a f, entonces $f \leqslant f_n$ para todo n.

3. Integral de Riemann

3.1. Definición de la integral de Riemann

La teoría básica de integración presentada en esta y la siguiente sección, está basada en las ideas presentadas en [**16**].

Antes de comenzar a definir la integral de Riemann, será necesario definir un par de conceptos previos.

Definición 9. Sean a y b números reales con $a < b$. Una partición de $[a, b]$ es una colección finita de puntos $P = \{x_0, x_1, x_2, \ldots, x_n\}$ de $[a, b]$ que cumple con que $x_0 = a < x_1 < x_2 < \ldots < x_n = b$.

Tal y como su nombre lo indica, esta colección finita de puntos nos induce una manera de partir el intervalo $[a, b]$ en una cantidad finita de subintervalos formados por los puntos adyacentes en la partición.

Definición 10. La norma de una partición $P = \{x_0, x_1, x_2, \ldots, x_n\}$ se define como

$$|P| := max\,\{x_i - x_{i-1} : i = 1, 2, ..., n\}.$$

Es evidente que, si δ es un número positivo y $|P| < \delta$, entonces $x_i - x_{i-1} < \delta$ para cada $i = 1, 2, \ldots, n$.

Definición 11. Sea $f : [a, b] \to \mathbb{R}$ una función y $P = \{x_0, x_1, x_2, \ldots, x_n\}$ una partición de $[a, b]$. Si para cada $i = 1, 2, \ldots, n$ se elige un punto $t_i \in [x_{i-1}, x_i]$ entonces, a la suma

$$\sum_{i=1}^{n} f(t_i)(x_i - x_{i-1})$$

se le denomina suma de Riemann de f correspondiente a la particion P con puntos intermedios t_i y se denota por el símbolo $S(f, P, \{t_i\})$.

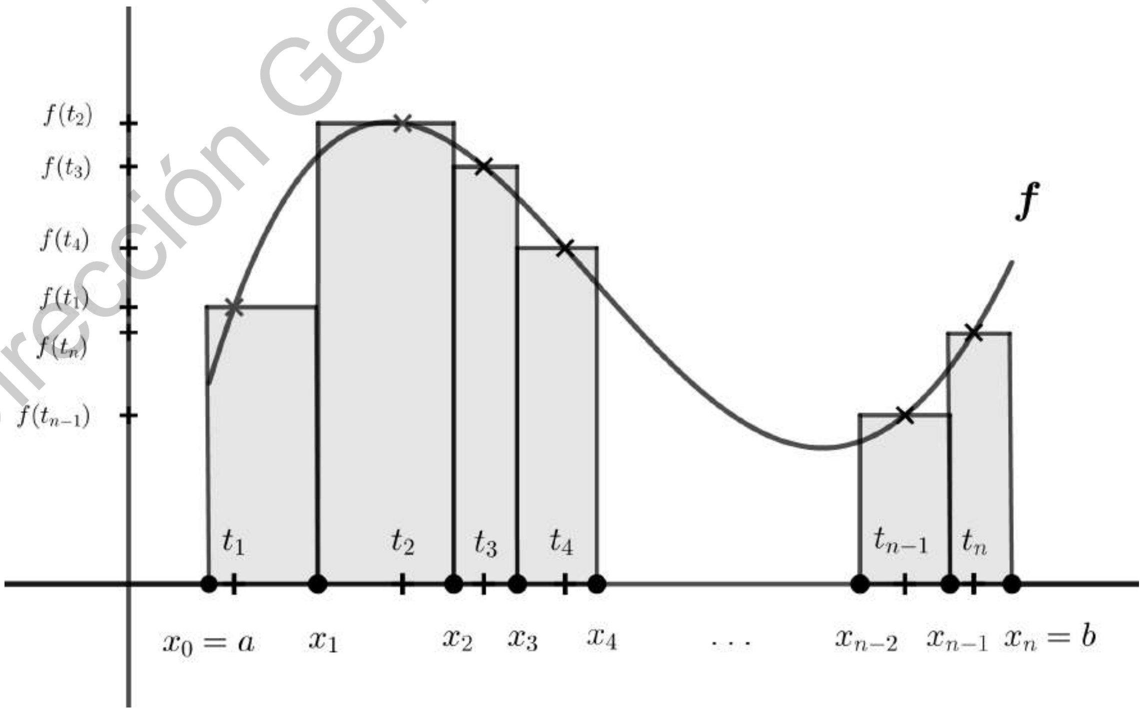

Figura 3: Sumas de Riemann

De la definición anterior notemos que si $f \geqslant 0$, entonces una suma de Riemann $S(f, P, \{t_i\})$ no es más que un intento por aproximar el área que se encuentra entre la gráfica de f y el eje x, tal y como se ilustra en la figura 3.

Al observar esta imagen, uno podría pensar que la precisión de la suma de Riemann como aproximación al área I bajo la curva, va a depender únicamente de la norma de la partición. Pues mientras más pequeña sea la norma de P, los rectángulos de altura $f(t_i)$ y base $x_i - x_{i-1}$ serán más finos y proporcionarán una mejor aproximación de I.

Sin embargo, no hay que olvidar el hecho de que la suma de Riemann depende también de los puntos t_i elegidos, haciendo que aún cuando se tiene la misma partición del intervalo, los valores de las sumas de Riemann varien dependiendo de la elección de estos puntos intermedios. Este inconveniente no se soluciona eligiendo una partición lo suficientemente fina como para que estas diferencias no representen un problema. Existirán funciones que sin importar cuan fina sea la partición, las sumas de Riemann no tienden a un valor en concreto, a estas las llamaremos funciones no integrables en el sentido de Riemann. Obviamente, a las funciones que si cumplan con estos requisitos, les llamaremos funciones Riemann integrables.

Definición 12. (Integral de Riemann) Sea $f : [a, b] \to \mathbb{R}$ una función acotada. Diremos que f es Riemann integrable si existe un número real I, tal que para todo $\varepsilon > 0$ existe un $\delta > 0$ de modo que si P es una partición de $[a, b]$ que cumple con $|P| < \delta$, entonces

$$|S(f, P, \{t_i\}) - I| < \varepsilon$$

para cualesquiera puntos t_i elegidos. Al número I se le conoce como integral de Riemann de f.

Tal y como lo habíamos discutido antes, solo serán integrables en el sentido de Riemann aquellas funciones que cumplan la condición de que a medida que la norma de la partición sea cada vez más pequeña, la $S(f, P, \{t_i\})$ va a converger a un único valor I, independientemente de los puntos t_i elegidos.

Esta es la primera definición formal que se dió de la integral de una función y proporcionó al fin el rigor que le faltaba al concepto de área bajo la curva. Fue dada por Rimann en 1867 **[13]**. A pesar de ello, esta no es la única definición de integral de Riemann, ni la más usada. De hecho, existen varias definiciones que son equivalentes y en algunos casos inclusive pueden llegar a resultar mas útiles.

3.2. Definición de la integral de Darboux

La definición de integral que se verá en esta sección, fue dada por Darboux en el año 1879 **[4]**. Es posterior a la definición de Riemann, pero presenta ideas muy interesantes que son útiles inclusive en cursos posteriores como teoría de la medida.

La teoría aquí presentada, no es completa y puede ser complementada en **[16]**.

Definición 13. Sea $f : [a.b] \longrightarrow \mathbb{R}$ una función acotada y $P = \{x_0, x_1, x_2, \ldots, x_n\}$ una partición de $[a, b]$. Para cada $i = 1, 2, \ldots, n$ definimos a

$$m_i = \inf \{f(t) : x_{i-1} \leqslant t \leqslant x_i\}$$

$$M_i = \sup \{f(t) : x_{i-1} \leqslant t \leqslant x_i\}.$$

Definimos a la suma inferior de f con respecto de P como

$$\sum_{i=1}^{n} m_i \left(x_i - x_{i-1}\right)$$

y la denotamos por $L\left(f, P\right)$. De manera similar, definimos a la suma superior de f con respecto de P es

$$\sum_{i=1}^{n} M_i \left(x_i - x_{i-1}\right)$$

y la denotamos por $U\left(f, P\right)$.

Se debe resaltar que esta vez, la suma inferior y superior dependen únicamente de la función y de la partición, a diferencia de las sumas de Riemann que también dependían de los puntos intermedios elegidos. Esto es algo crucial para entender el porqué se decide trabajar con la definición de integral que se dará a continuación, en lugar de simplemente quedarse con la definición de Riemann.

De manera visual, una suma inferior y una suma superior se verían como el área sombreada de la figura 4. Estas áreas están traslapadas y la de color fuerte repesenta la suma inferior, mientras que la de color más claro representa la suma superior.

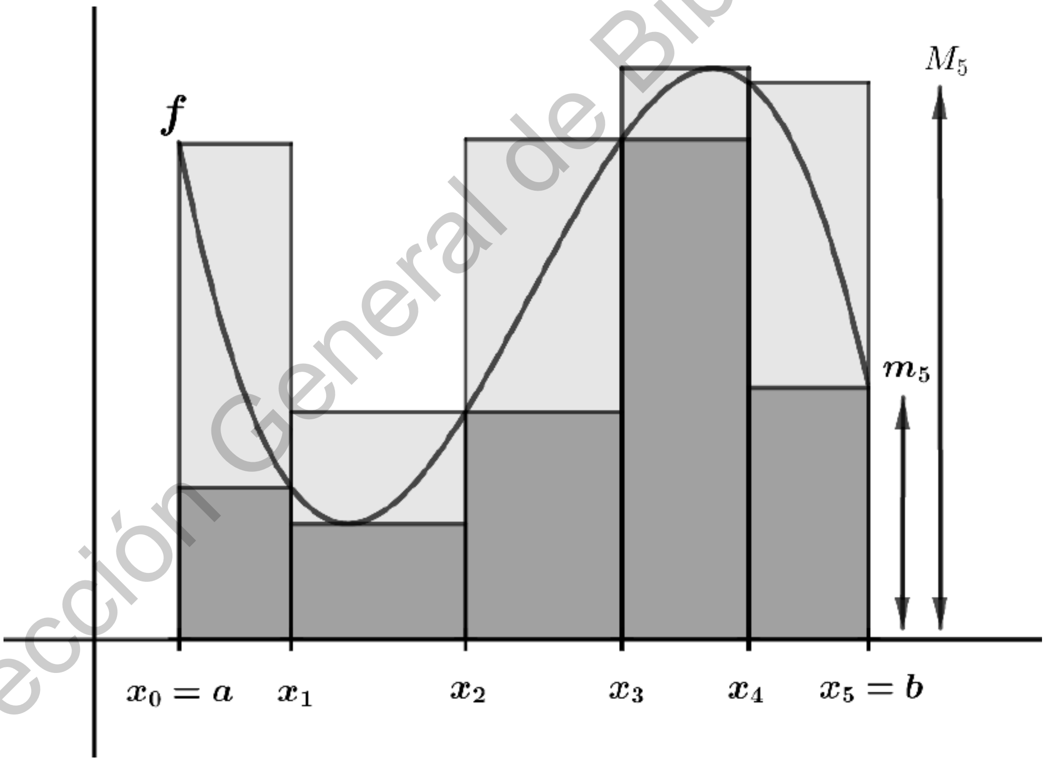

Figura 4: Suma superior e inferior.

A continuación, se inroducirá el concepto de refinamiento de una partición.

Definición 14. Sean P y Q dos particiones de un intervalo $[a, b]$. Diremos que Q es un refinamiento de P si se cumple que $P \subseteq Q$.

Se le denomina refinamiento debido a que Q tiene los mismos puntos de P y unos pocos más, haciendo que Q sea más " fina".

No es difícil ver que si P es una partición cualquiera, entonces siempre se cumple que $L(f,P) \leqslant U(f,P)$.

Un hecho un poco menos obvio, pero no muy difícil de demostrar, es el que involucra al refinamineto de una partición.

Teorema 4. *Sea $f:[a,b] \longrightarrow \mathbb{R}$ una función acotada y P y Q particiones de $[a,b]$ de modo que Q es un refinamiento de P. Entonces,*

$$L(f,P) \leqslant L(f,Q) \leqslant U(f,Q) \leqslant U(f,P).$$

Una demostración de este último teorema puede ser encontrada en **[16]** y lo que está indicando es que, a medida que se incluyan más puntos en la partición, las sumas inferiores van incrementando su valor, mientras que las sumas superiores lo disminuyen. De manera idónea, uno esperaría que, a medida que la cantidad de puntos aumente y la norma de las particion sea cada vez más pequeña, ambas sumas se aproximen al valor del área bajo la curva, una por abajo y la otra por arriba.

Entonces, basta con definir a la integral como el supremo de todas las sumas inferiores que existen, o como el ínfimo de las sumas superiores

Primero, habría que garantizar la existencia de dichas cantidades, y para ello, ambos conjuntos, el de las sumas inferiores y el de las sumas superiores, deben ser no vacios y estar acotados, el primero superiormente y el segundo inferiormente.

La no vacuidad está garantizada, pues para cualquier intervalo siempre existirá una partición y ya que la función es acotada, las sumas inferiores y superiores siempre están bien definidas.

Para demostrar que ambos conjuntos están acotados de manera debida en cada caso, se usa el siguiente hecho

Teorema 5. *Sea $f:[a,b] \longrightarrow \mathbb{R}$ una función acotada y P_1 y P_2 particiones cualesquiera de $[a,b]$. Entonces,*

$$L(f,P_1) \leqslant U(f,P_2)$$

Demostración. Consideremos a la partición $P = P_1 \cup P_2$.

Está claro que P es un refinamiento tanto de P_1 como de P_2. Por el teorema 4 podemos concluir que

$$L(f,P_1) \leqslant L(f,P) \leqslant U(f,P) \leqslant U(f,P_1)$$

y

$$L(f,P_2) \leqslant L(f,P) \leqslant U(f,P) \leqslant U(f,P_2).$$

Con lo cual $L(f,P_1) \leqslant U(f,P) \leqslant U(f,P_2)$. ■

Con este nuevo resultado, si fijamos P_1 y variamos a P_2, se puede notar que $L(f,P_1)$ es una cota inferior para el conjunto de las sumas superiores.

Si esta vez fijamos P_2 y variamos a P_1, se estará concluyendo que $U(f,P_2)$ es una cota superior para el conjunto de las sumas inferiores.

Ahora, la existencia del supremo de las sumas inferiores y el ínfimo de las sumas superiores está completamente justificada. Estamos listos para dar una segunda definición de integral.

Definición 15. (Integral de Darboux) Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función acotada. Definimos a la integral inferior de f como

$$\underline{\int_a^b} f := \sup \{L(f, P) : P \text{ es partición de } [a, b]\}$$

y a la integral superior de f como

$$\overline{\int_a^b} f := \text{ínf} \{U(f, P) : P \text{ es partición de } [a, b]\}.$$

Diremos que f es Darboux integrable o integrable en el sentido de Darboux, si se cumple que ambas integrales coinciden, en cuyo caso definimos a la integral de Darboux de f como

$$\int_a^b f := \underline{\int_a^b} f = \overline{\int_a^b} f.$$

También es común hacer uso de la notación

$$\int_a^b f(x)dx$$

para referirnos a la integral de Darboux.

Uno se podría preguntar, ¿por qué no se definió a la integral simplemente como el supremo de las sumas inferiores, o como el ínfimo de las sumas superiores?, ¿a caso no representan ambas el valor del área bajo la curva ? La respuesta a esto es no. A pesar de que ambas números siempre existen, no siempre valen lo mismo y es por ello que únicamente cuando sus valores coinciden, se dice que la función es integrable.

El ejemplo clásico de una función no integrable en el sentido de Darboux es la función $f : [0, 1] \longrightarrow \mathbb{R}$ dada por

$$f(x) = \begin{cases} 1 & \text{si } x \text{ es un número racional} \\ 0 & \text{si } x \text{ no es un número racinal} \end{cases}$$

Uno puede convencerse de este hecho siemplemente cálculando una suma superior y una suma inferior cualquiera y viendo que pasa con el ínfimo y el supremo.

La siguiente observación es una consecuencia inmediata del inciso III) del teorema 1, enunciado en el capítulo 2.

Observación 2. *Si f es una función acotada, entonces siempre se cumple que*

$$\underline{\int_a^b} f \leqslant \overline{\int_a^b} f.$$

El siguiente teorema da algunas de las propiedades básicas que cumple la integral de Darboux. La demostración de dichas propiedades puede ser encontrada con facilidad en la mayoría de libros de cálculo con cierto rigor matemático, por ejemplo, en [**16**].

Teorema 6. *Sean f y g dos funciones acotadas definidas en $[a, b]$ y k una constante real. Entonces*

I) *Si f y g son Darboux integrables, entonces $f + kg$ también lo es y se cumple que*

$$\int_a^b (f + kg) = \int_a^b f + k \int_a^b g.$$

II) *Si $f \leqslant g$, entonces*

$$\underline{\int_a^b} f \leqslant \underline{\int_a^b} g$$

y

$$\overline{\int_a^b} f \leqslant \overline{\int_a^b} g.$$

En particular, si f y g son integrables, entonces

$$\int_a^b f \leqslant \int_a^b g.$$

III) *Si c es un número entre a y b, entonces*

$$\underline{\int_a^c} f + \underline{\int_c^b} f = \underline{\int_a^b} f$$

y

$$\overline{\int_a^c} f + \overline{\int_c^b} f = \overline{\int_a^b} f.$$

En particular, si f es Darboux integrable, entonces

$$\int_a^c f + \int_c^b f = \int_a^b f.$$

IV) *Si f es Darboux integrable, entonces $|f|$ también lo es.*

V) *Siempre se cumple que*

$$\underline{\int_a^b (f + k)} = \underline{\int_a^b} f + k\,(b - a)$$

y

$$\overline{\int_a^b (f + k)} = \overline{\int_a^b} f + k\,(b - a)\,.$$

El inciso I del teorema anterior en realidad involucra dos propiedades de la integral de Darboux, la de separar la integral de la suma de dos funciones en la suma de las integrales y la de conmutar constantes con el símbolo de la integral, claro, en caso de que las funciones involucradas sean Darboux integrables. En el inciso V también ocurre algo interesante, pues en el caso de la integral inferior, en general, no siempre se cumple que la integral inferior de la suma de dos funciones, es igual a la suma de las integrales inferiores, pero si una de las funciones es una constante, entonces la igualdad sí es válida. No se ha de olvidar que en el caso de las funciones constantes, estas son siempre integrables, por lo que su integral inferior coincide con su integral. Una situación similar ocurre con la integral superior. Este inciso V es una consecuencia inmediata de la siguiente proposición y del corolario 1 del primer capítulo.

Proposición 2. *Sea $f : [a,b] \longrightarrow \mathbb{R}$ una función acotada y k una constante. Entonces,*

$$L(f+k,P) = L(f,P) + k(b-a)$$

y

$$U(f+k,P) = U(f,P) + k(b-a).$$

Demostración. Únicamente demostraremos el caso de las sumas inferiores.

Por la parte II) del corolario 1 sabemos que

$$L(f+k,P) = \sum_{i=1}^{n} (m_i + k)(x_i - x_{i-1})$$

donde

$$m_i = \text{ínf}\{f(t) : x_{i-1} \leqslant t \leqslant x_i\}.$$

Por lo tanto

$$L(f+k,P) = \sum_{i=1}^{n} m_i (x_i - x_{i-1}) + k(b-a).$$

■

Con la intergal de Darboux, ya se tendrían dos definiciones para el mismo concepto, pero ¿cuál resulta mejor?, ¿cuáles son las diferencias entre considerar una u otra?, ¿existen funciones que son integrables en el sentido de Riemann, pero no lo son en el sentido de Darboux, o viceversa? Respondiendo la última pregunta; no existen funciones que sean integrables con una definición, pero con la otra no. Es decir, f es Riemann integrable si y solo si f es Darboux integrable (este hecho es demostrado más adelante en el teorema 8) . Por lo tanto, la diferencia entre usar una u otra radica simplemente en las comodidades y facilidades que brindan cada una.

Se debe tener presente que, al definir las sumas de Riemann y las sumas superiores e inferiores, se hizo especial incapié en que las primeras dependían de los puntos intermedios elegidos. Esto quiere decir que para cada partición hay una infinidad de sumas de Riemann a considerar, mientras que en el caso de la integral de Darboux, para cada partición se consideran únicamente dos sumas, la inferior y la superior. Por ello y por otras razones es que, en general, es más usada esta segunda definición para desarrollar la teoría de integración, lo cual no quiere decir que la primera definición no sea útil.

3.3. Equivalencia entre la integral de Darboux y la integral de Riemann

Esta sección comienza con un resultado bien conocido de la integral de Darboux y cuyo segundo inciso resulta especialmente útil para saber cúando una función es Darboux integrable sin necesidad de conocer el supremo de las sumas inferiores o el ínfimo de las sumas superiores que en ocasiones, pueden llegar a resultar muy díficiles de cálcular.

Teorema 7. *Sea $f : [a,b] \longrightarrow \mathbb{R}$ una función acotada. Entonces,*

I) *para todo* $\varepsilon > 0$ *existe una partición* P *de* $[a,b]$ *tal que*

$$\underline{\int_a^b} f - L(f,P) < \varepsilon$$

y

$$U(f,P) - \overline{\int_a^b} f < \varepsilon.$$

II) f *es Darboux integrable si y solo si para todo* $\varepsilon > 0$ *exise una partición* P *de* $[a,b]$ *tal que*

$$U(f,P) - L(f,p) < \varepsilon.$$

Demostración. Únicamente demostraremos el primer inciso. Los interesados pueden consultar [**16**] para ver una demostración del segundo inciso. Sea $\varepsilon > 0$ dado, aplicando I) del teorema 2, existe una partición P_1 tal que

$$U(f,P_1) < \overline{\int_a^b} f + \varepsilon.$$

Pues la integal superior es el ínfimo de las sumas superiores. De manera similar, por II) del mismo teorema 2 sabemos que existe una partición P_2 tal que

$$\underline{\int_a^b} f - \varepsilon < L(f,P_2).$$

Pues a la integral inferior es el supremo de las sumas inferiores. Si llamamos P a $P_1 \cup P_2$. Enotnces P es un refinamiento tanto de P_1 como de P_2 y por el teorema 4 de la sección anterior, tenemos que

$$\underline{\int_a^b} f - \varepsilon < L(f,P_2) \leqslant L(f,P) \leqslant U(f,P) \leqslant U(f,P_1) < \overline{\int_a^b} f + \varepsilon.$$

Con lo cual

$$\underline{\int_a^b} f - L(f,P) < \varepsilon$$

y

$$U(f,P) - \overline{\int_a^b} f < \varepsilon.$$

■

Ahora, se discutirá un poco de la relación que siempre se da entre las sumas de Riemann, las sumas inferiores y las sumas superiores. Si $f : [a,b] \longrightarrow \mathbb{R}$ es una función acotada y P es una partición de $[a,b]$, entonces se cumple que

$$L(f,P) \leqslant S(f,P,\{t_i\}) \leqslant U(f,P) \tag{1}$$

para cualquier suma de Riemann sin importar los puntos t_i elegidos. Esto no es difícil de demostrar pues

$$m_i \leqslant f(t_i) \leqslant M_i$$

para todo $i = 1, 2, \ldots, n$. Por lo que

$$m_i\left(x_i - x_{i-1}\right) \leqslant f\left(t_i\right)\left(x_i - x_{i-1}\right) \leqslant M_i\left(x_i - x_{i-1}\right)$$

para todo $i = 1, 2, \ldots, n$. Por lo tanto

$$\sum_{i=1}^{n} m_i\left(x_i - x_{i-1}\right) \leqslant \sum_{i=1}^{n} f\left(t_i\right)\left(x_i - x_{i-1}\right) \leqslant \sum_{i=1}^{n} M_i\left(x_i - x_{i-1}\right).$$

Además de esto, si f resulta ser Darboux integrable, entonces siempre se cumple que

$$L(f, P) \leqslant \int_a^b f \leqslant U(f, P) \tag{2}$$

pues al ser el supremo de todas las sumas inferiores, es mayor o igual a todas ellas, pero también es el ínfimo de todas las sumas superiores, por lo que es menor o igual a todas ellas. De (1) y (2) podemos concluir, gracias a la proposición 1 de los números reales demostrada en el capítulo 2, que si f es una función Darboux integrable y P es una partición dada, entonces

$$\left|\int_a^b f - S\left(f, P, \{t_i\}\right)\right| \leqslant U(f, P) - L(f, P) \tag{3}$$

para cualesquiera puntos t_i elegidos.

Si de alguna manera se consiguiera hacer que $U(f, P) - L(f, P)$ fuera tan pequeño como se quiera a partir de cierta norma de la partición, entonces se estaría demostrando que f es Riemann integrable y además, que la integral de Riemann es la misma que la integral de Darboux.

Se debe de tener claro que el apartado I del teorema 7 nos garantiza que para todo $\varepsilon > 0$ existe al menos una partición que cumple $U(f, P) - L(f, P) < \varepsilon$. Pero una partición en particular es insuficiente, pues se está buscando que a partir de cierta norma en las particiones, esta diferencia entre las sumas superiores e inferiores, sea tan pequeña como se quiera. Sin embargo, este es un buen comienzo y de hecho, será parte importante en la demostración del siguiente resultado.

Lema 1. *Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función acotada. Entonces para todo $\varepsilon > 0$, existe un $\delta > 0$ tal que*

$$\underline{\int_a^b} f - L(f, P) < \varepsilon$$

y

$$U(f, P) - \overline{\int_a^b} f < \varepsilon$$

siempre y cuando $|P| < \delta$. En particular, si f es Darboux integrable, entonces para todo $\varepsilon > 0$, existe un $\delta > 0$ tal que

$$U(f, P) - L(f, P) < \varepsilon.$$

Demostración. Únicamente demostraremos la primera desigualdad con la integral inferior, la demostración para la integral superior es análoga.

Supóngase que f es una función no negativa y al final haremos un comentario que permitirá extender el resultado para funciones acotadas arbitrarias.

Sea $\varepsilon > 0$ dado. Por el teorema 7, sabemos que existe por lo menos una partición $P_0 = \{x_0, x_1, \ldots, x_n\}$ que cumple con

$$\underline{\int_a^b} f - L(f, P_0) < \frac{\varepsilon}{2}.$$

Esta partición podría proporcionarnos una primera aproximación al delta que buscamos. Tal vez, si la norma de las particiones son más pequeñas que el minímo de los $x_i - x_{i-1}$ de esta partición P_0, entonces la resta de la integral inferior menos la suma inferior también menor que ε. Exploremos pues, esta posibilidad.

Supongamos que $P_1 = \{y_1, y_2, \ldots, y_m\}$ es una partición de $[a, b]$ con $|P_1| < \text{mín}\{x_i - x_{i-1} \mid i = 1, 2, \ldots, n\}$.

Si recordamos un poco, la norma de una partición es el valor más grande de los $y_i - y_{i-1}$, por lo que la partición P_0 y P_1 deberián lucir como en la figura 5 (los elementos de P_0 están representados con puntos, mientras que los de P_1 están representados con rayas).

Para no perderse en la notación, sería bueno localizar algunos puntos importantes de la partición P_1, los cuales definimos a continuación.

$$y_{k_i} = \text{máx}\{y_j : y_j \in P_1 \ \wedge \ y_j \leqslant x_i\}$$

para $i = 1, 2, \ldots, n-1$.

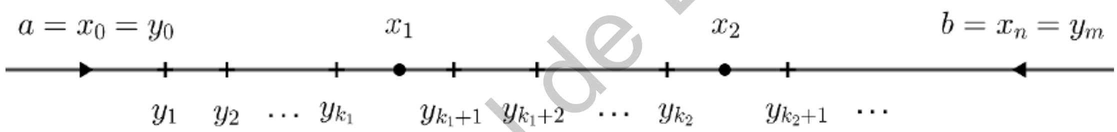

Figura 5: Particiónes P_0 y P_1

Una vez localizados estos puntos, hablemos de los m_i y M_i de cada partición. Para hacer una distinción entre ellos, denotemos por

$$m_i(P_0) := \text{ínf}\{f(t) : x_{i-1} \leqslant t \leqslant x_i\}$$

es decir, los m_i de la partición P_0.

De manera similar

$$m_i(P_1) := \text{ínf}\{f(t) : y_{i-1} \leqslant t \leqslant y_i\}.$$

Ahora, notemos que $m_1(P_0) \leqslant m_1(P_1), m_2(P_1), \ldots, m_{k_1}(P_1)$. Esto, gracias al inciso VIII) del teorema 1 que vimos en el capítulo 2, propiedades del ínfimo y supremo.

De aquí, se sigue que

$$\sum_{y_0 \leqslant y_i \leqslant y_{k_1}} m_i(P_1)(y_i - y_{i-1}) \geqslant \sum_{y_0 \leqslant y_i \leqslant y_{k_1}} m_1(P_0)(y_i - y_{i-1}) = m_1(P_0) \sum_{y_0 \leqslant y_i \leqslant y_{k_1}} y_i - y_{i-1}$$

pero la de la derecha es una suma donde se eliminan todos los términos, menos el primero y el último. Por lo que

$$\sum_{y_0 \leqslant y_i \leqslant y_k} m_i(P_1)(y_i - y_{i-1}) \geqslant m_1(P_0)(y_{k_1} - y_0).$$

Este razonamiento puede ser extendido para los demás puntos de la partición. Por ejemplo, para los puntos desde y_{k_1+1} hasta y_{k_2} tendríamos que

$$\sum_{y_{k_1+1}\leqslant y_i\leqslant y_{k_2}} m_i\left(P_1\right)\left(y_i - y_{i-1}\right) \geqslant m_2\left(P_0\right)\left(y_{k_2} - y_{k_1+1}\right).$$

Aunque, como se ilustra en la figura 6, para el cálculo de $L\left(f, P_1\right)$ aún necesitamos considerar los valores de $m_{k_i}\left(P_1\right)\left(y_{k_i+1} - y_{k_i}\right)$.

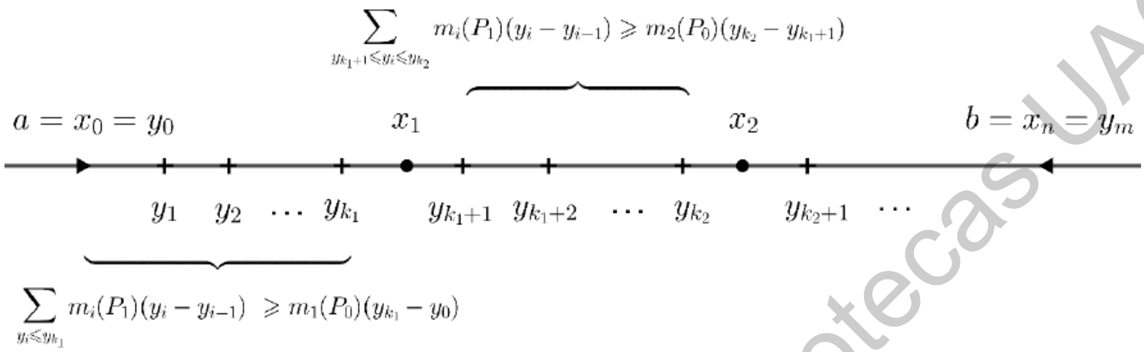

Figura 6: Comparación entre $L\left(f, P_0\right)$ y $L\left(f, P_1\right)$.

Entonces, podemos escribir $L\left(f, P_1\right)$ como

$$\begin{aligned} L\left(f, P_1\right) = &\sum_{y_0\leqslant y_i\leqslant y_{k_1}} m_i\left(P_1\right)\left(y_i - y_{i-1}\right) + m_{k_1+1}\left(P_1\right)\left(y_{k_1+1} - y_{k_1}\right) + \\ &\sum_{y_{k_1+1}\leqslant y_i\leqslant y_{k_2}} m_i\left(P_1\right)\left(y_i - y_{i-1}\right) + m_{k_2+1}\left(P_1\right)\left(y_{k_2+1} - y_{k_2}\right) + \\ &\vdots \\ &\sum_{y_{k_{n-1}+1}\leqslant y_i\leqslant y_m} m_i\left(P_1\right)\left(y_i - y_{i-1}\right). \end{aligned}$$

Luego

$$\begin{aligned} L\left(f, P_1\right) \geqslant &m_1\left(P_0\right)\left(y_{k_1} - y_0\right) + m_{k_1+1}\left(P_1\right)\left(y_{k_1+1} - y_{k_1}\right) + \\ &m_2\left(P_0\right)\left(y_{k_2} - y_{k_1+1}\right) + m_{k_2+1}\left(P_1\right)\left(y_{k_2+1} - y_{k_1}\right) + \\ &\vdots \\ &+ m_n\left(P_0\right)\left(x_n - x_{n-1}\right). \end{aligned}$$

Se sigue que

$$\begin{aligned} L\left(f, P_1\right) \geqslant &m_1\left(P_0\right)\left(x_1 - y_0\right) + m_{k_1+1}\left(P_1\right)\left(y_{k_1+1} - y_{k_1}\right) + \\ &m_2\left(P_0\right)\left(x_2 - x_1\right) + m_{k_2+1}\left(P_1\right)\left(y_{k_2+1} - y_{k_1}\right) + \\ &\vdots \\ &+ m_n\left(P_0\right)\left(x_n - x_{n-1}\right). \end{aligned}$$

Pues recordemos que f es no negativa, por lo tanto cada $m_i(P_0)$ es mayor que cero y lo único que estamos haciendo es multiplicarlos por algo más grande. Pero aquí ya aparece la suma inferior de f con recpecto de P_0 y tenemos que

$$\begin{aligned} L(f, P_1) \geqslant & L(f, P_0) + \\ & m_{k_1}(P_1)(y_{k_1+1} - y_{k_1}) + m_{k_2}(P_1)(y_{k_2+1} - y_{k_1}) + \dots \\ & + m_{k_{n-1}}(P_1)\left(y_{k_{n-1}+1} - y_{k_{n-1}}\right) \end{aligned} \tag{4}$$

Esto es lo mejor que se puede conseguir si solo suponemos que $|P_1| < \text{mín}\{x_i - x_{i-1} \mid i = 1, 2, \dots, n\}$. Pero ya casi hemos terminado, solo nos falta ajustar un poco más el valor de delta para acotar estos últimos términos que se resisten.

Esto se puede hacer simplemente observando que en (4) hay $n-1$ términos que hay que acotar. Además, recordemos que f es una función acotada, por lo que si definimos a M como

$$M := \sup\{f(t) : a \leqslant t \leqslant b\}$$

y elegimos a $\delta = \text{mín}\left\{|P_1|, \frac{\varepsilon}{2nM}\right\}$, entonces tendríamos que se cumple (4) y aún más

$$\begin{aligned} L(f, P_1) \geqslant & L(f, P_0) + \\ & m_{k_1}(P_1)(y_{k_1+1} - y_{k_1}) + m_{k_2}(P_1)(y_{k_2+1} - y_{k_1}) + \dots \\ & + m_{k_{n-1}}(P_1)\left(y_{k_{n-1}+1} - y_{k_{n-1}}\right) \\ & > L(f, P_0) - M\delta - M\delta - \dots - M\delta \\ & = L(f, P_0) - M(n-1)\delta \\ & > L(f, P_0) - \frac{\varepsilon}{2}. \end{aligned}$$

De esta manera, llegamos a que

$$0 \leqslant \underline{\int_a^b} f - L(f, P_1) < \underline{\int_a^b} f - L(f, P_0) + \frac{\varepsilon}{2} < \varepsilon$$

siempre y cuando $|P_1| < \delta$.

Para el caso cuando f sea una función acotada arbitraria, siempre podremos encontrar una constante $C > 0$ tal que $f + C$ sea una función positiva y por lo tanto, podemos aplicar lo que acabamos de demostrar a $f + C$ y decir que para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que si $|P| < \delta$, entonces

$$\underline{\int_a^b} (f + C) - L(f + C, P) < \varepsilon.$$

Pero gracias a la parte (5) del teorema 6 y a la proposición 2, sabemos que

$$\begin{aligned} \underline{\int_a^b} (f + C) - L(f + C, P) &= \left(\underline{\int_a^b} f + C(b-a)\right) - (L(f, P) + C(b-a)) \\ &= \underline{\int_a^b} f - L(f, P). \end{aligned}$$

■

Con este lema, ya se puede demostrar la equivalencia entre la integral de Riemann y la de Darboux.

Teorema 8. ***(De equivalencia de las integrales)*** *Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función acotada. f es Darboux integrable si y solo si f es Riemman integrable en cuyo caso, ambas integrales coinciden.*

Demostración. Como ya discutimos antes, la demostración de que si la función es Darboux integrable, enonces es Riemann integrable y las integrales coinciden, es consecuencia de (3) y del lema 1.

Ahora, supongamos que f es una función Riemann integrable y llamémosle I al valor de la integral. Vamos a demostrar que I queda por debajo de la integral inferior de f. Sea $\varepsilon > 0$ dado, entonces existe un $\delta > 0$ de modo que

$$\left|\sum_{i=1}^{n} f(t_i)\,(x_i - x_{i-1}) - I\right| < \varepsilon$$

siempre y cuando $|P| < \delta$ y los puntos t_i estén entre x_{i-1} y x_i para todo $i = 1, 2, \ldots, n$, o equivalentemente

$$-\varepsilon + I < \sum_{i=1}^{n} f(t_i)\,(x_i - x_{i-1}) < \varepsilon + I \tag{5}$$

siempre que $|P|$ sea menor que δ y los puntos t_i estén entre x_{i-1} y x_i para todo $i = 1, 2, \ldots, n$.

Tomemos una partición $P_0 = \{x_0 = a, x_1, \ldots, x_n = b\}$ con $|P| < \delta$ y pongamos atención al siguiente conjunto

$$\left\{\sum_{i=1}^{n} f(t_i)\,(x_i - x_{i-1})\ : t_i \in [x_{i-1}, x_i]\right\}.$$

Este conjunto describe a todas las sumas de Riemann de f posibles correspondientes a la partición P_0 y variando los puntos t_i elegidos.

El conjunto cláramente es no vacio y además por (5) sabemos que $-\varepsilon + I$ es conta inferior del conunto, por lo tanto su ínfimo existe y se cumple que

$$-\varepsilon + I \leqslant \text{ínf}\left\{\sum_{i=1}^{n} f(t_i)\,(x_i - x_{i-1})\ : t_i \in [x_{i-1}, x_i]\right\}$$

Ahora, basta con fijarnos que si definimos

$$A_i = \{f\,(t_i)\ : t_i \in [x_{i-1}, x_i]\},$$

entonces

$$\left\{\sum_{i=1}^{n} f(t_i)\,(x_i - x_{i-1})\ : t_i \in [x_{i-1}, x_i]\right\} = (x_1 - x_0)A_1 + (x_2 - x_1)A_2 + \ldots + (x_n - x_{n-1})A_n.$$

Con el lado derecho de la expresión interpretado como se explica en la definición 3 del capítulo 2. Con esta nueva manera de representar el conjunto de las sumas de Riemann, tenemos que

$$-\varepsilon + I \leqslant \text{ínf}\left((x_1 - x_0)A_1 + (x_2 - x_1)A_2 + \ldots + (x_n - x_{n-1})A_n\right)$$

y por el corolario 1 del capítulo 2, podemos concluir que

$$-\varepsilon + I \leqslant m_1\left(x_1 - x_0\right) + \ldots + m_n\left(x_n - x_{n-1}\right) = L\left(f, P_0\right).$$

Ya que la integral inferior es el supremo de todas las sumas inferiores, se sigue que

$$-\varepsilon + I \leqslant \underline{\int_a^b} f$$

y ya que ε fue un número positivo arbitrario, podemos decir que

$$I \leqslant \underline{\int_a^b} f.$$

De manera similar, de (5) se sigue que el conjunto de sumas de Riemann de f con respecto de P_0 y variando los puntos t_i elegidos, está acotado superiormente por $\varepsilon + I$, por lo que, siguiendo los pasos análogos a lo recientemente demostrado, llegaríamos a que

$$\overline{\int_a^b} f \leqslant I.$$

Pero recordemos que siempre se cumple que $\underline{\int_a^b} f \leqslant \overline{\int_a^b} f$ y por lo tanto

$$I \leqslant \underline{\int_a^b} f \leqslant \overline{\int_a^b} f \leqslant I$$

que es lo que se quería demostrar. ■

Con esto se conluye la sección y a partir de aquí, se usará el término integral de Riemann para referirnos también a la integral de Darboux, pues como se acaba de ver, son conceptos equivalentes.

3.4. Teoremas de intercambio entre límite e integral de Riemann (TCM y TCA)

En caso de aún no haber tomado un curso de teoría de la medida, es bueno saber que estos resultados surgen al estudiar un concepto de integral diferente a la integral de Riemann, llamada integral de Lebesgue, y lo que se presenta en esta sección, es un intento de formular dichos resultados fundamentales de ese tipo de integrales, pero adecuados para la integral de Riemann.

Estos dos teoremas permiten saber bajo qué condiciones es válido hacer el intercambio entre límite e integral. Al primer teorema que se va a presentar, se le conoce como teorema de la convergencia monótona, comúnmente abreviado como TCM . La idea es bastante simple, aunque requiere el mínimo conocimiento de la teoría de sucesiones de funciones. Dicho contenido puede ser encontrado en el capítulo 2 de conceptos preliminares.

Supongamos que se tiene una sucesión creciente de funciones (f_n) definidas en un intervalo $[a, b]$, donde cada f_n es Riemann integrable. Además, supongamos que

$$\lim_{n\to\infty} f_n = f$$

donde una función Riemann integrable.

Ya que $f_n \leqslant f_{n+1}$, por el inciso II) del teorema 6 de propiedades de la integral, se sabe que $\int_a^b f_n \leqslant \int_a^b f_{n+1}$. Por lo tanto, la sucesión de números $\left(\int_a^b f_n\right)$ es creciente. Además, es acotada, pues

$$\int_a^b f_n \leqslant \int_a^b f.$$

Por consecuencia

$$\lim_{n\to\infty} \int_a^b f_n$$

existe. La pregunta obligada sería

$$\text{¿Se cumple que } \lim_{n\to\infty} \int_a^b f_n = \int_a^b \lim_{n\to\infty} f_n = \int_a^b f \text{ ?} \tag{6}$$

Aunque para exponer la idea se asumió un comportamiento creciente de la sucesión de funciones, la pregunta puede ser hecha también para sucesiones decrecientes de funciones. Por este motivo, a este resultado se le conoce como teorema de la convergencia monótona.

Antes de continuar con los teoremas imporantes de esta sección, resulta necesario incluir en nuestra notación, otra manera de representar a las particiones. Como ya se definió, una partición de un intervalo $[a, b]$ es un subconjunto finito de puntos del intervalo, ordenados de menor a mayor, donde el primer punto es a y el último es b. Sin embargo, en el caso de las sumas de Riemann, además de hablar de una partición, también debemos tener en cuenta los puntos intermedios elegidos t_i para calcular la suma de Riemann, es por ello que se insitió en usar la notación $S(f, P, \{t_i\})$ para referirnos a ellas.

Definición 16. Sea $\{x_0, x_1, \ldots, x_k\}$ una partición del intervalo $[a, b]$ y t_i un punto en el intervalo $[x_{i-1}, x_i]$ para cada $i = 1, 2, \ldots, k$. Vamos a llamar partición etiquetada de $[a, b]$ al conjunto de parejas

$$P = \{([x_0, x_1], t_1), ([x_1, x_2], t_2), \ldots, ([x_{k-1}, x_k], t_k)\}$$

Como se observa, la partición etiquetada es otra manera de partir el intervalo, ya no como un subconjunto de puntos, si no como un conjunto de pares conformados por los sub-intervalos que son inducidos por una partición convencional y sus correspondientes puntos intermedios. Cuando se haga referencia a la norma de una partición etiquetada, nos referiremos al valor más grande de los $x_i - x_{i-1}$.

Esta idea de las particiones etiquetadas es especialmente útil para trabajar con sumas de Riemann y de hecho, esa es la razón para incluirlas en este trabajo.

El siguiente corolario es una consecuencia directa del lema 1, presentado en la sección 3.3 de equivalencia entre la integral de Riemann y la integral de Darboux y a pesar de tener una apariencia poco elegante, la idea que transmite es bastante simple. Básicamente, este resultado nos permite asegurar que, a partir de cierta norma en las particiones, cualquier suma de Riemann al sumarle una cantidad positiva ε, esta quedará por encima de la integral inferior. Todo esto sin suponer nada a cerca de la integrabilidad de la función f, únicamente se requiere que sea acotada para que la suma inferior esté bien definida.

Corolario 2. *Sea $f : [a,b] \longrightarrow \mathbb{R}$ una función acotada. Entonces para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que para toda $P = \{([x_0,x_1],t_1),([x_1,x_2],t_2),\ldots,([x_{k-1},x_k],t_k)\}$ partición etiquetada de $[a,b]$ con $|P| < \delta$, se cumple que*

$$\sum_{([x_{i-1},x_i],t_i)\in B} \underline{\int_{x_{i-1}}^{x_i}} f < \varepsilon + \sum_{([x_{i-1},x_i],t_i)\in B} f(t_i)(x_i - x_{i-1}). \tag{7}$$

para cualquier B subconjunto de P.

Demostración. Sea $\varepsilon > 0$ dado. Por el lema 1, sabemos que existe un $\delta > 0$ de modo que

$$\underline{\int_a^b} f - L(f,P) < \varepsilon$$

para cualquier partición con $|P| < \delta$. Lo que en términos de m_i, sería

$$\underline{\int_a^b} f - \sum_{([x_{i-1},x_i],t_i)\in P} m_i(x_i - x_{i-1}) < \varepsilon.$$

Gracias a la propiedad III) del teorema 6, podemos partir la integral inferior de la siguiente manera

$$\begin{aligned}\underline{\int_a^b} f - \sum_{([x_{i-1},x_i],t_i)\in P} m_i(x_i - x_{i-1}) &= \sum_{([x_{i-1},x_i],t_i)\in P} \underline{\int_{x_{i-1}}^{x_i}} f - \sum_{([x_{i-1},x_i],t_i)\in P} m_i(x_i - x_{i-1}) \\ &= \sum_{([x_{i-1},x_i],t_i)\in P} \left(\underline{\int_{x_{i-1}}^{x_i}} f - m_i(x_i - x_{i-1}) \right).\end{aligned}$$

Pero cada término de esa suma es mayor o igual a cero, pues la integral inferior es mayor o igual a la sumas inferiores. Por lo que

$$\sum_{([x_{i-1},x_i],t_i)\in B} \left(\underline{\int_{x_{i-1}}^{x_i}} f - m_i(x_i - x_{i-1}) \right) \leqslant \sum_{([x_{i-1},x_i],t_i)\in P} \left(\underline{\int_{x_{i-1}}^{x_i}} f - m_i(x_i - x_{i-1}) \right) < \varepsilon.$$

Para cualquier B subconjunto de P. De donde

$$\sum_{([x_{i-1},x_i],t_i)\in B} \underline{\int_{x_{i-1}}^{x_i}} f < \varepsilon + \sum_{([x_{i-1},x_i],t_i)\in B} m_i(x_i - x_{i-1}).$$

Y ya que $m_i \leqslant f(t_i)$ para todo $t_i \in [x_{i-1}, x_i]$, tenemos lo que queríamos demostrar.

■

***Observación* 3.** *Sea $\varepsilon > 0$ arbitrario y $\delta > 0$ como en el corolario 2. Entonces, basta con que $x_i - x_{i-1} < \delta$ para todos los elementos de B para que se cumpla la desigualdad (7).*

Lo que sigue es mostrar un resultado característico de las particiones etiqueadas. El lema de Cousin fue demostrado por primera vez por el matemático belga Cousin en el siglo XIX y juega un papel importantísimo en otro tipo de integral llamada la integral de Henstock. Para comprender lo que dice este lema, es necesario definir primero lo que significa que una particón etiquetada sea $\delta - fina$.

Definición 17. Sea $\delta : [a, b] \longrightarrow \mathbb{R}^+$ una función positiva. Decimos que una partición etiquetada $P = \{([x_0, x_1], t_1), ([x_1, x_2], t_2\}, \ldots, ([x_{k-1}, x_k], t_k)\}$ es $\delta - fina$ si

$$x_i - x_{i-1} < \delta(t_i)$$

para $i = 1, 2, \ldots, k$

Es decir, P es $\delta - fina$ si la longitud de cada intervalo es menor a δ evaluada en el punto intermedio.

El lema de Cousin permite garantizar que, sin importar cuál sea nuestra función δ (mientras esta sea estríctamente positiva), siempre va a existir al menos una partición del intervalo que sea $\delta - fina$. La demostración del lema de Cousin que se presenta a continuación, está basada entéramente en las ideas presentadas en **[6]** y los argumentos usados en ella son bastante parecidos a los que se usan en pruebas convencionales de resultados de cálculo como el teorema del valor intermedio entre otros.

Las personas que tengan noción de lo que son los conjuntos cerrados, pueden revisar el capítulo 5, donde el autor de este trabajo propone una demostración original de este hecho.

Lema 2. *(De Cousin)* *Sea $\delta : [a, b] \longrightarrow \mathbb{R}$ una función positiva, entonces existe una partición etiquetada del intervalo $[a, b]$ que es $\delta - fina$.*

Demostración. Consideremos el conjunto

$$S = \{x \in [a, b] : \text{existe una partición } \delta - fina \text{ de [a,x]}\}.$$

Claramente S es no vacío, pues $x \in S$. Además, S está acotado superiormente por b. Gracias a la propiedad de la mínima cota superior, sabemos que S tiene supremo al cual denotaremos por α.

A continuación, nos disponemos a demostrar que $\alpha \in S$, es decir, que el intervalo $[a, \alpha]$ tiene una partición etiquetada $\delta - fina$. Por el teorema 2, sabemos que exise un $x_0 \in S$ tal que

$$\alpha - \frac{\delta(\alpha)}{2} < x_0.$$

Ya que $x_0 \in S$, eso significa que existe una partición etiquetada P de $[a, x_0]$ que es $\delta - fina$. Lo que quiere decri que

$$P \cup \{[x_0, \alpha], \alpha\}$$

es una partición etiquetada $\delta - fina$ de $[a, \alpha]$. Con lo cual concluimos que $\alpha \in S$.

Por último, demostraremos que $\alpha = b$. Supongamos que no es así y que $\alpha < b$. Tomemos un número $c \in [a, b]$ tal que

$$\alpha < c < \text{mín}\,(\alpha + \delta(\alpha), b)\,.$$

Entonces

$$P \cup \{[\alpha, c], \alpha\}$$

resulta ser una partición etiquetada de $[a, c]$ que es $\delta - fina$, lo cual supone una contradicción con el hecho de que α es el supremo de S. Por lo tanto $\alpha = b$. ■

El siguiente lema está completamente inspirado en el trabajo de Brian S. Thomson presentado en **[17]** para la revista de divulgación matemática The American Mathematical Monthly. Sin embargo, existe una diferencia crucial entre la prueba presentada por Thomson y la presentada en este trabajo. Dicha diferencia permite extender el alcance de este lema, para que brinde no solo una demostración prácticamente inmediata del TCM, si no que también nos ayudará en la demostración para el otro resultado principal de esta sección.

Al tratarse de una demostración en términos elementales, sin mucha teoría necesaria para ser entendida, puede volverse un poco técnica y tal vez resulte necesario leerla un par de veces para que sea completamente clara.

Lema 3. *Sea (f_n) una sucesión de funciones acotadas definidas en un intervalo $[a, b]$ que decrecen puntualmente a $f = 0$. Entonces, se cumple que*

$$\lim_{n\to\infty} \underline{\int_a^b f_n} = 0$$

Demostración. Empecemos tomando un $\varepsilon > 0$ arbitrario y definamos la constante $\theta = \frac{\varepsilon}{b-a+1}$. A continuación, gracias al corolario 2, para cada $n \in N$ podemos hallar un número real $\delta_n > 0$ de tal modo que

$$\sum_B \underline{\int_{x_{i-1}}^{x_i} f_n} < \frac{\theta}{2^n} + \sum_B f_n\,(t_i)\,(x_i - x_{i-1}) \tag{8}$$

para cualquier subconjuno B de una partición etiquetada P con $|P| < \delta_n$.

Ahora, para cada x en $[a, b]$ definimos al número natural $N\,(x)$ como el primer número natural para el cual la sucesión $f_n(x) < \theta$.

En la Figura 7 se muestran los primeros términos de una sucesión a modo de ejemplo. En ese caso ilustrativo $N\,(x) = 3$, pues $f_1\,(x)\,, f_2\,(x) > \theta$, pero $f_3\,(x) < \theta$.

De esta manera, agruparemos a los x de $[a, b]$ que tengan el mismo valor de $N\,(x)$ mediante los siguientes conjuntos

$$E_j = \{x \in [a, b] \mid N(x) = j\}$$

para $j = 1, 2, \ldots$.

Volviendo al ejemplo de la Figura 7, los conjuntos E_j quedarían como en la Figura 8

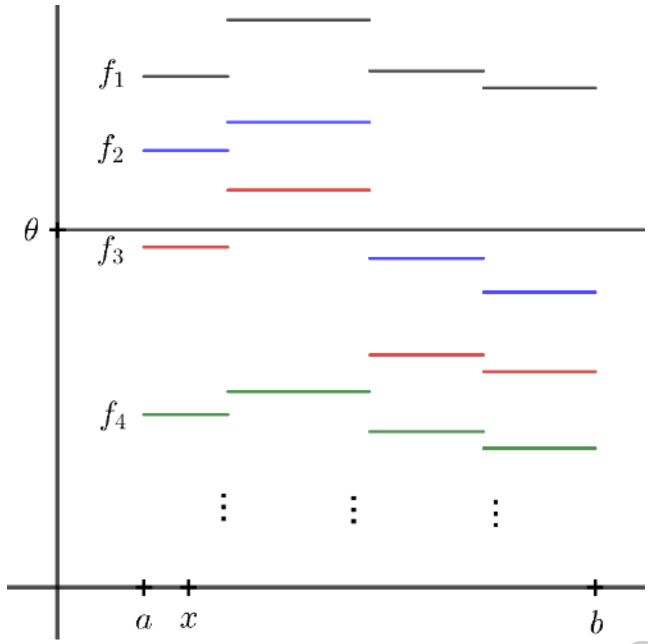

Figura 7: Ejemplo ilustrativo.

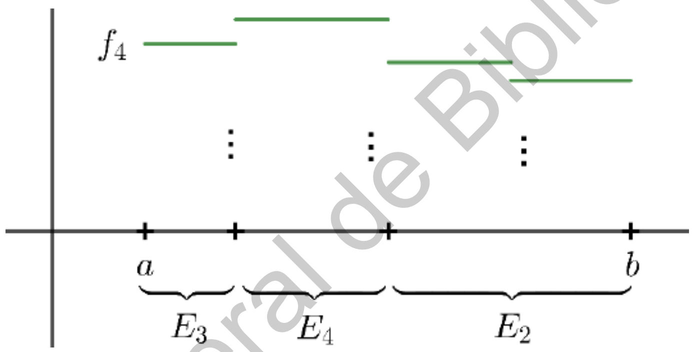

Figura 8: Imagen recortada del ejemplo ilustrativo

Así, cada elemento del intervalo $[a,b]$ pertenece a algún E_j.

Lo que sigue es usar estos conjuntos para definir la función $\delta : [a,b] \longrightarrow \mathbb{R}^+$ de la siguiente manera

$$\delta(x) = \delta_j \text{ si } x \in E_j.$$

Continuando con el ejemplo, la Figura 9 muestra una posible gráfica de la funcion δ.

Ya que δ es una función positiva, podemos aplicar el lema de Cousin para obtener una partición etiquetada $P = \{([x_0, x_1], t_1), ([x_1, x_2], t_2\}, \ldots, ([x_{k-1}, x_k], t_k)\}$ de tal modo que $x_i - x_{i-1} < \delta(t_i)$ para cada $i = 1, 2, \ldots, k$.

Lo siguiente será agrupar los pares de la partición etiquetada que compartan el mismo $N(t_i)$. Sea N el número más grande de los $N(t_i)$ con $i = 1, 2, \ldots, k$. Definimos a P_j como sigue

$$P_j = \{([x_{i-1}, x_i], t_i) \in P : N(t_i) = j\}$$

para $j = 1, 2, \ldots, N$. De este modo, los P_j son conjuntos disjuntos y además $P = P_1 \cup P_2 \cup \ldots \cup P_N$.

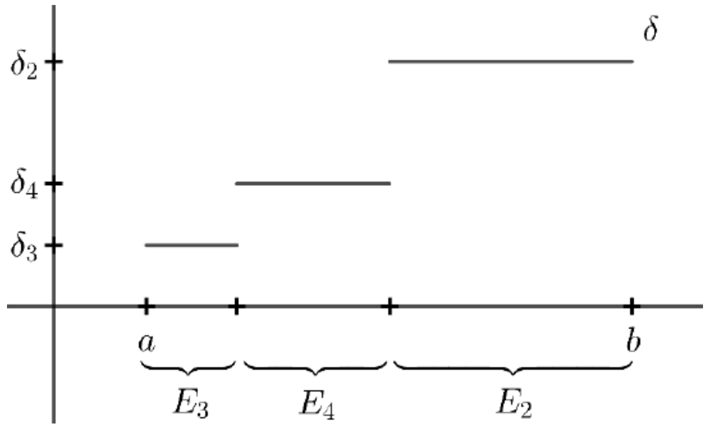

Figura 9: Función δ.

Veamos lo que pasa si tomamos un número natural m mayor o igual a N y calculamos la integral inferior de f_m. Pues gracias a la propiedad III demostrada en el teorema 6, podemos descomponer la integral inferior en suma de integrales inferiores por suintervalos de la siguiente manera

$$\begin{aligned} 0 \leqslant \underline{\int_a^b f_m} &= \sum_{i=1}^{k} \left(\underline{\int_{x_{i-1}}^{x_i} f_m} \right) \\ &= \sum_{j=1}^{N} \left(\sum_{([x_{i-1},x_i],t_i)\in P_j} \underline{\int_{x_{i-1}}^{x_i} f_m} \right) \end{aligned}$$

Debido a que $m \geqslant N$, tenemos que $f_m \leqslant f_j$ para cada $j = 1, 2, \ldots, N$ y por lo tanto $\underline{\int_{x_{i-1}}^{x_i} f_m} \leqslant \underline{\int_{x_{i-1}}^{x_i} f_j}$ para cada j, dejándonos con

$$\underline{\int_a^b f_m} \leqslant \sum_{j=1}^{N} \left(\sum_{([x_{i-1},x_i],t_i)\in P_j} \underline{\int_{x_{i-1}}^{x_i} f_j} \right)$$

Ya que $N(t_i) = j$ para todos los $t_i \in P_j$, esto implica dos cosas; la primera es que, por el lema de Cousin, $x_i - x_{i-1} < \delta(t_i) = \delta_j$. Lo que quiere decir que P_j es un subconjunto de una partición con norma menor a δ_j, por lo cual (8) se cumple para cada $j = 1, 2, \ldots, N$ y por lo tanto, podemos concluir que

$$\sum_{j=1}^{N} \left(\sum_{([x_{i-1},x_i],t_i)\in P_j} \underline{\int_{x_{i-1}}^{x_i} f_j} \right) \leqslant \sum_{j=1}^{N} \left(\frac{\theta}{2^j} + \sum_{([x_{i-1},x_i],t_i)\in P_j} f_j(t_i)(x_i - x_{i-1}) \right)$$

La segunda cosa que podemos concluir de que $N(t_i) = j$, es que $f_j(t_i) < \theta$, pues así se definió $N(t_i)$, como el primer entero para el cual el valor de la sucesión (f_n) es menor a θ en el punto t_i. Por lo tanto, se sigue que

$$
\begin{aligned}
\sum_{j=1}^{N}\left(\sum_{([x_{i-1},x_i],t_i)\in P_j} \underline{\int_{x_{i-1}}^{x_i}} f_j\right) &< \sum_{j=1}^{N}\left(\frac{\theta}{2^j}+\sum_{([x_{i-1},x_i],t_i)\in P_j}\theta\left(x_i-x_{i-1}\right)\right)\\
&=\theta\sum_{j=1}^{N}\left(\frac{1}{2^j}+\sum_{([x_{i-1},x_i],t_i)\in P_j}\left(x_i-x_{i-1}\right)\right)\\
&<\theta\left(\sum_{j=1}^{\infty}\frac{1}{2^j}+\sum_{j=1}^{N}\sum_{([x_{i-1},x_i],t_i)\in P_j}\left(x_i-x_{i-1}\right)\right)\\
&=\theta\left(1+(b-a)\right)\\
&=\varepsilon
\end{aligned}
$$

de manera que si $m \geqslant N$, entonces

$$0 \leqslant \underline{\int_a^b} f_m < \varepsilon.$$

$\blacksquare$

Como se había comentado antes, esta demostración está completamente basada en la dada por Thomson en **[17]**, con la diferencia de que Thomson asume la integrabilidad de cada función f_n de la sucesión y demuestra que el límite de las integrales converge a cero. Lo cual, como se verá en el siguiente teorema, ya representa una demostración para el TCM. Sin embargo, al introducir el corolario 2, se pudo mejorar un poco el resultado removiendo la restricción de que cada f_n sea integrable, pero quedandose en su lugar con la integral inferior. Debería ser obvio que el lema 3 es un resultado más general que el presentado por Thomson. Pues en caso de ser integrables las f_n, entonces se cumple que

$$\lim_{n\to\infty}\int_a^b f = \lim_{n\to\infty}\underline{\int_a^b} f_n = 0.$$

En **[9]**, se puede encontrar otra demostración del lema 3 proporcionada por Wilhelmus Luxemburg. El único inconveniente con esa prueba (inconveniente para este trabajo) es que hace uso del teorema de Dini para decir que una sucesión de funciones continuas definidas en un intervalo $[a,b]$ que convergen a una función continua, lo hacen de manera uniforme. Dicho teorema usa teoría de conjuntos compactos y la noción de otro tipo de convergencia, los cuales, son conceptos que no seán estudiades en este trabajo, pues se aleja del objetivo del mismo que es el de presentar demostraciones lo más elementales posibles.

Ahora la demostración del TCM es prácticamente inmediata del lema 3.

Teorema 9. ***(TCM)*** *Sea (f_n) una sucesión monótona de funciones definidas en $[a,b]$ que son Riemann integrables y que convergen puntualmente a una función f Riemann integrable. Entonces,*

$$\lim_{n\to\infty}\int_a^b f_n = \int_a^b \lim_{n\to\infty} f_n = \int_a^b f$$

Demostración. Comencemos demostrando el caso cuando la sucesión (f_n) es creciente. Definamos la sucesión (g_n) dada por $g_n = f - f_n$. No es difícil ver que la sucesión (g_n) está acotada y además decrece a cero. Por el lema 3, tenemos que

$$0 = \lim_{n\to\infty} \int_a^b g_n = \lim_{n\to\infty} \int_a^b f_n - \int_a^b f.$$

Para el caso cuando (f_n) es decreciente, simplemente notemos que $g_n = f - f_n$ define una sucesión creciente de funciones Riemann intergables que convergen puntualmente a cero. Por la primera parte de esta demostración, podemos concluir que

$$\lim_{n\to\infty} \int_a^b g_n = \int_a^b \lim_{n\to\infty} g_n = \int_a^b 0 = 0.$$

■

Por supuesto, si el alcance del lema 3 se quedara aquí, en realidad no representaría una mejoría a lo ya presentado por Thomson en **[17]**. Es en el segundo resultado principal de esta sección, el teorema de la convergencia acotada (TCA), donde se alcanza todo el potencial del lema 3.

La idea de usar el lema 3 para demostrar el TCA se debe a Wilhelmus Luxemburg en **[9]**, por lo que la siguiente demostración está inspirada en su trabajo.

Teorema 10. ***(TCA)*** *Sea (f_n) una sucesión de funciones Riemann integrables definidas en $[a, b]$ que convergen puntualmente a una función f que también es Riemann integrable. Supongamos que existe una constante $M > 0$ de modo que $|f_n| \leqslant M$ para todo $n \in N$, entonces*

$$\lim_{n\to\infty} \int_a^b |f - f_n| = 0.$$

En particular

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b f.$$

Demostración. Comencemos definiendo la sucesión de funciones (g_n) dada por $g_n = |f - f_n|$. Por la propiedad 4 de la integral de Riemann, demostrada en el teorema 6, podemos decir que cada g_n es integrable. Y ya que $|f_n| < M$ para toda n, esto implica que $|f| \leqslant M$ y por lo tanto $g_n \leqslant |f_n| + |f| < 2M$. Además, por como está definida, la sucesión (g_n) converge puntalmente a cero (no necesariamente de manera decreciente).

Ahora, para cada $x \in [a, b]$ y $n \in \mathbb{N}$, definimos a

$$h_n(x) := \sup \{g_n(x), g_{n+1}(x), g_{n+2}(x), \ldots\}.$$

Entonces, la sucesión (h_n) está acotada superiormente por $2M$ y además $0 \leqslant g_n \leqslant h_n$ para toda $n \in \mathbb{N}$.

Por el teorema 3 del capítulo 2, sabemos que $h_{n+1}(x) \leqslant h_n(x)$ para todo $x \in [a, b]$ y para todo $n \in \mathbb{N}$, por lo que es una sucesión decreciente. Además, por el mismo teorema, sabemos que para cada $x \in [a, b]$

$$\lim_{n\to\infty} h_n(x) = \lim_{n\to\infty} g_n(x) = 0$$

Con lo cual, podemos aplicar el lema 3 a la sucesión (h_n) y decir que

$$\lim_{n\to\infty} \underline{\int_a^b} h_n = 0$$

Pero $0 \leqslant \int_a^b g_n = \underline{\int_a^b} g_n \leqslant \underline{\int_a^b} h_n$. Terminando así, la demostración. ■

Gracias a estos dos teoremas, el TCM y el TCA, ahora se tienen las condiciones bajo las cuales es válido hacer el intecambio entre límite e integral de Riemann. Son notables las diferencias que existen entre ambos resultados. En el TCA, logramos quitar la restricción de que la sucesión debe de ser monótona, pero agregando el supuesto de que tiene que existir una constante acotando la sucesión. Esta nueva condición de la cota, en relidad no representa una gran restricción si lo comparamos con lo supuesto en el TCM. Si (f_n) es una sucesión de funciones Riemann integrables definidas en $[a, b]$ que crecen a una función f Riemann integrable, entonces $f_1 \leqslant f_2 \leqslant ... \leqslant f$ y la sucesión estaría acotada inferiormente por $M_1 = \inf\{f_1(x) \mid x \in [a, b]\}$ y superiormente por $M_2 = \sup\{f(x) \mid x \in [a, b]\}$, tomando M igual al máximo entre $|M_1|$ y $|M_2|$, tendríamos una cota para toda la sucesión. Esto quiere decir que, en realidad, el TCA implica el TCM. La razón para incluir ambos teoremas, es debido a que en este trabajo se busca dar un panorama, lo más amplio posible, de los resultados más importantes de la integral de Lebesgue, para aquellas personas que aún no hayan tomado un curso de teoría de la medida.

En caso de ya estar familiarizado con la integral de Lebesgue, es interesante notar el papel que juega el lema 3 en la demostración del TCA para la integral de Riemann. Es el mismo que el que desempeña el TCM para la integral de Lebesgue, en la demostración del teorema de la convergencia dominada.

Aún queda algo más que comentar a cerca de estos dos teoremas. Se debe discutir la importancia que tiene la hipótesis de que el límite f de la sucesión de funciones debe ser Riemann integrable. Para las personas cuyo acercamiento a estos resultados se haya dado por primera vez en este documento puede parecer una condición obviamente necesaria, pues cómo se puede hablar de intercambiar límite con integral si para empezar, no se sabe si el límite es una función integrable. Sin embargo, no se debe de olvidar que estos resultados provienen de una idea de integral más general y por lo tano, puede haber hipótesis que ya no sean necesarias de asumir. Debido a que hablar de esto sería algo de interes mayormente para las personas que conozcan los resultados originales de la integral de Lebesgue, este contenido será incluido en la sección 4.6. Aún así, no será necesario saber nada de la integral de Lebesgue o teoría de la medida para entender esa pequeña primera parte de la sección.

Por último, otros trabajos que también abordan este tema del intercambio entre límite e integral de Riemann son los siguientes. En **[3]** y **[5]** Cunningham y Gordon dan una demostración de que el TCD es equivalente a un lema conocido como lema de Arzela, luego bastaría con dar una demostración del lema de Arzela. La demostración de este lema la hicieron por primera vez Arzela en 1885 y de manera independiente Osgood en 1887 **[12]**. Estas pruebas se consideran bastante técnicas y pesadas. Afortunadamente, se han dado pruebas más elementales de este lema (ver **[8]**). Finalmente, en **[7]** Kastelman introduce una leve noción de medida para conjuntos abiertos y así logra dar una demostración del TCD un poco técnica, pero accesible si se tiene noción de los conjuntos abiertos y sus propiedades.

3.5. Teoremas de convergencia para la integral impropia de Riemann

En esta sección se explorará la posibilidad de extender los teoremas de convergencia monótona y acotada para cuando el intervalo de definición de la sucesión de funciones no es un intervalo acotado. Esto, mediante el uso de la integral impropia de Riemann.

La integral impropia de Riemann es un intento por extender la integral de Riemann a situaciones donde bien, el intervalo de definición de la función no esté acotado, o bien, la función en cuestión, no esté acotada en su intervalo de definición.

Aquí únicamente trabajaremos con la primera de estas situacones, la otra puede abordarse de manera similar.

Comencemos definiendo lo que significa la integral impropia de Riemann de una función.

Definición 18. Sea $f : [a, \infty) \longrightarrow \mathbb{R}$ una funcion para la cual $\int_a^m f$ existe para todo $m \in [a, \infty)$. Si se cumple que

$$\lim_{m\to\infty} \int_a^m f$$

existe, es decir, si existe un númer real I tal que para todo $\varepsilon > 0$ existe un $M > a$ tal que

$$\left| \int_a^m f - I \right| < \varepsilon.$$

para todo $m \geqslant M$, entonces se dice que la integral impropia de f existe y a dicho límite I se le conoce como integral impropia de f y se denota por

$$\int_a^\infty f.$$

Como ee aprecia, la intergal impropia es una manera muy natural de extender la integral de Riemann para funciones cuyo intervalo de definición no esté acotado, simplemente viéndola como el límite de integrales definidas cuando estas tienden a infinito.

Resulta ser que, la integral impropia de Riemann cumple las mismas propiedades básicas que la integral de Riemann para intervalos acotados, esto gracias a las propiedades tanto del límite como de la misma integral de Riemann.

Teorema 11. *Sean $f, g : [a, \infty) \longrightarrow \mathbb{R}$ dos funciones cuya integral impropia de Riemann existe y k un número real. Entonces,*

I) *La integral impropia de $f + kg$ existe y se cumple que*

$$\int_a^\infty f + kg = \int_a^\infty f + k \int_a^\infty g.$$

II) *Si $f \leqslant g$, entonces*

$$\int_a^\infty f \leqslant \int_a^\infty g.$$

III) *Si* $c \in [a, \infty)$, *entonces*
$$\int_a^c f + \int_c^\infty f = \int_a^\infty f.$$

Comparado con el teorema 6, solo faltaría la propiedad **??**). En general, no es cierto que si la inegral impropia de Riemann de f existe, entonces la integral impropia de $|f|$ existe.

Con estas propiedades se puede hacer la siguiente observación que surge de la definición de intergal impropia.

***Observación* 4.** *Si la integral impropia de* $f : [a, \infty) \longrightarrow \mathbb{R}$ *existe, entonces para todo* $\varepsilon > 0$ *existe un* $M > a$ *tal que*
$$\left| \int_M^\infty f \right| = \left| \int_a^M f - \int_a^\infty f \right| = \left| \int_a^M f - I \right| < \varepsilon.$$

A continuación, se dará el único resultado extra que se necesita para la demostración del TCA y el TCM.

Teorema 12. *Sean* $f, g : [a, \infty) \longrightarrow \mathbb{R}$ *dos funciones no negativas tales que* $f \geqslant g$ *y*
$$\int_a^m f \ , \ \int_a^m g$$
existen para todo $m \in [a, \infty)$*. Entonces, si la integral impropia de* g *existe, la integral impropia de* f *también existe.*

Demostración. Definamos a
$$F(x) = \int_a^x f$$
y a
$$G(x) = \int_a^x g.$$

Es fácil ver que tanto F como G son funciones crecientes y por lo tanto el límite cuando x tiende a infinito de $F(x)$ existirá si y solo si F está acotada. Pero, por la propiedadII del teorema 11, sabemos que
$$\int_a^\infty g$$
es una cota para F. Por lo tanto
$$\lim_{x \to \infty} F(x) = \lim_{x \to \infty} \int_a^x f = \int_a^\infty f$$
existe. ■

Ahora, ya se puede ver la demostración del TCA para la integral impropia de Riemann, la cual, como veremos a continuación, se basa fuertemente en el TCA para intervalos acotados.

Teorema 13. ***(TCA para la integral impropia de Riemann)*** *Sea (f_n) una sucesión de funciones definidas en $[a,\infty)$ que convergen puntualmente a una función f. Supongamos que las integrales impropias de Riemann de f y de cada f_n existen. Además, supongamos que existe una función $g:[a,\infty)\longrightarrow\mathbb{R}$, cuya integral impropia de Riemann existe, tal que $|f_n|\leqslant g$ para todo n. Entonces,*

$$\lim_{n\to\infty}\int_a^\infty |f-f_n| = 0.$$

En particular

$$\lim_{n\to\infty}\int_a^\infty f_n = \int_a^\infty f.$$

Demostración. Antes que nada, deberíamos de garantizar la existencia de la integral impropia de $|f-f_n|$ para cada n. Recordemos que, en general, la integral impropia del valor absoluto de una función no siempre existe aún cuando la integral impropia de la función exista. Pero esto es fácil de probar, pues para todo $n\in\mathbb{N}$

$$|f-f_n|\leqslant|f|+|f_n|\leqslant 2g. \tag{9}$$

Aplicando el teorema 12 y la propiedad IV del teorema 6, concluimos que $|f-f_n|$ tiene integral impropia.

Ahora, sea $\varepsilon>0$ arbitrario. Por la observación 4, sabemos que existe un $M>a$ tal que

$$\int_M^\infty g = \left|\int_M^\infty g\right| < \frac{\varepsilon}{4}.$$

Gracias a (9), concluimos que

$$\int_M^\infty |f-f_n| \leqslant 2\int_M^\infty g < \frac{2\varepsilon}{4} = \frac{\varepsilon}{2}.$$

para todo $n\in N$. A continuación, notemos que en el intervalo $[a,M]$ la sucesión (f_n) está acotada por K, donde K es cualquier cota superior de g en $[a,M]$. Por lo tanto, podemos aplicar el TCA para la sucesión (f_n) restringida únicamente al intervalo $[a,M]$ y decir que exise un $n\in\mathbb{N}$ tal que

$$\int_a^M |f-f_n| < \frac{\varepsilon}{2}.$$

para todo $n\geqslant N$. De esta manera

$$\int_a^\infty |f-f_n| = \int_a^M |f-f_n| + \int_M^\infty |f-f_n| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon.$$

■

Al comparar este TCA con el TCA para intervalos acotados la diferencia es clara. Aquí ya no basta que la sucesión se encuentre acotada por una constante, si no que debe de estar acotada por una función integrable. En realidad, se puede cambiar la constante por una función integrable en el TCA para intervalos acotados, pues en ese caso, ambas condiciones son equivalentes. Sin embargo, para intervalos no acotados, las funciones constantes ya no son funciones integrables.

Para terminar con esta sección, el TCM para la integral impropia de Riemann.

Teorema 14. ***(TCM para la integral impropia de Riemann)*** *Sea (f_n) una sucesión monótona de funciones definidas en $[a,\infty)$ que convergen puntualemnte a una función f. Supongamos que la integral impropia de f y de cada f_n existen. Entonces*

$$\lim_{n\to\infty}\int_a^\infty f_n = \int_a^b f. \tag{10}$$

Demostración. Supongamos que (f_n) es creciente. Por lo tanto $f_1 \leqslant f_2 \leqslant \ldots \leqslant f$ y si definimos $g_n = f - f_n$ vemos que la sucesión (g_n) es una sucesión decreciente de funciones que converge puntualmente a cero cuya integral impropia de Riemann existe para cada $n \in \mathbb{N}$. Por lo tanto, toda la sucesión es positiva y se encuentra acotada por g_1, aplicando el TCA para la integral impropia de Riemann, tenemos que

$$0 = \lim_{n\to\infty}\int_a^\infty g_n = \lim_{n\to\infty}\int_a^\infty f_n - \int_a^\infty f.$$

Si (f_n) es decreciente, entonces $g_n = f_n - f$ define una sucesión decreciente de funciones que converge puntualmente a cero y concluimos lo mismo que en el caso anterior.

∎

4. Integral de Lebesgue

Dentro del análisis matemático la integral de Lebesgue nace de una rama llamada teoría de la medida. Esta teoría es fundamental en otras disciplinas de las matemáticas como la probabilidad, estadística y geometría. Entender el objetivo de la teoría de la medida es bastante sencillo. Esta, simplemente busca asignarle a los conjuntos una cierta medida cuya intención es, representar de alguna manera, el tamaño de los conjuntos. Por ejemplo, supongamos que tenemos un saco con 50 canicas idénticas en su interior y estamos buscando la manera de que, a cada subconjunto de canícas del saco original, se le asigne una medida. En ese caso, lo más razonable sería usar como medida la función que cuenta el número de canicas que hay en los subconjuntos.

Más adelante se presentará una definición formal de medida y aunque el concepto de medida es muy general y en principio puede ser aplicada a conjuntos arbitrarios, este trabajo se concentrará en una idea de medida muy específica para ser usada sobre los subconjuntos de los números reales. Desafortunadamente, aplicar la misma idea de medida que vimos en el ejemplo reciente de asignarle como medida a los subconjuntos de los números reales el número de elementos que posean, parece que llevaría la teoría por un camino poco útil e interesante. La medida que se usará en este trabajo para subconjuntos de $\mathbb{R}$ es conocida como medida de Lebesgue y esta puede ser pensada como la extensión de la medida que a los intervalos de la forma $[a, b]$ les asigna como medida su longitud, $b - a$. Sin embargo, dentro de los números reales existen subonjuntos espantosos y la duda sería, si también es posible asignarle una medida de Lebesgue a estos subconjuntos. Antes de dar una respuesta a esta incognita y de definir formalmente lo que se quiere decir por medida, se intentará esbozar que relación existe entre la poca idea que tenemos hasta ahora de medida y la integral.

Supongamos que se tiene una función f definida en el intervalo $[a, b]$ que toma un valor constante $k > 0$. Además, supongamos que existe una función μ que va del conjunto potencia de $[a, b]$ a $\mathbb{R}^+$ de modo que, $\mu(A)$ representa, de alguna manera, el tamaño del conjunto A para cualquier $A \subseteq [a, b]$. Lo que sabemos hasta ahora, gracias a la integral de Riemann, es que la integral de la función f representa el área que se encuentra entre el eje x y la gráfica de f que en este caso corresponde al área de un rectángulo de base $b - a$ y altura k. Por lo tanto, la integral de f debería de ser $k(b - a)$. En otras palabras

$$\int f = k \cdot \mu([a, b]) = k \cdot \mu\left(f^{-1}(\{k\})\right)$$

donde f^{-1} representa la imagen inversa. Con esto, simplemente se está queriendo decir que la integral de f es, la medida de la base del rectángulo, multiplicado por su altura. Claramente tendríamos un problema si se intentara extender esta idea de integral a funciones que tomaran más de un valor. Sin ir más lejos, con este mismo razonamiento y si la función f toma dos valores positivos k_1 y k_2, entonces la integral de f debería de ser

$$k_1 \cdot \mu\left(f^{-1}(\{k_1\})\right) + k_2 \cdot \mu\left(f^{-1}(\{k_2\})\right) \tag{11}$$

Si la gráfica de la función f fuera como en la parte de arriba de la figura 10, parece razonable pensar que la integral de f debería ser $k_1(c - a) + k_2(b - c)$, que

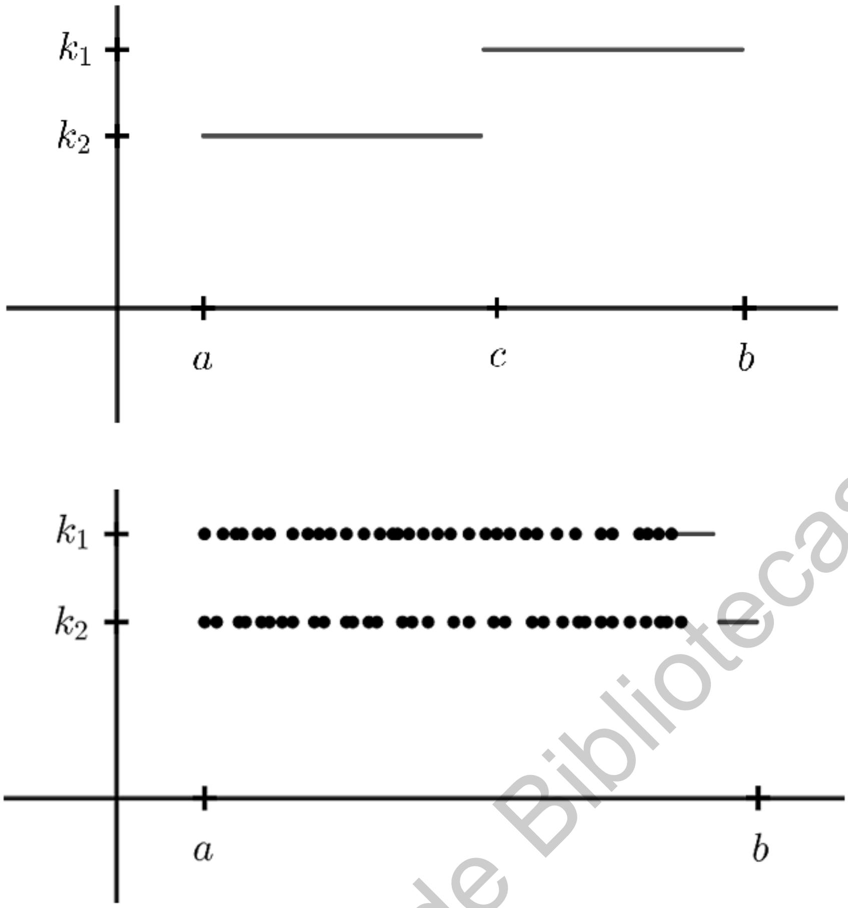

Figura 10: Dos funciones que toman únicamente dos valores k_1 y k_2.

en este caso, coincide con la expresión (11) . Sin embargo, al observar la imagen de abajo de la misma figura, vuelve a presentarse el mismo problema de hace un momento, el de definir la medida para cualquier subconjunto de $[a, b]$, pues se debe tener presente estas ideas se empezaron bajo la suposición de que dicha medida μ existe. La existencia de tal medida es un hecho muy profundo y complejo que llega hasta lo que aceptamos como axiomas en la teoría de conjuntos y por lo tanto, se escapa del objetivo de este trabajo. Se ha probado que no es posible definir una medida en el conjunto potencia que generalice la idea de longitud de un intervalo.

Dicho esto, y ya que no es posible definir una medida como nos gustaría para todos los subconjuntos de $\mathbb{R}$, viene la pregunta ¿para qué subconjuntos de $\mathbb{R}$ es posible definir su medida, extendiendo el concepto de longitud de un intervalo? Para responder esta pregunta, resulta necesario definir primero lo que es una $\sigma - \acute{a}lgebra$ de un conjunto.

Ya que el principal objetivo de este trabajo es hacer una comparasión entre la integral de Riemann y la integral de Lebesgue, no es prioridad desarrollar toda la teoría necesaria de la integral de Lebesgue, por lo que algunos resultados pueden estar sin demostración. Sin embargo, siempre habra una referencia a la cual se pueda acudir para recibir más detalles al respecto.

4.1. Sigma álgebra

Una sigma álgebra es una colección de subconjuntos la cual cumple con las propiedades mínimas para ser candidata a poder ser el espacio donde definir una medida.

Definición 19. Sea X un conjunto. Una $\sigma - álgebra$ de X es una colección $\mathcal{F}$ de subconjuntos de X que cumple con las siguientes propiedades:

i) $X \in \mathcal{F}$

ii) Si $A \in \mathcal{F}$ entonces $A^c \in \mathcal{F}$

iii) Si (A_n) es una sucesión de elementos de $\mathcal{F}$, entonces $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

Al par $(X, \mathcal{F})$ se le llama espacio medible y a los elementos de $\mathcal{F}$ se les llaman conjuntos o elementos medibles.

La anterior es una definición general de $\sigma - álgebra$, pero es de particular en este trabajo, hallar una $\sigma - álgebra$ para $\mathbb{R}$. Sin embargo, no cualquier $\sigma - álgebra$ va a ser útil, pues el conjunto $\{\emptyset, \mathbb{R}\}$ representa una $\sigma - álgebra$ de $\mathbb{R}$, no obstante, esta no sirve para nuestros propósitos. Aún más, el conjunto potencia es en sí una $\sigma - álgebra$, pero como ya se discutió previamente, esta tampoco resulta útil. De hecho, la $\sigma - álgebra$ que va a servir es una llamada $\sigma - álgebra$ de Lebesgue. Lamentablemente, no será posible dar una definición completamente formal de esta $\sigma - álgebra$, pues requiere un extenso desarrollo de la teoría de la medida de Lebesgue, por lo que en este trabajo únicamente se intentará dar una idea de como sería esta $\sigma - álgebra$. Sin embargo, se debe de tener claro que la manera de presentar la $\sigma - álgebra$ de Lebesgue que se está dando en este trabajo no es la usual y tampoco es la correcta, pero consideramos que es la que más se adecua para los porpósitos del estudio. Los interesados en ver una definición formal de la $\sigma - álgebra$ de Lebesgue pueden consultar el capítulo 9 de [**2**].

Antes de intentar "definir" la $\sigma - álgebra$ de Lebesgue, se necestian considerar algunos aspectos de la medida de Lebesgue y a la $\sigma - álgebra$ de Borel. Esta última parte es precisamente por la cual se verá primero el siguiente resultado.

Teorema 15. *Si $(\mathcal{F}_\alpha)_{\alpha \in I}$ es una colección arbitraria de sigmas álgebras de X, entonces*

$$\mathcal{F} = \bigcap_{\alpha \in I} \mathcal{F}_\alpha$$

es también una sigma álgebra.

Demostración. Veamos que $\mathcal{F}$ cumple con las tres propiedades para ser una sigma álgebra.

i) Es inmediato, pues $X \in F_\alpha$ para todo α, de manera que X se encuentra en la intersección de todas ellas.

ii) Sea $A \in \mathcal{F}$. Entonces, $A \in \mathcal{F}_\alpha$ para cada α y por lo tanto, $A^c \in \mathbb{F}_\alpha$ para todo α. De este modo, A^c partenece a la intesección de todas las sigmas álgebras.

iii) Sea (A_n) una sucesión de elementos de $\mathcal{F}$. Esto implica que $A_n \in \mathcal{F}_\alpha$ para todo n y para cada α, por lo que $\cup_{n=1}^{\infty} A_n \in \mathcal{F}_\alpha$ para cada α y así, $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$. ■

Definición 20. La $\sigma - \acute{a}lgebra$ en $\mathbb{R}$ generada por la intersección de todas las $\sigma - \acute{a}lgebras$ que contienen a todos los intervalos de la forma (a, b) se llama $\sigma - \acute{a}lgebra$ de Borel y se denota por $\mathcal{B}(\mathbb{R})$. Si $[a, b]$ es un intervalo fijo, entonces $\mathcal{B}([a, b])$ denota a la $\sigma - \acute{a}lgebra$ que contiene a todos conjuntos medibles de $\mathcal{B}(\mathbb{R})$ que son subconjuntos de $[a, b]$. A los elementos de $\mathcal{B}(\mathbb{R})$ y $\mathcal{B}([a, b])$ se les conocen también como Borelianos y los Borelianos en $[a, b]$, respectivamente.

Claramente existe al menos una $\sigma - \acute{a}lgebra$ que cumple con contener a todos los intervalos de la forma (a, b), el conjunto potencia. Luego, debido al teorema 15, la $\sigma - \acute{a}lgebra$ de Borel está bien definida y representa la $\sigma - \acute{a}lgebra$ más "pequeña" que contiene a todos los intervalos de la forma (a, b).

4.2. Medida de Lebesgue

Definición 21. Sea $(X, \mathcal{F})$ un espacio medible. Una medida en $(X, \mathcal{F})$ es una función $\mu : \mathcal{F} \to \overline{\mathbb{R}}^{+} \bigcup \{0\}$ tal que

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

II) $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$, si (A_n) es una sucesión disjunta de elementos de $\mathcal{F}$

A la terna $(X, \mathcal{F}, \mu)$ se le conoce como espacio de medida.

Nuevamente, esta definición de medida es general. Lo que interesa para este estudio es poder definir una medida en la $\sigma - \acute{a}lgebra$ de Borel de modo que, en el caso de la medida para los intervalos de la forma $[a, b]$, esta coincida con lo que se espera, que sería asociarles el número $b - a$, para posteriormente, ahora sí, definir la $\sigma - \acute{a}lgebra$ de Lebesgue. La medida que permite hacer esto es conocida como la medida de Lebesgue y generalmente se denota por λ. La construcción de esta medida no forma parte del objetivo de este trabajo, por lo que se puede consultar **[2]** para tener más detalles acerca de este tema.

Teniendo una vaga idea de lo que la de medida de Lebesgue representa y junto con la $\sigma - \acute{a}lgebra$ de Borel, estamos listos para "definir" la $\sigma - \acute{a}lgebra$ de Lebesgue.

Definición 22. Definimos a la $\sigma - \acute{a}lgebra$ de Lebesgue, denotada por $\mathcal{L}(\mathbb{R})$, como todos los conjuntos Borel medibles y agregamos los subconjuntos de conjuntos de medida cero. Es decir, E pertenece a $\mathcal{L}(\mathbb{R})$ si y solo si E es un Boreliano o $E \subseteq A$ de modo que $\lambda(A) = 0$. A los elementos de $\mathcal{L}(\mathbb{R})$ se les conoce como elementos o conjuntos Lebesgue medibles.

En lo que resta del documento se estará trabajando únicamente con el espacio de medida $(\mathbb{R}, \mathcal{L}(\mathbb{R}), \lambda)$ conformado por los números reales, la $\sigma - \acute{a}lgebra$ de Lebesgue y la medida de Lebesgue o con el espacio de medida $([a, b], \mathcal{L}([a, b]), \lambda)$ que son, el intervalo cerrado $[a, b]$, los conjuntos Lebesgue medibles en $[a, b]$ y la medida de Lebesgue restringida a los conjuntos Lebesgue medibles en $[a, b]$.

El siguiente resultado básico lo cumple cualquier medida, aunque aquí lo enunciaremos para el caso de la medida de Lebesgue. Una demostración para este teorema puede ser encontrada en el capítulo 3 de **[2]**.

Teorema 16. *Sea* $(\mathbb{R}, \mathcal{L}(\mathbb{R}), \lambda)$ *el espacio de medida conformado por los números reales, la* $\sigma -$ *álgebra de Lebesgue y la medida de Lebesgue. Si* (A_n) *es una sucesión creciente de conjuntos, es decir,* $A_n \subseteq A_{n+1}$*, entonces*

$$\lambda\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \lambda(A_n).$$

A diferencia del teorema anterior, el resultado que viene a continuación es inherente a la integral de Lebesgue y aunque no es difícil de demostrar, sí está íntimamente relacionado con el como se define la medida de Lebesgue, por lo que los interesados pueden encontrar información al respecto en el capítulo 13 de **[2]**. Sin embargo, la idea que intenta expresar es bastante simple; los conjuntos medibles pueden ser aproximados tanto como se quieran por conjuntos cerrados que quedan por "debajo" o por conjuntos abiertos que quedan por "encima".

Teorema 17. *Consideremos el espacio de medida* $(\mathbb{R}, \mathcal{L}(\mathbb{R}), \lambda)$ *el conformado por el intervalo cerrado* $[a, b]$*, los conjuntos Lebesgue medibles en* $[a, b]$ *y la medida de Lebesgue y sea* E *un conjunto Lebesgue medible en* $[a, b]$*. Entonces,*

i) *Para todo* $\varepsilon > 0$ *existe una conjunto cerrado* F *tal que* $F \subseteq E$ *y*

$$\lambda(E \setminus F) = \lambda(E \cap F^c) < \varepsilon.$$

ii) *Para todo* $\varepsilon > 0$ *existe un conjunto abierto* O *tal que* $E \subseteq O$ *y*

$$\lambda(O \setminus E) = \lambda(O \cap E^c) < \varepsilon$$

Por úlimo, se tiene la siguiente terminología.

Definición 23. Decimos que una propiedad P se cumple casi en todo punto ("*c.t.p.*"), si existe un $E \subseteq \mathbb{R}$ medible, tal que $\lambda(E) = 0$ y la propiedad P se cumple en E^c.

En otras palabra, una propiedad se cumple *c.p.t.* si esta se cumple siempre, salvo quizá, en algún conjunto de medida cero.

4.3. Funciones medibles

Antes de definir la integral de Lebesgue de una función, definamos el siguiente conjunto de funciones.

Definición 24. Una función $f : \mathbb{R} \to \overline{\mathbb{R}}$ se dice Lebesgue medible si cumple que $\{x \in \mathbb{R} : f(x) \geqslant \alpha\} \in \mathcal{L}(\mathbb{R})$ para todo $\alpha \in \mathbb{R}$.

De manera similar, una función $f : [a, b] \longrightarrow \overline{\mathbb{R}}$ es dice Lebesgue medible si cumple que $\{x \in \mathbb{R} : f(x) \geqslant \alpha\} \in \mathcal{L}([a, b])$ para todo $\alpha \in \mathbb{R}$.

En este documento, siempre que se haga referencia a funciones medibles y no se especifique nada más, nos estaremos refiriendo a funciones Lebesgue medibles.

Definición 25. Al conjunto de todas las funciones $f : \mathbb{R} \to \overline{\mathbb{R}}$ Lebesgue medibles se les denota por $M(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ y al conjunto de funciones $f \in M(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ tales que $f \geqslant 0$ se les denota por $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$. Mientras que al conjunto de todas las funciones $f : [a, b] \longrightarrow \overline{\mathbb{R}}$ Lebesgue medibles se les denota por $M(\mathbb{R}, \mathcal{L}([a, b]))$ y al conjunto de funciones $f \in M(\mathbb{R}, \mathcal{L}([a, b]))$ tales que $f \geqslant 0$ se les denota por $M^+(\mathbb{R}, \mathcal{L}([a, b]))$.

Estos conjuntos serán de vital importancia para este estudio pues serán dichas funciones sobre las cuales definiremos la intergal de Lebesgue. Tal y como ocurrió en el caso de los conjuntos medibles, no será posible definir la integral de Lebesgue para todas las funciones existentes, se debe de ser un poco más selectivos para evitar perder propiedades importantes de la integral.

El siguiente teorema establece que bajo las operaciones básicas entre funciones medibles, la mesurabilidad se preserva.

Teorema 18. *Sean f y g funciones Lebsegue medibles y c una constante, entonces cf, f^2, $f+g$, fg, $|f|$, $\text{mín}\,(f,g)$ y $\text{máx}\,(f,g)$ son funciones Lebsegue medibles.*

Así mismo, se tiene el siguiente resultado que habla del límite de una sucesión de funciones medibles.

Teorema 19. *Sea (f_n) una sucesión de funciones en Lebsegue medibles tales que*

$$\lim_{n\to\infty} f_n(x) = f(x).$$

Entonces, f también es una función Lebesgue medible.

Otro resultado importante de las funciones Lebesgue medibles es el siguiente.

Teorema 20. *Sea (f_n) una sucesión de funciones Lebsegue medibles. Entonces, para cada $x \in \mathbb{R}$, las funciones definidas por*

$$\begin{aligned}
&\sup\{f_n(x) : n \in \mathbb{N}\}\\
&\inf\{f_n(x) : n \in \mathbb{N}\}\\
&\limsup f_n(x)\\
&\liminf f_n(x).
\end{aligned}$$

son Lebesgue medibles.

Demostraciones de los tres teoremas anteriores pueden ser encontradas en el capítulo 2 de [**2**]

Como último resultado de esta sección, aprovecharemos el hecho de no habernos quedado únicamente con la $\sigma - \acute{a}lgebra$ de Borel y haberla extendido un poco más para tener la $\sigma - \acute{a}lgebra$ de Lebesgue mediante el siguiente resultado.

Teorema 21. *Sea f una función Lebesgue medible. Si $g = f$ casi en todo punto, entonces g también es medible.*

Demostración. Sea $\alpha \in \mathbb{R}$ arbitrario. Notemos que el conjuno

$$B = \{x \in \mathbb{R} : f(x) \neq g(x)\}$$

es un conjunto medible, pues este tiene medida cero. Por lo que $\{x \in \mathbb{R} : g(x) \geqslant \alpha\}$ puede ser expresado como

$$\begin{aligned}
&(\{x \in \mathbb{R} : f(x) \geqslant \alpha\} \cap B^c) \cup (\{x \in \mathbb{R} : g(x) \geqslant \alpha\} \cap B) =\\
&\{x \in \mathbb{R} : f(x) \geqslant \alpha \ \wedge \ g(x) = f(x)\} \cup \{x \in \mathbb{R} : g(x) \geqslant \alpha \ \wedge \ g(x) \neq f(x)\}.
\end{aligned}$$

Donde ambos conjuntos de la unión son medibles; el primero al ser resultado de la intersección de conjuntos medibles y el segundo por ser subconjunto de un conjunto de medida nula. ■

Una función que será de mucha ayuda a lo largo del desarrollo de toda la teoría es la función característica de un conjunto.

Definición 26. Sea $E \subseteq \mathbb{R}$. Definimos la función característica de E, denotada por χ_E, como

$$\chi_E(x) = \begin{cases} 1 & \text{si } x \in E \\ 0 & \text{si } x \in E^c \end{cases}$$

Claramente, la función característica de E será una función medible si y solo si E es un conjunto medible.

4.4. La intergal de Lebesgue y algunas de sus propiedades básicas

El primer paso para definir la integral de Lebesgue de una función medible f, es definir la integral de funciones más simples.

Definición 27. Una función φ que pertenece a $M(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ se llama simple, si solo toma un número finito de valores ninguno de los cuales es ∞ ó $-\infty$. En otras palabras

$$\varphi : \mathbb{R} \longrightarrow \{a_1, a_2, ..., a_n\} \subseteq \mathbb{R} \text{ donde } a_j \neq a_i \text{ si } i \neq j.$$

Notemos que

$$\varphi = \sum_{i=1}^{n} a_j \chi_{E_j} \tag{*}$$

donde $E_j = \varphi^{-1}(a_j)$ y χ_{E_j} es la función característica de E_j. A esta manera de representar a φ se le llama representación estándar de φ.

La razón para contemplar primero este tipo de funciones simples, es por el hecho de que definir su integral resulta algo muy natural. Posteriormente y basándose en esto, se podrá definir la integral de Lebesgue para cualquier función medible.

Definición 28. Sea φ una función simple que pertenece a $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ con representación estándar $(*)$. Definimos la integral de Lebesgue de φ con respecto de λ como

$$\int \varphi d\lambda = \sum_{i=1}^{n} a_j \lambda(E_j)$$

Haciendo la convención de que $0 \cdot \infty = 0$.

Resulta algo muy razonable el definir la integral de las funciones simples como esa suma, pues es es la suma de las áreas de los "rectángulos" formados por el eje x y la gráfica de la función simple φ. Por el momento, se restringen estas ideas a las funciones simples positivas debido a que se quieren evitar situaciones donde aparezcan expresiones del tipo $\infty - \infty$, lo cual, no está definido.

A continuación, se definirá lo que significa la integral de una función simple sobre un conjunto E y un resultado de este nuevo concepto.

Definición 29. Sea φ una función simple que pertenece a $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ y λ la medida de Lebesgue. Para cualquier conjunto E que sea Lebesgue medible, se cumple que

$$\varphi\chi_E$$

es una función medible y además, también es una función simple. Por lo tanto, su integral existe. Definimos y denotamos a la integral de φ sobre el conjunto E como

$$\int_E \varphi d\lambda := \int \varphi\chi_E \, d\lambda$$

Teorema 22. *Sea φ una función simple que pertenece a $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ y sea λ la medida de Lebesgue. Entonces, la función $\mu : \mathcal{L}(\mathbb{R}) \longrightarrow \overline{\mathbb{R}}^+ \cup \{0\}$ definida como*

$$\mu(E) = \int_E \varphi d\lambda$$

es una medida.

Aunque ahora el teorema 22 no sea muy útil para el objetivo actual de definir la integral de Lebesgue para funciones medibles arbitrarias, será de ayuda más adelante cuando se vea la demostración del teorema de la convergencia monótona.

El último escalón que queda por subir antes de poder definir la integral de Lebesgue para funciones medibles arbitrarias, es definir la integral de Lebesgue para funciones medibles no negativas.

Definición 30. Si f es una función en $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$, definimos la integral de Lebesgue de f con respecto de λ como

$$\int f d\lambda = \sup\left\{\int \varphi d\lambda \,:\, \varphi \in M^+(\mathbb{R}, \mathcal{L}(\mathbb{R})) \text{ es simple y } 0 \leqslant \varphi(x) \leqslant f(x) \ \forall x \in \mathbb{R}\right\} \tag{12}$$

Además, si E es un conjunto Lebesgue medible, definimos la integral de f sobre E como

$$\int_E f d\lambda := \int f\chi_E \, d\lambda.$$

Con la siguiente definición quedará claro por qué resultó útil definir primero la integral de las funciones medibles no negativas.

Definición 31. Sea f una función que pertenece a $M(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Definimos a la parte positiva y negativa de f, denotadas por f^+ y f^-, como

$$f^+(x) := \text{máx}(0, f(x))$$

$$f^-(x) := \text{máx}(0, -f(x))$$

Haciendo la observación de que $f = f^+ - f^-$, por el teorema 18 queda claro que f es una función medible si y solo si f^+ y f^- son funciones medibles. Además f^+ y f^- son funciones no negativas y por lo tanto, su integral de Lebesgue está bien definida.

Definición 32. Sea f una funcion que pertenece a $M(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ y λ la medida de Lebesgue. Diremos que f es integrable con respecto de λ o simplemente que f es Lebesgue integrable si

$$\int f^+ d\lambda, \int f^- d\lambda < \infty$$

en cuyo caso, definimos a la inegral de Lebesgue de f con respecto de λ como

$$\int f d\lambda := \int f^+ d\lambda - \int f^- d\lambda.$$

Además, si E es un conjunto Lebesgue medible, definimos la integral de f sobre el conjunto E como

$$\int_E f d\lambda := \int_E f^+ d\lambda - \int_E f^- d\lambda.$$

Recapitulando un poco como fue la construcción hasta llegar a definir la integral de Lebesgue; primero se definió lo que es una función simple, después, se dijo qué es la integral de Lebesgue para funciones simples no negativas y finalmente, se definió la integral de Lebesgue para funciones medibles arbitrarias, pero no negativas, que es realmente aquí donde descansa la idea de la integral de Lebesgue. Pero ¿por qué pedirle a la función f la propiedad de ser medible si el supremo en (12) queda perfectamente bien definido ya sea si la función f es medible o no? Resulta ser que, si se quita esa restricción y se define la integral de Lebesgue para cualquier función, entonces se pierden propiedades importantes de la integral, por ejemplo, la propiedad de linealidad que se verá a continuación. Una demostración de estas propiedades pueden ser encontrada en [**2**], a excepción de la propiedad III) que se sigue de aplicar el teorema 25, enunciado más adelante, a $f - g$.

Teorema 23. *Sean $f, g \in M(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, k una constante real y λ la medida de Lebesgue. Supongamos que f y g son ambas Lebesgue integrables. Entonces,*

I*) La función $f + kg$ es Lebesgue integrable y se cumple que*

$$\int (f + kg)\, d\lambda = \int f\lambda + k \int g d\lambda.$$

II*) Si $f \leqslant g$, entonces*

$$\int f d\lambda \leqslant \int g d\lambda.$$

III*) Si $f = g$ casi en todo punto, entonces*

$$\int f d\lambda = \int g d\lambda.$$

El siguiente resultado se sigue inmediatamente de la definición de integral de Lebesgue para funciones medibles no negativas. Aunque puede no parecer la gran cosa, afirma que una función medible es integrable siempre que exista una función integrablle que la "domine". Cuestión que no sucede con la integral de Riemann, aunque una función Riemann integrable domine a otra, esta útlima no necesariamente es Riemann integrable.

Teorema 24. *Sean f y g funciones en $M(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ tales que g es Lebesgue integrable y $|f| \leqslant g$, entonces f también es Lebesgue integrable.*

Demostración. Vamos a demostrar que $\int f^+ d\lambda < \infty$ y lo mismo se sigue para la integral de f^-. Sea $\varphi \in M^+(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ tal que $0 \leqslant \varphi \leqslant f^+$. Entonces, $\varphi \leqslant g$ y por el inciso III) del teorema 23 concluimos que

$$\int \varphi d\lambda \leqslant \int g d\lambda < \infty.$$

Ya que φ fue una función medible simple y arbitraria que queda por debajo de f^+, concluimos que $\int f^+ d\lambda \leqslant \int g d\lambda < \infty$. ■

Cuando la integral de una función medible y no negativa es cero, esto debería ser consecuencia de que el conjunto de puntos para los cuales la función es mayor que cero, tiene una medida muy pequeña. En concreto, el siguiente resultado nos dice que esta situación se da si y solo si la medida de ese conjunto es cero.

Teorema 25. *Sea $f : \mathbb{R} \longrightarrow \overline{\mathbb{R}}$ una función en $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$. Entonces, $\int f d\lambda = 0$ si y solo si f es cero casi en todo punto.*

Demostración. Supongamos que $\int f d\lambda = 0$. Para cada $n \in \mathbb{N}$ definimos a

$$E_n = \left\{ x \in \mathbb{R} : f(x) > \frac{1}{n} \right\}$$

Claramente $E_n \subseteq E_{n+1}$ y además, $\frac{1}{n}\chi_{E_n} \leqslant f$ para todo $n \in \mathbb{N}$. Por lo tanto,

$$0 = \int f d\lambda \geqslant \int \frac{1}{n}\chi_{E_n} d\lambda = \frac{1}{n}\lambda(E_n) \geqslant 0$$

De donde $\lambda(E_n) = 0$ para todo $n \in \mathbb{N}$. Luego, aplicando el teorema 16 concluimos que

$$0 \leqslant \lambda(\{x \in \mathbb{R} : f(x) > 0\}) = \lambda\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \lambda(E_n) = 0.$$

Ahora, supongamos que $f = 0$ casi en todo punto. Simplemente hay que notar que para cualquier función simple $\varphi \in M^+(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ tal que $0 \leqslant \varphi \leqslant f$, se cumple que $\int \varphi d\lambda = 0$. En caso contrario, existiría algún E_j en la prepresentación estandar (*) de φ tal que $\lambda(E_j) > 0$ y $a_j > 0$. Lo que implicaría que $f \geqslant a_j \neq 0$ en un conjunto de medida mayor a cero. ■

El siguiente resultado recuerda a la propiedad V) del teorema 6 de propiedades de la integral de Riemann, con la gran diferencia de que aquí, estamos hablando de funciones definidas en todo $\mathbb{R}$, minetras que aquel resulatdo solo es cierto cuando la función está definida en intervalos acotados de la forma $[a, b]$.

Teorema 26. *Una función $f \in M(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ es integrable si y solo si $|f|$ es integrable, en cuyo caso*

$$\int |f| d\lambda = \int f^+ d\lambda + \int f^- d\lambda.$$

Demostración. Recordemos que f es Lebesgue integrable si y solo si $\int f^+ d\lambda$, $\int f^- d\lambda < \infty$ si y solo si $\int f^+ d\lambda + \int f^- d\lambda = \int (f^+ + f^-) d\lambda < \infty$. Pero $|f| = f^+ + f^-$, por lo que f es Lebesgue integrable si y solo si $\int |f| d\lambda < \infty$. ■

Un resultado importante de las funciones Lebesgue integrables es el siguiente.

Teorema 27. *Sea f una función en $M(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ integrable. Entonces, para todo $\varepsilon > 0$, existe un $\delta > 0$ tal que si $\lambda(E) < \delta$ esto implica que*

$$\left| \int_E f d\lambda \right| < \varepsilon.$$

Demostración. Demostremos el caso cuando f es no negativa. Sea $\varepsilon > 0$ arbitrario, entonces existe una función simple φ en $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ tal que $0 \leqslant \varphi \leqslant f$ y además

$$0 \leqslant \int f d\lambda - \int \varphi d\lambda = \int (f - \varphi)\, d\lambda < \frac{\varepsilon}{2}.$$

Ya que $0 \leqslant (f - \varphi) \cdot \chi_E \leqslant f - \varphi$ para cualquier conjunto E, se sigue que

$$0 \leqslant \int_E (f - \varphi)\, d\lambda < \frac{\varepsilon}{2}$$

para cualquier conjunto medible E. De donde

$$0 \leqslant \int_E f d\lambda < \frac{\varepsilon}{2} + \int_E \varphi d\lambda.$$

Ahora, si E es un conjunto medible tal que $\lambda(E) < \frac{\varepsilon}{2M}$, donde $M = \sup\{\varphi(x) : x \in \mathbb{R}\}$, entonces se cumple que

$$0 \leqslant \int_E f d\lambda < \frac{\varepsilon}{2} + \int_E \varphi d\lambda < \frac{\varepsilon}{2} + \int_E M d\lambda = \frac{\varepsilon}{2} + M\lambda(E) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

El caso para funciones medibles arbitrarias es inmediato, pues si f es Lebesgue integrable, entonces $|f|$ también lo es y además se cumple que

$$\left| \int_E f d\lambda \right| \leqslant \int_E |f|\, d\lambda$$

por lo que aplicando lo recien demostrado a $|f|$ habríamos terminado. ■

4.5. Teoremas de intercambio entre límite e integral de Lebesgue (TCM, TCD y lema de Fatou)

Una parte importante de la teoría de integración de Lebesgue son sus teoremas que dictan bajo qué condiciones es válido hacer el intercambio entre el símbolo de la integral y el de límite. El primer resultado de este tipo que se va a presentar, es conocido como teorema de la convergencia mónotona, o abreviando, TCM.

Teorema 28. ***(Teorema de la convergencia monótona)*** *Sea (f_n) una sucesión de funciones en $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ tal que la sucesión (f_n) crece a f. Entonces,*

$$\lim_{n\to\infty} \int f_n d\lambda = \int f d\lambda = \int \lim_{n\to\infty} f_n d\lambda.$$

Demostración. Para empezar, deberíamos asegurarnos de que f es una función medible para poder hablar de su integral, pero esto se sigue del teorema 19. Además, como

$$f_1 \leqslant f_2 \leqslant \dots f_n \leqslant \dots f,$$

esto implica que la sucesión de números reales extendidos $\left(\int f_n d\lambda\right)$ es creciente y se cumple que

$$\lim_{n\to\infty} \int f_n d\lambda \leqslant \int f d\lambda.$$

Ahora, tomemos un $\alpha \in (0,1)$ y φ una función simple en $M^+(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ tal que $0 \leqslant \varphi \leqslant f$.

A continuación, definamos los conjuntos

$$A_n = \{x \in \mathbb{R} : f_n(x) \geqslant \alpha\varphi(x)\}.$$

Entonces, se cumple que $A_n \subseteq A_{n+1}$ para todo $n \in \mathbb{N}$, pues (f_n) es creciente. Además, cada A_n es un conjunto medible, pues

$$A_n = \{x \in \mathbb{R} : f_n(x) - \alpha\varphi(x) \geqslant 0\}$$

donde f_n y $\alpha\varphi$ son ambas funciones medibles. Por otro lado, debido a que $f_n(x) \longrightarrow f(x) \geqslant \varphi(x) > \alpha\varphi(x)$ para todo $x \in \mathbb{R}$, se sigue que

$$\mathbb{R} = \bigcup_{n=1}^{\infty} A_n.$$

Luego, gracias a como está definido el conjunto A_n, tenemos que

$$\int_{A_n} \alpha\varphi d\lambda \leqslant \int_{A_n} f_n d\lambda \leqslant \int f_n d\lambda. \tag{13}$$

Si definimos a $\mu : \mathcal{B}(\mathbb{R}) \longrightarrow \overline{\mathbb{R}^+} \cup \{0\}$ como

$$\mu(E) = \int_E \alpha\varphi d\lambda$$

sabemos, por el teorema 22, que μ define una medida y se sigue por el teorema 16 que

$$\int \alpha\varphi d\lambda = \mu(\mathbb{R}) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \int_{A_n} \alpha\varphi d\lambda.$$

Con esto, y haciendo tender n a infinito en (13), llegamos a que

$$\alpha \int \varphi d\lambda \leqslant \lim_{n\to\infty} \int f_n d\lambda.$$

Tomando ahora el límite cuano α tiende a 1 por la izquierda, obtenemos

$$\int \varphi d\lambda \leqslant \lim_{n\to\infty} \int f_n d\lambda.$$

Y ya que φ es una función simple, medible y abitraría que cumple con $0 \leqslant \varphi \leqslant f$, se concluye que

$$\int f d\lambda \leqslant \lim_{n\to\infty} \int f_n d\lambda$$

■

De este resultado se sigue que, cuando (f_n) es una sucesión de funciones Lebesgue medibles no positivas que decrecen a una función f, entonces el cambio entre límite e integral también es válido. Cosa que no es verdad para sucesiones decrecientes de funciones no negativas o para sucesiones crecientes de funciones no positivas. Por ejemplo, si (f_n) es la sucesión de funciones definida en $\mathbb{R}$ dada por

$$f_n(x) = \frac{1}{n}$$

es claro que la sucesión decrece a cero y cada término es no negativo, pero

$$\lim_{n\to\infty} \int f_n d\lambda = \infty \neq 0 = \int 0 d\lambda. \tag{14}$$

Del TCM se siguen varios resultados interesantes como los siguienes.

Corolario 3. *Sea (f_n) una sucesión de funciones en $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$. Entonces,*

$$\sum_{n=1}^{\infty} \int f_n d\lambda = \int \left(\sum_{n=1}^{\infty} f_n \right) d\lambda.$$

Demostración. Es inmediato, pues si para cada $k \in \mathbb{N}$ definimos a $g_k = \sum_{n=1}^{k} f_n$, entonces (g_k) es una sucesión de funciones que crecen a $g = \sum_{n=1}^{\infty} f_n$. Luego, por el TCM, concluimos que

$$\begin{aligned}\sum_{n=1}^{\infty} \int f_n d\lambda = \lim_{k\to\infty} \sum_{n=1}^{k} \int f_n d\lambda &= \lim_{k\to\infty} \int \sum_{n=1}^{k} f_n d\lambda \\ &= \lim_{k\to\infty} \int g_k d\lambda = \int g d\lambda = \int \left(\sum_{n=1}^{\infty} f_n \right) d\lambda\end{aligned}$$

■

Lo que el corolario anterior nos transmite es que podemos intercambiar el símbolo de la integral por el de la suma, si las funciones involucradas son medibles y no negativas.

Teorema 29. *Sea f una función en $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ Lebesgue integrable. Entonces, la función $\mu : \mathcal{L}(\mathbb{R}) \longrightarrow \overline{\mathbb{R}^+} \cup \{0\}$ definida por*

$$\mu(E) = \int_E f d\lambda$$

define una medida.

Demostración. Es claro que $\mu(\varnothing) = 0$.

Sea (A_n) una sucesión de conjuntos medibles y disjuntos entre sí. Entonces, $g_n = f \cdot \chi_{A_n}$ define una sucesión de funciones no negativas, por lo que podemos aplicar el corolario anterior y concluir que

$$\begin{aligned}\mu\left(\bigcup_{n=1}^{\infty} A_n \right) &= \int_{\cup A_n} f d\lambda = \int \left(\sum_{n=1}^{\infty} f \cdot \chi_{A_n} \right) d\lambda \\ &= \sum_{n=1}^{\infty} \int f \cdot \chi_{A_n} d\lambda = \sum_{n=1}^{\infty} \int_{A_n} f d\lambda = \sum_{n=1}^{\infty} \mu(A_n).\end{aligned}$$

■

Hasta antes de este punto, se sabía que el resultado anterior era cierto si la función involucrada era medible, simple y no negativa. Pero ahora con el TCM se pudo demostrar que esto se cumple siempre que la función sea medible y no negativa.

El siguiente de esta serie de teoremas representativos de la integral de Lebesgue es el lema de Fatou.

Teorema 30. ***(Lema de Fatou)*** *Sea* $(\mathbb{R}, \mathcal{L}(\mathbb{R}), \lambda)$ *el espacio de medida formado por los reales, la* $\sigma-$*álgebra de Lebesgue y la medida de Lebesgue y* (f_n) *una sucesión de funciones en* $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ *. Entonces,*

$$\int \liminf f_n \, d\lambda \leqslant \liminf \int f_n d\lambda.$$

Demostración. Definamos a $g_m = \inf \{f_m, f_{m+1}, \ldots\}$. Luego

$$g_m \leqslant f_n$$

para todo $n \geqslant m$. Por lo tanto,

$$\int g_m d\lambda \leqslant \int f_n d\lambda$$

para todo $n \geqslant m$. Además, la sucesión (g_m) es creciente. Tomando ínfimo de esta última expresión llegamos a que

$$\int g_m d\lambda = \inf_{n \geqslant m} \int g_n d\lambda \leqslant \inf_{n \geqslant m} \int f_n d\lambda.$$

Ahora, tomemos el límite cuando m tiende a ínfinito para concluir que

$$\lim_{m \to \infty} \int g_m d\lambda \leqslant \lim_{m \to \infty} \inf_{n \geqslant m} \int f_n d\lambda = \liminf \int f_n d\lambda.$$

Por otro lado, como ya habíamos comentado, (g_m) es una sucesión creciente de funciones medibles no negativas que, de hecho, converjen a $\liminf f_n$. Por el teorema de la convergencia monótona concluimos que

$$\int \liminf f_n \, d\lambda = \lim_{m \to \infty} \int g_m d\lambda \leqslant \liminf \int f_n d\lambda.$$

■

Como se puede apreciar, el lema de Fatou tiene condiciones mucho menos restrictivas a cerca de la sucesión de funciones. Concretamente, ya no se está suponiendo que la sucesión converja puntualmente a una función límite. Por lo mismo, ya no podemos asegurar un cambio entre límite e integral, quedándonos en su lugar con esa desigualdad con limítes inferiores. Sin embargo, esto es suficiente, por ejemplo, para dar una demostración inmediata del TCM, pues si se vuelve un poco atras se puede ver que la parte extensa de la demostración de ese teorema era la desigualdad (14) que, de hecho, se sigue inmediatamente del lema de Fatou. Esto quiere decir que en realidad el lema de Fatou y el TCM son resultados equivalentes.

Como último de estos resultados característicos de la integral de Lebesgue, se tiene el teorema de la convergencia dominada.

Teorema 31. ***(Teorema de la convergencia dominada)*** *Sea (f_n) una sucesión de funciones en $M(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ que convergen puntualmente a f. Supongamos que existe una función $p \in M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$ Lebesgue integrable tal que $|f_n| \leqslant p$ para todo $n \in \mathbb{N}$. Entonces, f y cada f_n son Lebesgue integrables y además*

$$\lim_{n\to\infty} \int |f - f_n| \, d\lambda = 0.$$

En particular

$$\lim_{n\to\infty} \int f_n d\lambda = \int f d\lambda.$$

Demostración. Por el teorema 26 y por la propiedad 2 del teorema 23, concluimos que cada f_n y f son integrables. Ahora, definamos para cada $m \in \mathbb{N}$

$$g_m = \sup_{n\geqslant m} \{|f - f_n|\}.$$

Por el teorema 20, sabemos que cada g_m es medible y aún más

$$|f - f_n| \leqslant |f| + |f_n| \leqslant 2p$$

para todo $n \in \mathbb{N}$. Con lo cual, concluimos que cada g_m está dominada por una función intergable y por lo tanto, es integrable. No es difícil ver que la sucesión (g_m) decrece a cero, razón por la cual, la sucesión $h_k = g_1 - g_k$ es una sucesión creciente de funciones en $M^+(\mathbb{R}, \mathcal{L}(\mathbb{R}))$. Aplicando el teorema de la convergencia monótona, concluimos que

$$\int g_1 d\lambda = \int \left(\lim_{k\to\infty} h_k\right) d\lambda = \lim_{k\to\infty} \int h_k d\lambda = \lim_{k\to\infty} \int g_1 d\lambda - \int g_k d\lambda.$$

De donde

$$\lim_{k\to\infty} \int g_k d\lambda = 0.$$

Pero,

$$0 \leqslant \lim_{k\to\infty} \int |f - f_k| \, d\lambda \leqslant \lim_{k\to\infty} \int g_k d\lambda = 0.$$

En particular

$$0 \leqslant \lim_{k\to\infty} \left| \int f d\lambda - \int f_k d\lambda \right| \leqslant \lim_{k\to\infty} \int |f - f_k| \, d\lambda = 0.$$

Terminando la demostración. ■

4.6. Comparación entre la integral de Riemann y la integral de Lebesgue y caracterización de las funciones Riemann integrables

Esta sección comienza retomando el tema que se mencionó al final de la sección 3.4 donde se exponen el TCM y TCD para la integral de Riemann. Para analizar mejor la diferencia entre las hipótesis del TCM y TCD de ambas integrales, enunciaremos el TCM para la intergal de Riemann a continuación.

(TCM para la integral de Riemann) *Sea (f_n) una sucesión monótona de funciones definidas en $[a,b]$ que son Riemann integrables y que convergen puntualmente a una función f Riemann integrable. Entonces, el intercambio entre integral de Riemann y límite es válido.*

Hay dos cambios que se pueden notar casi de inmediato; La sucesión de funciones no necesariamente tiene que tratarse de una sucesión creciente de funciones no negativas, o de una sucesión decreciente de funciones no positivias. Simplemente, tienen que ser funciones Riemann integrables. Sin embargo, no se debe de olvidar que la integral de Riemann solo está definida para funciones acotadas, es decir, para cualquier sucesión monótona, digamos creciente, se puede encontrar una constante $K > 0$ tal que $g_n = f_n + K$ sea una sucesión creciente de funciones no negativas. Razón por la cual, en el caso de la integral de Riemann, suponer que el cambio entre integral y límite se cumple para sucesiones crecientes de funciones Riemann integrables no negativas, es equivalente a decir que se cumple para sucesiones crecientes de funciones Riemann integrables arbitrarias. El segundo cambio resulta más interesante. En el caso de la integral de Riemann, se tiene que garantizar la integrabilidad de la función límite f y para ver la necesidad de este supuesto, recurramos al siguiente ejemplo: Sea $Q : \mathbb{N} \longrightarrow \mathbb{Q} \cap [0,1]$ una biyección entre los números naturales y los racionales del intervalo $[0,1]$. A $q_n = Q(n)$ se le conoce como numeración de los racionales. Definamos la sucesión de funciones en $[0,1]$

$$f_n(x) = \begin{cases} 1 & \text{si } x = q_j \text{ para algún } j \leqslant n \\ 0 & \text{cualquier otro caso} \end{cases}$$

Es claro que la sucesión (f_n) es creciente y además, cada término de la sucesión es Riemann integrbale con integral igual a cero. Sin embargo, la función límite conocida como función de Dirichlet

$$f(x) = \begin{cases} 1 & \text{si } x \in \mathbb{Q} \cap [0,1] \\ 0 & \text{cualquier otro caso} \end{cases}$$

resulta no ser Riemann integrable como ya se discutió en su momento cuando se definió la integral de Darboux en la sección 3.2. Por lo tanto, a pesar de ser una sucesión creciente de funciones Riemann integrables, el límite de la sucesión resulta no ser Riemann intergrable. Motivo por el cual, no tiene sentido hablar de un cambio ente límite e integral.

Continuando con lo discutido recien de la función de Dirichlet, este es un ejemplo de una función que no es Riemann integrable, pero que si es Lebesgue integrable. Precisamente la función de Dirichlet es una función medible y aún más, es cero casi en todo punto. Por lo que

$$\int_{[a,b]} f d\lambda = 0.$$

La pregunta ahora sería si existen funciones que sean Riemann integrables, pero que no lo sean en el sentido de Lebesgue. La respuesta a esto se dá a continuación, en el único resultado de esta sección. Dicho resultado está inspirado en las ideas presentadas en **[11]** . En esta demostración, se usan mayormente caracterizaciones y teoría de la integral de Darboux, por lo que los interesados pueden consultar la secciones 3.2 y 3.3 para ver la teoría básica de la integral de Darboux y la equivalencia entre la integral de Darboux y la de Riemann, respectivamente. Pero antes de ver la demostración, es necesario establecer un poco de notación.

Sea $f:[a,b]\longrightarrow\mathbb{R}$ una función acotada, $P=\{x_0,x_1,\ldots,x_n\}$ una partición de $[a,b]$ y m_i y M_i interpretados como se dan en la definición 13 para $i=1,2,\ldots,n$. Definimos a $g_P,G_P:[a,b]\longrightarrow\mathbb{R}$ como

$$g_P(t)=m_i \text{ si } t\in(x_{i-1},x_i] \text{ para } i=1,\ldots,n$$

y

$$G_P(t)=M_i \text{ si } t\in(x_{i-1},x_i] \text{ para } i=1,\ldots,n.$$

No es difícil ver que g_P y G_P son funciones medibles y simples. Además, se cumple que

$$\int_{[a,b]} g_P d\lambda = L(f,P)$$

y

$$\int_{[a,b]} G_P d\lambda = U(f,P).$$

Teorema 32. *Sea $f:[a,b]\longrightarrow\mathbb{R}$ una función acotada.*

I) *Si f es Riemann integrable, entonces f es Lebesgue medible y se cumple que*

$$\int_a^b f(x)dx=\int_{[a,b]} f d\lambda.$$

II) *La función f será Riemann integrable si y solo si el conjunto de puntos de discontinuidades de f tiene medida de cero.*

Demostración. (I): Ya que f es Riemann integrable, gracias al inciso (II) del teorema 7, podemos encontrar una sucesión de particiones Q_k tales que

$$U(f,Q_k)-L(f,Q_k)<\frac{1}{k} \tag{15}$$

para $k=1,2,\ldots$. Por el teorema 4, si definimos a $P_k=\cup_{n=1}^{k}Q_n$, entonces se sigue cumpliendo (15) y P_{k+1} es un refinamiento de P_k para todo $k\in\mathbb{N}$. Esto implica que

$$g_{P_1}\leqslant g_{P_2}\leqslant\ldots\leqslant f\leqslant\ldots\leqslant G_{P_2}\leqslant G_{P_1}.$$

Por lo tanto, al tratarse de sucesiones monótonas y acotadas, (g_{P_k}) y (G_{P_k}) convergen puntualmente digamos a las funciones g y G respectivamente. Aún más, g y G son funciones medibles, pues son el límite de funciones medibles y además se cumple que $g\leqslant f\leqslant G$. Si aplicamos ahora el teorema de la convergencia dominada a las sucesiones (g_{P_k}) y (G_{P_k}) podemos decir que

$$\lim_{k\to\infty} L(f,P_k)=\lim_{k\to\infty}\int_{[a,b]} g_{P_k}d\lambda=\int_{[a,b]} g d\lambda$$

y

$$\lim_{k\to\infty} U(f,P_k)=\lim_{k\to\infty}\int_{[a,b]} G_{P_k}d\lambda=\int_{[a,b]} G d\lambda.$$

A continuación, vamos a demostrar que $\lim_{k\to\infty} U(f,P_k)=\overline{\int_a^b f(x)dx}$ y se podría demostrar lo análogo con la integral inferior. Tomemos un $\varepsilon>0$ arbitrario, esto quiere decir que podemos encontrar un $k_0 \in \mathbb{N}$ tal que $\frac{1}{k_0} < \varepsilon$. Ahora, veamos que si $k \geqslant k_0$, entonces

$$0 \leqslant U(f, P_k) - \overline{\int_a^b} f(x)dx \leqslant U(f, P_k) - \underline{\int_a^b} f(x)dx \leqslant U(f, P_k) - L(f, P_k) < \frac{1}{k} < \varepsilon.$$

Esto debido a la relación que se da entre la integral inferior, la integral superior y las sumas superiores e inferiores. De aquí concluimos varias cosas. Para empezar, recordemos que la integral inferior y superior de f coinciden por ser Riemann integrable y por lo tanto, tenemos que

$$\int_{[a,b]} g d\lambda = \underline{\int_a^b} f(x)dx = \int_a^b f(x)dx = \overline{\int_a^b} f(x)dx = \int_{[a,b]} G d\lambda.$$

Además, recordemos que $g \leqslant f \leqslant G$ de manera que $G - g \geqslant 0$, pero como acabamos de ver

$$\int_{[a,b]} (G - g)\, d\lambda = \int_{[a,b]} G d\lambda - \int_{[a,b]} g d\lambda = 0.$$

Gracias al teorema 25, concluimos que $G = g$ *c.t.p.* Lo que quiere decir que f también es igual a G casi en todo punto. Esto implica, por el teorema 21, que f es medible y aún más

$$\int_{[a,b]} f d\lambda = \int_{[a,b]} G d\lambda = \int_a^x f(x)dx.$$

(II): Supóngase que f es Riemann integrable. Haciendo uso de lo demostrado en el apartado (I) de este resultado, tenemos que $g = f = G$ casi en todo punto. Es decir, existe un conjunto medible E tal que $g = f = G$ en E^c y $\lambda(E) = 0$. Definamos a H como $(\cup_{k=1}^{\infty} P_k) \cup E$. Si lo gramos demostrar que f es continua en H^c habremos terminado. Sea $t \in H^c$ y $\varepsilon > 0$ arbitrario. Ya que $g_{P_k} \longrightarrow g$ y $G_{P_k} \longrightarrow G$, entonces existe un $k \in \mathbb{N}$ tal que

$$g(t) - g_{P_k}(t) = |g(t) - g_{P_k}(t)| < \frac{\varepsilon}{2}$$

y

$$G_{P_k}(t) - G(t) \leqslant |G_{P_k}(t) - G(t)| < \frac{\varepsilon}{2}.$$

Por otro lado, ya que $t \in H^c$, esto quiere decir que $g(t) = G(t)$. De donde

$$G_{P_k}(t) - g_{P_k}(t) < \varepsilon.$$

Pero $G_{P_k}(t) = M_i$ y $g_{P_k}(t) = m_i$ para algún intervalo (x_{i-1}, x_i) de la partición P_k. Ahora simplemente habremos de notar que si $t, y \in (x_{i-1}, x_i)$, entonces $m_i \leqslant f(t), f(y) \leqslant M_i$ y por lo tanto se cumple que

$$|f(t) - f(y)| < M_i - m_i = G_{P_k}(t) - g_{P_k}(t) < \varepsilon.$$

Acabamos de encontrar un intervalo abierto que contiene a t donde las imagenes de la función f en el intervalo, caen tan cerca como se quiera de la imagen de f en t, por lo que concluimos que f es continua en t.

Para terminar, supongamos ahora que f es continua casi en todo punto. Sea (P_k) una sucesión de partciones tales que P_{k+1} es un refinamiento de P_k y las normas de las particiones P_k decrecen a cero. Definimos las sucesiones (g_{P_k}) y (G_{P_k}) como en la parte (I). Vamos a demostrar que $g = f = G$ casi en todo punto. Sea $t \in [a, b]$ tal que f es continua en t y sea $\varepsilon > 0$ arbitrario, entonces existe un $\delta > 0$ tal que

$$\text{si } |t - x| < \delta \text{ entonces se cumple que } |f(t) - f(x)| < \frac{\varepsilon}{2} \tag{16}$$

Ya que $|P_k| \longrightarrow 0$, existe un $k_1 \in \mathbb{N}$ tal que $|P_k| < \delta$ para toda $k \geqslant k_1$. Por otro lado, ya que $\lim_{k\to\infty} g_{P_k}(t) = g(t)$, esto quiere decir que existe un $k_2 \in \mathbb{N}$ tal que si $k \geqslant k_2$, entonces se cumple que

$$|g(t) - g_{P_k}(t)| < \frac{\varepsilon}{2}$$

Tomando $k_3 = \max(k_1, k_2)$ provocamos que $t \in (x_{i-1}, x_i) \subseteq (t - \delta, t + \delta)$ para algún intervalo (x_{i-1}, x_i) de la partición P_{k_3} y de (16) se sigue que

$$f(t) - \frac{\varepsilon}{2} < f(y) \text{ para todo } y \in (x_{i-1}, x_i).$$

Lo que implica que

$$f(t) - \frac{\varepsilon}{2} < m_i.$$

Así

$$\begin{aligned} |f(t) - g(t)| &\leqslant |f(t) - m_i| + |m_i - g(t)| \\ &= |f(t) - m_i| + \left| g_{P_{k_3}}(t) - g(t) \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Por lo que $g = f$ en cualquier punto donde la función f sea continua, es decir, $f = g$ casi en todo punto. De manera similar se demuestra lo propio con G y concluimos que $g = G$ casi en todo punto. Por último, para ver que f es Riemann integrable usaremos el criterio (II) del teorema 7. Gracias el teorema de la convergencia dominada, podemos decir que

$$0 \leqslant \lim_{k\to\infty} U(f, P_k) - L(f, P_k) = \lim_{k\to\infty} \int_{[a,b]} G_{P_k} d\lambda - \int_{[a,b]} g_{P_k} d\lambda = \int_{[a,b]} (G - g)\, d\lambda = 0$$

Por lo tanto, para cada $\varepsilon > 0$ podemos encontrar una partición P_k tal que $U(f, P_k) - L(f, P_k) < \varepsilon$. Así, f es Riemann integrable. ■

Este teorema no solo despeja la duda que planteamos casi al inicio de la sección de si existen funciones Riemann integrables que no sean Lebesgue integrables. Dejando claro que la integral de Lebsegue es más general que la integral de Riemann. Un hecho más impresionante es la caraterización que brinda este resultado de las funciones que son Riemann integrables en función de la medida del conjunto de puntos de discontinuidades de la función en cuestión. De cursos de cálculo integral, se sabe que una función continua es siempre Riemann integrbale. De hecho, para los alumnos de cálculo integral resulta fácil probar que las funciones con una cantidad finita de discontinuidades son también Riemann integrables. Pero de tener un númer finito de discontinuidades, se tiene luego a la función de Dirichlet, que no es Riemann integrable, pero que es discontinua en todos los elementos de su dominio. La pregunta de si existe un punto intermedio entre estas dos situaciones donde la función aún sea Riemann integrable, se da con este resultado.

Para terminar esta sección, se discutirá un poco de la integral impropia de Riemann. Aún cuando el teorema anterior garantiza que en intervalos acotados toda función Riemann integrable es Lebesgue integrable en cuyo caso ambas integrales coinciden, una cosa diferente ocurre cuando hablamos de la integral impropia de Riemann. Haciendo memoria, la integral impropia de Riemann es un intento por extender la integral de Riemann a funciones definidas en intervalos no acotados. Esto se hace mediante la idea de límites como se define en la sección 3.5.

Si se considera la función definida en $[0, \infty)$

$$g(x) = \begin{cases} 1 & \text{si } x = 0 \\ \frac{(-1)^{n+1}}{n} & \text{si } x \in (n-1, n] \text{ para } n = 1, 2, \ldots \end{cases}$$

Calcular la integral impropia de Riemann de esta función es de hecho equivalente a calcular el valor de la serie

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

la cual es una serie alternante que converge. Sin embargo, esta función no es Lebesgue integrable, pues al intentar calcular la integral de la parte positiva de g, uno se dá cuenta que es la suma de los términos impares de la serie geométrica que, como ya sabemos, es infinito. Algo similar ocurre con la parte negativa de g, resultando en que g no puede ser Lebesgue integrable.

4.7. Teorema fundamental del cálculo para la integral de Lebesgue

Los contenidos de esta sección están completamente inspirados en las ideas presentadas en [**14**].

El objetivo de esta sección es formular los resultados más representativos de la integral de Riemann, el primer y segundo teorema fundamental del cálculo, pero adecuados a la teoría de integración de Lebesgue. Estos dos resultados, el primer y segundo teorema fundamental del cálculo, nos dicen bajo qué condiciones la derivada y la integral se consideran procesos opuestos. Para dar una formulación lo más familiar posible, en este capítulo se usará la siguiente notación para referirnos a la integral de Lebesgue.

Definición 33. Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función Lebesgue integrable. Definimos la expresión

$$\int_c^d f(x) d\lambda := \int_{[c,d]} f d\lambda$$

para cualquier $c, d \in [a, b]$ con $c < d$.

***Observación* 5.** *No es difícil comprobar que si $h(x) = f(x + k)$, entonces*

$$\int_c^d h(x)\, d\lambda = \int_{c+k}^{d+k} f(x)\, d\lambda$$

siempre que $c < d$ y que $c, c + k, d, d + k \in [a, b]$.

Para poder expresar la relación que hay entre la derivada de una función y la integral de Lebesgue, resulta necesario hablar primero de las funciones crecientes y las funciones de variación acotada.

4.7.1. Funciones crecientes y funciones de variación acotada

Los siguientes dos resultados se dejan sin demostración ya que estas resultan demasiado técnicas e incluyen aspectos de la medida de Lebesgue que no van acordes con el objetivo de este trabajo. Sin embargo, una demostración de estos puede ser encontrada en [**14**].

Lema 4. ***(Lema de Vitali para conjuntos medibles)*** *Sea E un conjunto Lebesgue medible con medida finita. Si $\mathcal{H}$ es una colección de intervalos, ya sean abiertos, cerrados o semi abiertos, pero sin haber intervalos que consten de un único punto, que cumple que para cada $x \in E$ y para todo $\delta > 0$ existe un intervalo I en $\mathcal{H}$ tal que $x \in I$ y $\lambda(I) < \delta$, entonces para todo $\varepsilon > 0$ existe una colección finita de intervalos $I_1, I_2, \ldots, I_n$ de $\mathcal{H}$ tales que*

$$\lambda\left(E \setminus \bigcup_{j=1}^{n} I_j\right) < \varepsilon$$

Donde $\lambda(A \setminus B)$ se interpreta como $\lambda(A \cap B^c)$. Es decir, podemos encontrar una colección finita de intervalos de $\mathcal{H}$ que cubran a todo E, salvo quizá, por un conjunto de medida menor a ε. A la colección de intervalos $\mathcal{H}$ se le suele llamar cobertura de E en el sentido de Vitali.

Teorema 33. *Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función creciente. Entonces, f es diferenciable casi en todo punto.*

Ahora que se sabe que la derivada de una función creciente existe *c.t.p.*, nos falta comprobar que esta sea una función medible.

Teorema 34. *Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función creciente. Entonces, su derivada es una función medible y se cumple que*

$$\int_a^b f'(x)d\lambda \leqslant f(b) - f(a).$$

Demostración. Comencemos extendieno un poco el dominio de f definiendo $f(x) = f(b)$ para $b < x \leqslant b + 1$. Ahora, por el teorema anterior, sabemos que existe un $E \subseteq [a, b]$ tal que $\lambda(E) = 0$ y f es diferenciable en E^c. Para cada $x \in [a, b]$ definimos a

$$g(x) = \begin{cases} \lim\limits_{h \to 0} \dfrac{f(x+h) - f(x)}{h} & \text{si } x \in E^c \\ 0 & \text{si } x \in E \end{cases}$$

Entonces, g representa la derivada de f y la sucesión (g_n) dada por

$$g_n(x) = \begin{cases} \frac{f\left(x+\frac{1}{n}\right) - f(x)}{\frac{1}{n}} & \text{si } x \in E^c \\ 0 & \text{si } x \in E \end{cases}$$

claramente es una sucesión de funciones medibles que convergen puntualmente a g. Por el teorema 19, concluimos que g es medible.

Para demostrar la desigualda del teorema, comencemos observando que $g_n \geqslant 0$, pues f es creciente. Aplicando el lema de Fatou a la sucesión (g_n) tenemos que

$$\begin{aligned}
\int_a^b g(x)\, d\lambda &\leqslant \liminf_{n\in\mathbb{N}} \left(\int_a^b g_n(x)\, d\lambda \right) \\
&= \liminf_{n\in\mathbb{N}} \left(n \int_a^b f\left(x + \frac{1}{n} \right) d\lambda - n \int_a^b f(x) d\lambda \right) \\
&= \liminf_{n\in\mathbb{N}} \left(n \int_{a+\frac{1}{n}}^{b+\frac{1}{n}} f(x) d\lambda - n \int_a^b f(x) d\lambda \right) \\
&= \liminf_{n\in\mathbb{N}} \left(n \int_b^{b+\frac{1}{n}} f(x) d\lambda - n \int_a^{a+\frac{1}{n}} f(x) d\lambda \right) \\
&\leqslant \liminf_{n\in\mathbb{N}} \left(n \int_b^{b+\frac{1}{n}} f(x)\, d\lambda - n \int_a^{a+\frac{1}{n}} f(a)\, d\lambda \right) \\
&= \liminf \left(f(b) - f(a) \right) = f(b) - f(a)
\end{aligned}$$

■

A continuación, se define lo que significa que una función sea de variación acotada.

Definición 34. Sea $f : [a,b] \longrightarrow \mathbb{R}$ una función y $Q = \{x_0, x_1, \ldots, x_n\}$ una partición del intervalo $[a,b]$. Definimos

$$\begin{aligned}
p(Q) &= \sum_{i=1}^{n} \left[f(x_i) - f(x_{i-1}) \right]^+ \\
n(Q) &= \sum_{i=1}^{n} \left[f(x_i) - f(x_{i-1}) \right]^- \\
t(Q) &= p(Q) + n(Q) = \sum_{i=1}^{n} \left| f(x_i) - f(x_{i-1}) \right|
\end{aligned}$$

donde s^+ se interpreta como s si $s \geqslant 0$ y cero en caso contrario, s^- se interpreta como $-s$ si $s \leqslant 0$ y cero en caso contrario. Claramente $p - n = f(b) - f(a)$.

A continuación, definimos los números reales extendidos

$$\begin{aligned}
P &= \sup \{ p(Q) : Q \text{ es partición de } [a,b] \} \\
N &= \sup \{ n(Q) : Q \text{ es partición de } [a,b] \} \\
T &= \sup \{ t(Q) : Q \text{ es partición de } [a,b] \}
\end{aligned}$$

y las llamamos, variación positiva, variación negativa y variación total de la función f en el intervalo $[a,b]$, respectivamente. En ocasiones, usaremos la notación $P_a^b(f)$, $N_a^b(f)$ y $T_a^b(f)$ para que quede clara esta dependencia. Si se cumple que $T < \infty$, diremos que la función f es de variación acotada.

***Observación* 6.** *No es difícil ver que si $x \leqslant y$, entonces se cumple que $P_a^x(f) \leqslant P_a^y(f)$, $N_a^x(f) \leqslant N_a^y(f)$ y $T_a^x(f) \leqslant T_a^y(f)$.*

Respecto a las funciones de variación acotada, tenemos el siguiente resultado inmediato.

Lema 5. *Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función de variación acotada. Entonces,*

$$T_a^b(f) = P_a^b(f) + N_a^b(f)$$

y

$$P_a^b(f) - N_a^b(f) = f(b) - f(a).$$

Demostración. Es claro que $P, N \leqslant T < \infty$. Por el teorema 1, concluimos que

$$\begin{aligned} P_a^b + N_a^b &= \sup \{p(Q) : Q \text{ es partición de } [a, b]\} + \sup \{n(Q) : Q \text{ es partición de } [a, b]\} \\ &= \sup \{p(Q) + n(Q) : Q \text{ es partición de } [a, b]\} = T. \end{aligned}$$

La otra igualdad se obtiene de manera similar. ■

Es en el siguiente resultado donde se establece la relación que existe entre las funciones de variación acotadas y las funciones crecientes.

Teorema 35. *Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función. Entonces, f es de variación acotada si y solo si f es la diferencia de dos funciones crecientes.*

Demostración. Supongamos que f es de variación acotada. Entonces, definimos a $g(x) = P_a^x(f)$ y $h(x) = N_a^x(f)$. Por la observación 6, sabemos que tanto g como h son funciones crecientes. Por otro lado, gracias al lema 5, sabemos que $f(x) = g(x) - h(x) + f(a)$. De esta manera, hemos expresado a f como la diferencia de dos funciones crecientes, g y $h - f(a)$.

Supongamos ahora que $f = g - h$, donde g y h son dos funciones crecientes. Entonces, para cualquier partición Q del intervalo $[a, b]$, se sigue por la desigualdad del triángulo que

$$\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| = \sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| + \sum_{i=1}^{n} |h(x_i) - h(x_{i-1})|$$

y por ser g y h crecientes, tenemos que

$$\sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| + \sum_{i=1}^{n} |h(x_i) - h(x_{i-1})| = (g(b) - g(a)) + (h(b) - h(a))$$

de donde concluimos que $T \leqslant (g(b) + h(b)) - (g(a) - h(a))$. ■

Gracias a este resultado, tenemos el siguiente corolario para funciones de variación acotada.

Corolario 4. *Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función de variación acotada. Entonces, la derivada de f existe casi en todo punto.*

Demostración. Como acabamos de ver en el teorema anterior, f es la diferencia de dos funciones crecientes, digamos g y h. Por el teorema 34, sabemos que existen conjuntos medibles E y F con $\lambda(E) = \lambda(F) = 0$ tales que g es diferenciable en E^c y h es diferenciable en F^c. Tomando $O = E \cup F$, tenemos que $f' = g' - h'$ en O^c y además $\lambda(O) = \lambda(E \cup F) = 0$. ■

Esta es toda la teoría necesaria de funciones de variación acotada que se verá en este trabajo.

4.7.2. Primer teorema fundamental del cálculo para la integral de Lebesgue

El lema que viene a continuación, tiene su análogo en la teoría de integración de Riemann.

Lema 6. *Sea $f : [a,b] \longrightarrow \mathbb{R}$ una función Lebesgue integrable. Entonces, la función*

$$F(x) = \int_a^x f(t) d\lambda$$

es una función continua y de varación acotada en $[a,b]$.

Demostración. La continuidad de f es prácticamente inmediata, pues para cualquier $\varepsilon > 0$, por el teorema 27, existe un $\delta > 0$ tal que si $y \in (x-h, x+h)$ y $\lambda((x-h, x+h)) < \delta$, entonces se cumple que

$$|F(x) - F(y)| = \left| \int_{[\text{mín}(x,y), \text{máx}(x,y)]} f d\lambda \right| < \varepsilon$$

sin importar el $x \in [a,b]$ elegido.

Para ver que f es de variación acotada, tomemos $Q = \{x_0, x_1, \ldots, x_n\}$ una partición cualquiera de $[a,b]$. Entonces,

$$\sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| = \sum_{i=1}^{n} \left| \int_{x_{i-1}}^{x_i} f(t) d\lambda \right| \leqslant \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |f(t)|\, d\lambda = \int_a^b |f(t)|\, d\lambda.$$

Por el teorema 26, sabemos que este es un valor finito. Como Q fue un partición arbitraria de $[a,b]$, concluimos que $T_a^b(F) \leqslant \int_a^b |f(t)|\, d\lambda$. ■

Lema 7. *Sea $f : [a,b] \longrightarrow \mathbb{R}$ una función Lebesgue integrable. Si*

$$\int_a^x f(t) d\lambda = 0$$

para cada $x \in [a,b]$, entonces f es cero casi en todo punto.

Demostración. Supongamos que no es así y que el conjunto

$$E = \{x \in [a,b] : f(x) > 0\}$$

tiene medida $\varepsilon > 0$. Entonces, por el teorema 17, sabemos que existe un conjunto cerrado $G \subseteq E$ tal que $\lambda(E \cap G^c) < \frac{\varepsilon}{2}$. Esto implica que

$$\lambda(G) + \lambda(E \cap G^c) = \lambda((E \cap G) \cup (E \cap G^c)) > \varepsilon.$$

De donde $\lambda(G) > \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}$. Sea O el conjunto abierto dado por $O = (a,b) \cap G^c$. Luego,

$$0 = \int_a^b f(t) d\lambda = \int_O f d\lambda + \int_G f d\lambda.$$

y por lo tanto

$$\int_O f d\lambda = -\int_G f d\lambda \neq 0.$$

Pero al ser O abierto, es la unión numerable de intervalos abiertos disjuntos $((c_n, d_n))_{n=1}^{\infty}$. Por el corolario 3, concluimos que

$$\int_O f d\lambda = \sum_{n=1}^{\infty} \int_{c_n}^{d_n} f(t) d\lambda.$$

y se sigue que

$$\int_{c_n}^{d_n} f(t) d\lambda \neq 0$$

para algún $n \in \mathbb{N}$. Lo cual es una contradicción, pues

$$\int_{c_n}^{d_n} f(t) d\lambda = \int_a^{d_n} f(t) d\lambda - \int_a^{c_n} f(t) d\lambda = 0$$

Se llega a una contradicción de manera similar si se asume que f es menor que cero para un cojunto de medida no nula. ■

A continuación, se presenta el último lema antes de ver el primer teorema fundamental del cálculo para la integral de Lebesgue, que básicamente, es el primer teorema fundamental del cálculo para funciones acotadas. Notemos que en las hipótesis no decimos nada a cerca de la integrabilidad de la función f, debido a que esta se sigue por ser f acotada.

Lema 8. *Sea $f : [a, b] \longrightarrow \mathbb{R}$ acotada y medible. Entonces, la función*

$$F(x) = \int_a^x f(t) d\lambda$$

es diferenciable casi en todo punto y además, $F'(x) = f(x)$ casi en todo punto.

Demostración. Por el lema 6 y el corolario 4, concluimos que F es diferenciable casi en todo punto. Ahora, sea $K \in \mathbb{N}$ tal que $|f| < K$. Para cada $n \in \mathbb{N}$ definimos

$$g_n(x) = \frac{F\left(x + \frac{1}{n}\right) - F(x)}{\frac{1}{n}}.$$

Entonces, para cada $x \in [a, b]$ se tiene que

$$|g_n(x)| = \left| n \int_x^{x+\frac{1}{n}} f(t) d\lambda \right| \leqslant n \int_x^{x+\frac{1}{n}} K d\lambda = K.$$

Por lo tano, la sucesión (g_n) está dominada y además, converge casi en todo punto a F'. Por el teorema de la convergencia dominada, se sigue que, para cada $c \in [a,b]$

$$\begin{aligned}
\int_a^c F'(t)d\lambda &= \lim_{n\to\infty}\int_a^c g_n d\lambda \\
&= \lim_{n\to\infty}\left(n\int_a^c F\left(t+\frac{1}{n}\right)d\lambda - n\int_a^c F(t)d\lambda\right) \\
&= \lim_{n\to\infty}\left(n\int_{a+\frac{1}{n}}^{c+\frac{1}{n}} F(t)d\lambda - n\int_a^c F(t)d\lambda\right) \\
&= \lim_{n\to\infty}\left(n\int_{c}^{c+\frac{1}{n}} F(t)d\lambda - n\int_a^{a+\frac{1}{n}} F(t)d\lambda\right) \\
&= \lim_{h\to 0^+}\left(\frac{1}{h}\int_{c}^{c+h} F(t)d\lambda - \frac{1}{h}\int_a^{a+h} F(t)d\lambda\right) \\
&= \lim_{h\to 0^+}\frac{I(c+h)-I(c)}{h} - \lim_{h\to 0^+}\frac{I(a+h)-I(a)}{h}
\end{aligned}$$

Donde $I(x) = \int_a^x F(x)d\lambda$. Además, dado que F es continua, entonces es Reimann integrable y podemos aplicar el primer teorema fundamental del cálculo para decir que I es diferenciable en a y en b y finalmente concluir que

$$\int_a^c F'(t)d\lambda = I'(c) - I'(a) = F(c) - F(a) = F(c) = \int_a^c f(t)d\lambda.$$

Por lo tanto,

$$\int_a^c (F' - f)(t)\, d\lambda = 0$$

para cada $c \in [a,b]$. Por el lema 7, concluimos que $F' = f$ casi en todo punto. ■

Finalmente, se puede presentar una demostración del primer teorema fundamental del cálculo adecuado a la integral de Lebesgue.

Teorema 36. ***(Primer teorema fundamental del cálculo para la integral de Lebesgue).*** *Sea $f : [a,b] \longrightarrow \mathbb{R}$ Lebesgue integrable. Entonces, la función*

$$F(x) = \int_a^x f(t)d\lambda$$

es diferenciable casi en todo punto y además, $F'(x) = f(x)$ casi en todo punto.

Demostración. Nuevamente, la diferenciabilidad de F casi en todo punto, se sigue del corolario 4 y el lema 6. Demostremos que el enunciado es cierto cuando $f \geqslant 0$. Empecemos definiendo

$$f_n(x) = \begin{cases} f(x) & \text{si } x \leqslant n \\ n & \text{si } x > n \end{cases}$$

Entonces, se cumple que $f - f_n \geqslant 0$ y por lo tanto G_n definida como

$$G_n(x) = \int_a^x (f - f_n)(t)\, d\lambda$$

es una función creciente que además, por el teorema 34, sabemos que tiene derivada casi en todo punto y por ser creciente, esta derivada sera no negativa cuando exista. Ahora, por el lema 8 aplicado a f_n tenemos que $F_n' = f_n$ casi en todo punto, donde

$$F_n(x) = \int_a^x f_n(t)d\lambda$$

y se sigue que si G_n y F_n son diferenciables en x, entonces

$$G_n^{'}(x) + F_n^{'}(x) = (G_n + F_n)^{'}(x) = F'(x)$$

y por lo tanto

$$F'(x) \geqslant F_n^{'}(x)$$

De donde $F' \geqslant f_n$ casi en todo punto. Ya que n fue arbitrario, podemos concluir que

$$F' \geqslant f \tag{17}$$

casi en todo punto. De donde

$$\int_a^b F'(t)\, d\lambda \geqslant \int_a^b f(t)\, d\lambda = F(b) = F(b) - F(a).$$

Por el teorema 34, sabemos que

$$\int_a^b F'(t)\, d\lambda \leqslant F(b) - F(a)$$

y por lo tanto, concluimos que

$$\int_a^b (F' - f)\, d\lambda = 0.$$

Pero acabamos de demostrar en (17) que $F' - f \geqslant 0$ casi en todo punto. Esto implica que $F' - f = 0$ casi en todo punto.

No hay que olvidar, que al inicio supusimos $f \geqslant 0$. El caso general se sigue de observar que $f = f^+ - f^-$, donde $f^+, f^- \geqslant 0$ y por lo tanto, si definimos a

$$F^+(x) = \int_a^x f^+(t) d\lambda$$
$$F^-(x) = \int_a^x f^-(t) d\lambda.$$

entonces podemos aplicar lo recien demostrado y concluir que existe un conjunto medible O tal que $\lambda(O) = 0$ y $(F^+)^{'} = f^+$ y $(F^-)^{'} = f^-$ en O^c. Tomemos un $x \in O$, entonces se cumple

$$f(x) = f^+(x) - f^-(x) = \left(F^+\right)^{'}(x) - \left(F^-\right)^{'}(x) = \left(F^+ - F^-\right)^{'}(x) = F'(x).$$

Por lo tanto, $F' = f$ casi en todo punto. ∎

Por el primer teorema fundamental del cáulculo, se sabe que esta relación entre la derivada de la integral de Riemann y el integrando se da siempre y cuando el integrando sea una función continua. Pero, como ya se ha visto en el capítulo de comparación entre la integral de Riemann y la integral de Lebesgue, si una función es Riemann integrbale, entonces es Lebesgue integrable y ambas integrales coinciden, por lo que, desde un inicio el teorema fundamental del cálculo ya valía para la integral de Lebesgue. Es decir, si $f : [a, b] \longrightarrow \mathbb{R}$ es Riemann integrable y continua en $c \in [a, b]$, entonces

$$F(x) = \int_{[a,x]} f d\lambda$$

es una función continua y diferenciable en c con derivada $F'(c) = f(c)$.

Lo que se acaba de demostrar ahora con este llamado primer teorema fundamental del cálculo para la integral de Lebesgue es, para empezar, que la derivada de la función definida por la integral de Lebesgue, llamada F en el teorema, existe casi en todo punto. Pero además, se demostró que esta relación entre la derivada de la función F y el integrando f se cumple casi en todo punto, aún cuando el integrando no sea una función continua.

4.7.3. Segundo teorema fundamental del cálculo para la integral de Lebesgue

Antes de intentar formular el segundo teorema fundamental del cálculo adecuado a la integral de Lebesgue, se debe de estudiar un poco un concepto intimamente relacionado con las funciones definidas por integrales, la continuidad absoluta.

Definición 35. Sea $f : [a,b] \longrightarrow \mathbb{R}$ una función. Diremos que f es absolutamente continua si para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que para cualquier colección finita de intervalos ajenos $\left\{ \left(x_1^-, x_1^+\right), \left(x_2^-, x_2^+\right), \ldots, \left(x_n^-, x_n^+\right) \right\}$ con

$$\sum_{i=1}^{n} x_i^+ - x_i^- < \delta$$

se cumple que

$$\sum_{i=1}^{n} \left| f\left(x_i^+\right) - f\left(x_i^-\right) \right| < \varepsilon$$

Es inmediato ver que si una función es absolutamente continua, entonces es continua y que la combinación lineal de un número finito de funciones absolutamente continuas es absolutamente continua. Aún más, el siguiente lema afirma que este tipo de funciones también son de variación acotada.

Lema 9. *Sea $f : [a,b] \longrightarrow \mathbb{R}$ una función absolutamente continua, entonces f es de variación acotada.*

Demostración. Sea $\varepsilon = 1$, entonces existe un $\delta > 0$ como en la definición de continuidad absoluta. A continuación, consideremos la colección finita de intervalos $\left\{ \left(a, a + \frac{\delta}{2}\right), \ldots, \left(a + \frac{n\delta}{2}, a + \frac{(n+1)\delta}{2}\right), \ldots, \left(a + \frac{K\delta}{2}, b\right) \right\}$, donde K es el máximo entero positivo para el cual $a + \frac{K\delta}{2} < b$. Tomemos una partición $Q = \{x_0, \ldots, x_m\}$ y definimos los conjuntos

$$A_1 = \left(a, a + \frac{\delta}{2}\right) \bigcap \left(\bigcup_{i=1}^{m} (x_{i-1}, x_i) \right)$$

$$\vdots$$

$$A_n = \left(a + \frac{(n-1)\delta}{2}, a + \frac{n\delta}{2}\right) \bigcap \left(\bigcup_{i=1}^{m} (x_{i-1}, x_i) \right)$$

$$\vdots$$

$$A_K = \left(a + \frac{K\delta}{2}, b\right) \bigcap \left(\bigcup_{i=1}^{m} (x_{i-1}, x_i) \right)$$

No es difícil ver que cada A_j es la unión finita de intervalos abiertos. Llamémosle $B_j = \left\{(r_1, s_1), (r_2, s_2), \ldots, \left(r_{k_j}, s_{k_j}\right)\right\}$ a la colección ordenada de intervalos abiertos que conforman a A_j. Entonces, en cada B_j se cumple que

$$\sum_{(r_i,s_i)\in B_j} s_i - r_i < \delta.$$

Por último, veamos que

$$\begin{aligned}\sum_{j=1}^{m} |f(x_j) - f(x_{j-1})| &\leqslant \sum_{j=1}^{K}\left(\sum_{(r_i,s_i)\in B_j} |f(s_i) - f(r_i)|\right)\\ &< \sum_{j=1}^{K} 1 = K.\end{aligned}$$

Ya que Q fue arbitraria, concluimos que $T_a^b(f) \leqslant K$. ■

Tenemos el siguiente corolario inmediato.

Corolario 5. *Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función absolutamente continua, entonces f tiene derivada casi en todo punto.*

Demostración. Se sigue del teorema anterior y el corolario 4 ■

Algo interesante que se puede notar de la teoría desarrollada en esta y la sección anterior, es el hecho de que si las funciones están definidas en un intervalo de la forma $[a, b]$, entonces se tienen las siguientes implicaciones: Continuidad absoluta implica de variación acotada implica con derivada casi en todo punto.

El siguiente lema asegura que, si la derivada de una función absolutamente continua es cero casi en todo punto, entonces la función debe de ser constante. Cuestión que no se puede asegurar para funciones arbitrarias. Es decir, una función puede tener derivada igual a cero casi en todo punto, pero no ser constante. Un ejemplo muy sencillo de esto es la función

$$f(x) = \begin{cases} 1 & \text{si } x \in [0,1] \\ 2 & \text{si } x \in (1,2] \end{cases}$$

Tal vez se podría pensar que el ejemplo anterior falla únicamente debido a la discontinuidad de la función, pero esto no es así. Existen ejemplos de funciones que son uniformemente continuas (y por lo tanto, continuas) cuya derivada es cero casi en todo punto, pero que no son constantes. El ejemplo más famoso de una función con estas características, es la función de Cantor (vease **[10]**), que tiene derivada igual a cero en el complemento del conjunto de Cantor (y por lo tanto, tiene derivada cero casi en todo punto), pero la imagen de dicha función son todos los números entre cero y uno.

Lema 10. *Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función absolutamente continua. Si $f' = 0$ casi en todo punto, entonces f es constante.*

Demostración. Vamos a demostrar que $f(a) = f(c)$ para todo $c \in (a, b]$. Ya que $f' = 0$ *c.t.p.*, sabemos que existe un conjunto $E \subseteq (a, c)$ tal que $f' = 0$ en E y $\lambda(E) = c - a$. Tomemos ahora ε y τ números positivos arbitrarios. Para cada $x \in E$, existe un intervalo $[x, x + h]$ tal que

$$\left| \frac{f(x+h) - f(x)}{h} \right| < \tau.$$

Equivalentemente, $|f(x+h) - f(x)| < h\tau$. Notemos que la colección de intervalos $[x, x+h]$ forman una cobertura en el sentido de Vitali de E y por lo tanto, por el lema 4, podemos encontrar una subcolección finita de intervalos de este tipo $\{[x_n, y_n]\}_{n=1}^{k}$ no superpuestos que cubren a E, excepto tal vez, por un conjunto A de medida tan pequeña como se quiera. Nosotros tomaremos A de tal manera que $\lambda(A) < \delta$, donde δ es el número correspondiente al ε en la definición de continuidad absoluta. A continuación, supongamos que los intervalos de la colección $\{[x_n, y_n]\}_{n=1}^{k}$, están ordenados como se muestra en la figura 11, es decir, con $x_n \leqslant x_{n+1}$ y renombremos a como y_0 y c como x_{k+1} para que las operaciones sean más compactas. Entonces, tenemos que

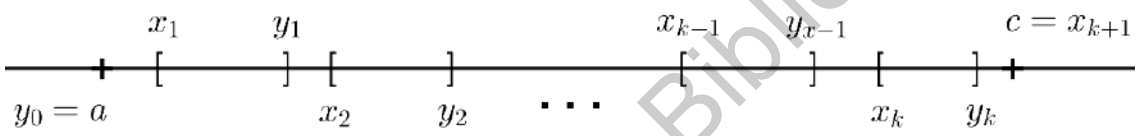

Figura 11: Subcolección de la cubierta de Vitali de $[a, c]$

$$\sum_{n=1}^{k+1} x_n - y_{n-1} < \delta$$

y por continuidad aboluta se sigue que

$$\sum_{n=1}^{k+1} |f(x_n) - f(y_{n-1})| < \varepsilon.$$

Por otro lado, para cada $n = 1, \ldots, k$ se cumple que

$$|f(y_n) - f(x_n)| < \tau(y_n - x_n)$$

y por lo tanto

$$\sum_{n=1}^{k} |f(y_n) - f(x_n)| < \tau(c - a).$$

De donde concluimos que

$$\begin{aligned} |f(a) - f(c)| &= |f(y_0) - f(x_{k+1})| \\ &\leqslant |f(y_0) - f(x_1)| + |f(x_1) - f(y_1)| + \ldots + |f(y_k) - f(x_{k+1})| \\ &< \varepsilon + \tau(c - a). \end{aligned}$$

Ya que ε y τ fueron números positivos arbitrarios, concluimos que $f(a) = f(c)$. ■

Ahora, se presenta el resultado principal de esta sección, el segundo teorema fundamental del cálculo adecuado a la integral de Lebesgue.

Teorema 37. *Una función $F : [a,b] \longrightarrow \mathbb{R}$ es de la forma $F(x) = \int_a^x F'(t)\, d\lambda + F(a)$ si y solo si es absolutamente continua.*

Demostración. La necesidad se sigue inmediatamente del teorema 27. Para la segunda implicación, supongamos que F es absolutamente continua. Entonces, F es de variación acotada y podemos escribir a F como

$$F = F_1 - F_2$$

donde F_1 y F_2 son ambas funciones crecientes. Además, tanto F como F_1 y F_2 son diferenciables casi en todo punto y se cumple que

$$|F'| = |F_1' - F_2'| \leqslant |F_1'| + |F_2'| = F_1' + F_2'$$

casi en todo punto. Por lo tanto,

$$\int |F'|\, d\lambda \leqslant \int (F_1' + F_2')\, d\lambda.$$

Y por el teorema 34 podemos decir que

$$\int |F'|\, d\lambda \leqslant F_1(b) + F_2(b) - F_1(a) - F_2(a).$$

Lo que, junto con el teorema 26, nos lleva a concluir que F' es Lebesgue inetgrable. Llamemos

$$G(x) = \int_a^x F'(t) d\lambda.$$

Entonces, por la primera parte de la demostración, G es absolutamente continua y por lo tanto $F - G$ también lo es. Por el primer teorema fundamental del cálculo para la integral de Lebesgue, se sigue que $0 = F' - G' = (F - G)'$ casi en todo punto y finalmente, gracias al lema 10 concluimos que $F - G$ es constante y

$$F(x) = G(x) + c = \int_a^x F'(t)\, d\lambda + c \text{ para algún } c \in \mathbb{R}.$$

De hecho, es fácil ver que $c = F(a)$, por lo que

$$F(x) = \int_a^x F'(t)\, d\lambda + F(a)$$

■

Para dar una mejor comparación entre este teorema y su análogo con la integral de Riemann, se dará el siguiente corolario y además, se enunciará el segundo teorema fundamental del cálculo para la integral de Riemann.

Corolario 6. ***(Segundo teorema fundamental del cálculo para la integral de Lebesgue)*** *Si $f : [a,b] \longrightarrow \mathbb{R}$ es una función tal que $f = g'$ casi en todo punto para alguna función absolutamente continua g, entonces f es medible, integrable y*

$$\int_a^x f(t) d\lambda = g(x) - g(a).$$

Demostración. Aplicando el teorema 37 a g, tenemos que

$$\int_a^x g'(t)d\lambda = g(x) - g(a).$$

Pero $f = g'$ casi en todo punto, por lo que

$$\int_a^x f(t)d\lambda = \int_a^x g'(t)d\lambda.$$

$\blacksquare$

Teorema 38. ***(Segundo teorema fundamental del cálculo)*** *Si $f : [a, b] \longrightarrow \mathbb{R}$ es una función Riemann integrable tal que $f = g'$ para alguna función g, entonces*

$$\int_a^x f(t)dt = g(x) - g(a).$$

Como se puede apreciar, se han incluido en el corolario 6 dos funciones f y g para que se parezca más a su análogo de la intergal de Riemann, pues en el caso de la teoría de integración de Riemann, la intención de este teorema al incluir dos funciones es establecer un método para hallar la integral de una función en términos de antiderivadas, más que indicar bajo qué condiciones la integral es el proceso inverso a la derivada. Sin embargo, no es difícil ver que ambos resultados, el teorema fundamental del cálculo para la integral de Lebesue y Riemann, se pueden abordar desde el punto de vista de qué condiciones son las necesarias para garantizar que la integral es el proceso inverso de la derivada. Como era de esperarse, la teoría de integración de Lebesgue impone condiciones menos restrictivas para garantizar esto. Solamente se necesita que la función g sea absolutamente continua. Mientras que en la teoría de integración de Riemann, no solo tenemos que asegurar la diferenciabilidad de g en todo $[a, b]$, si no que también la integrabiliadad de la función g'.

5. Lema de Cousin

En este breve capítulo se dará una demostración diferenta a la presentada en la sección 3.4 del lema de cousin. El siguiente lema, para ser entendido en su totalidad, hace falta conocer el concepto de conjunto cerrado, por lo que aquellos que no estén familiarizados con este término pueden consultar el capítulo 11 de [**1**].

Lema 11. *Sea (E_n) una sucesión de conjuntos cerrados , acotados y no vacíos de $\mathbb{R}$. Supongamos que la sucesión es decreciente, es decir, $E_{n+1} \subseteq E_n$ para todo $n \in \mathbb{N}$. Entonces, la intersección de todos ellos*

$$\bigcap_{n=1}^{\infty} E_n$$

es no vacía.

Demostración. Para cada $n \in \mathbb{N}$ definamos a $s_n = \sup\{x : x \in E_n\}$. Claramente la sucesión de números (s_n) es decreciente, pues $E_{n+1} \subseteq E_n$, y además está acotada por $\inf\{x : x \in E_1\}$. Por lo que se trata de una sucesión convergente, digamos que converge a un número s. Una cosa importante que habremos de resaltar, es el hecho de que cada s_n pertenece al E_n correspondiente, pues al tratarse del supremo, inmediatamente esto lo convierte en un punto de acumulación del conjunto y por ser E_n cerrado, este contiene a todos sus puntos de acumulación.

Si logramos demostrar que s pertenece a todos los conjuntos E_n habremos terminado, pues esto significará que s se encuentra en la intersección de todos ellos. Para ello tomemos un $N \in \mathbb{N}$ arbitrario y simplemente notemos que s es un punto de acumulación de E_N, pues dado un $\varepsilon > 0$, existe un $M \in \mathbb{N}$ tal que

$$|s_m - s| < \varepsilon \text{ para cualquier } m \geqslant M. \tag{18}$$

De manera que si n es mayor que M y N, entonces s_n cumple (18) y además $s_n \in E_n \subseteq E_N$, convirtiendo a s en un punto de acumulación de E_N y terminando la demostración.

■

El lema de Cousin involucra el concepto de particiones etiquetadas para una mayor comodidad en la demostración y en el planteamiento del resultado. Este concepto se define en la sección 3.4, pero para ser más prácticos lo enunciaremos a continuación de igual manera.

Definición 36. Sea $\{x_0, x_1, \ldots, x_k\}$ una partición del intervalo $[a, b]$ y t_i un punto en el intervalo $[x_{i-1}, x_i]$ para cada $i = 1, 2, \ldots, k$. Vamos a llamar partición etiquetada de $[a, b]$ al conjunto de parejas

$$P = \{([x_0, x_1], t_1), ([x_1, x_2], t_2), \ldots, ([x_{k-1}, x_k], t_k)\}$$

Lema de Cousin: *Si $\delta : [a, b] \longrightarrow \mathbb{R}$ es una función estríctamente positiva, entonces siempre existe una partición etiquetada*

$$P = \{([x_0, x_1], t_1), \ldots, ([x_{k-1}, x_k], t_k)\}$$

de $[a, b]$ *tal que*

$$x_i - x_{i-1} < \delta(t_i) \text{ para cada } i = 1, \ldots, k$$

Se le denomina a P *como partición* $\delta - fina$ *de* $[a, b]$.

Demostración. Demostrémoslo por contradicción. Supongamos que existe una función δ que sea estríctamente positiva, pero que no admita una partición etiquetada que sea $\delta - fina$. Supongamos, sin perdida de generalidad que δ está definida en el intervalo $[0, 1]$. Lo siguiente será notar que, sin importar cual sea la partición $P = \{x_0, x_1, \ldots, x_n\}$, siempre va a existir algún subintervalo inducido por la partición con la propiedad de que

$$\delta(t) \leqslant x_i - x_{i-1} \text{ para todo } t \in [x_{i-1}, x_i] \tag{19}$$

De lo contrario, para cada intervalo podríamos encontrar un $t_i \in [x_{i-1}, x_i]$ tal que $x_i - x_{i-1} < \delta(t_i)$, por lo que P junto con los t_i formarían una partición $\delta - fina$ de $[0, 1]$. Al intervalo que cumpla con (19) diremos que es un intervalo con la propiedad D.

Lo siguiente será definir la sucesión de conjuntos (E_n) empezando por

$$E_1 = [0, 1].$$

Ya que $P_1 = \{0, 1\}$ representa una partición de $[0, 1]$, entonces $\delta(t) \leqslant 1$ para todo $t \in E_1$, de lo contrario, P_1 sería una partición $\delta - fina$. Para definir a E_2 partimos el intervalo $[0, 1]$ en dos intervalos de igual longitud $[0, \frac{1}{2}]$ y $[\frac{1}{2}, 1]$. Notemos que estos dos intervalos inducen una partición $P_2 = \left\{0, \frac{1}{2}, 1\right\}$ de $[0, 1]$ y por lo tanto, debe de existir al menos uno de ellos que tenga la propiedad D. Definamos

$$\begin{aligned} E_2 = &\text{ La unión de los intervalos inducidos por la partición } P_2 \\ &\text{ que cumplen la propiedad } D. \end{aligned}$$

De manera similar que con el paso anterior, hay que darnos cuenta de que $\delta(t) \leqslant \frac{1}{2}$ para todo $t \in E_2$, pues este conjunto está formado por intervalos que cumplen con la propiedad D para la partición P_2. Para el siguiente paso, no dividiremos todos los subintervalos inducidos por P_2, si no que únicamente dividiermos aquellos intervalos que se hallen en E_2. Es decir, si en el paso anterior el único intervalo que cumplió con la propiedad D fue el intervalo $[0, \frac{1}{2}]$, entonces en este paso vamos a dividir a la mitad ese intervalo e inducir una partición P_3 generada por los intervalos $[0, \frac{1}{4}]$, $[\frac{1}{4}, \frac{1}{2}]$ y $[\frac{1}{2}, 1]$. Aplicando el mismo razonamiento que en el paso anterior, tenemos que debe de existir almenos uno de estos intervalos que cumpla con la propiedad D. Sin embrago, habrá que notar que los únicos intervalos inducidos por P_3 que pueden cumplir con la propiedad D, son los que acabamos de "partir". Es decir, solo los intervalos que se encuentren contenidos en E_2 pueden cumplir con la propiedad D, pues así es com definimos a E_2. En el ejemplo que planteamos hace un momento, únicamente $[0, \frac{1}{4}]$ y $[\frac{1}{4}, \frac{1}{2}]$ podrían cumplir la propiedad D. Entendiendo la idea, definir a E_3 no es un problema

$$\begin{aligned} E_3 = &\text{ La unión de los intervalos inducidos por la partición } P_3 \\ &\text{ que cumplen la propiedad } D. \end{aligned}$$

De este modo, $E_3 \subseteq E_2 \subseteq E_1$ y además $\delta(t) \leqslant \frac{1}{4}$ para todo $t \in E_3$, pues nuevamente, este se encuentra formado por intervalos inducidos por P_3 con la propiedad D . Continuando con este proceso de dividir en cada paso únicamente los intervalos que en el paso anterior hayan cumplido la propiedad D, generaríamos una sucesión decreciente de conjuntos cerrados (E_n), pues cada uno de ellos es la unión finita de intervalos cerrados, no vacíos y acotados. Por el lema 11, concluimos que la intersección de todos ellos es no vacía. Sea z que pertenece a la intersección de todos los E_n, esto quiere decir que $z \in E_n$ para toda $n \in \mathbb{N}$ y por lo tanto

$$\delta(z) \leqslant \frac{1}{2^{n-1}}$$

para todo n. De donde $\delta(z) = 0$, lo cual es una contradicción, pues quedamos que δ es estríctamente positiva.

■

6. Conclusiones

Uno de los objetivos de este trabajo era hacer un acercamiento para alumnos de primeros semestres de la carrera lo más amigable y accesible posible a los resultados principales de la integral de Lebesgue adecuados para la integral de Riemann. Como es de esperarse, no es posible trasladar de manera inmediata este tipo de resultados inherentes a la integral de Lebesgue, que expresan las condiciones necesarias para garantizar la validez del intercambio entre límite e integral, a la teoría de integración de Riemann. Sin embargo, es posible, mediante una modificación en las hipótesis, brindar resultados que contengan la esencia de sus homólogos en la teoría de integración de Lebesgue. Si bien, la sección 3.4, que es donde exponemos los teoremas de intercambio entre límite e integral de Riemann, contiene un lema un poco técnico, este se encuentra al alcance de un alumno de cálculo integral, dejando el contenido del resto del capítulo mucho más accesible.

El segundo objetivo de este trabajo era proporcionar contenidos que usualmente no son abarcados en un primer curso de teoría de la medida para aquellas personas deseosas de expandir sus conocimientos de la integral de Lebesgue. De este modo, en el capítulo 4, mostramos los análogos al primer y segundo teorema fundamental del cálculo adecuados para la integral de Lebesgue y no solo eso, sino que, además, como parte de esta "comparación" entre ambas integrales, expusimos un resultado que ligaba de forma clara y precisa la integral de Riemann con la integral de Lebesgue y la medida de Lebesgue. Dicho resultado brinda una de las caracterizaciones más interesantes de las funciones Riemann integrables en función de la medida del conjunto de puntos de discontinuidades y expande un poco más la visión general que se tiene de la teoría de integración.

Referencias

[1] R. G. Bartle. *The elements of real analysis*. Wiley, 2da edición, 1976.

[2] R. G. Bartle. *The elements of integration and Lebesgue measure*. Wiley, 1995.

[3] F. Cunningham Jr. *Taking Limits under the Integral Sign*. Mathematics Magazine, 179-184, 1967.

[4] G. Darboux. *Addition au Mémoire sur les fonctions discontinues*. Annales de l'Ecole Normale, 195-202, 1879.

[5] R. A. Gordon. *A Convergence Theorem for the Riemann Integral*. Mathematics Magazine, 141-147. 2000.

[6] E. T. Guarneros. *El lema de Cousin: Aplicaciones al análisis real*. Benemérita Universidad Autónoma de Puebla, 2-3, 2012.

[7] H. Kastelman. *Riemann Integration of Limit Functions*. The American Mathematical Monthly, 182-187, 1970.

[8] J. W. Lewin. *A truly elementary approach to the bounded convergence theorem*. The American Mathematical Monthly 93, 395-397. 1986

[9] W. A. J. Luxemburg. *Arzela's Dominated Convergence Theorem for the Riemann Integral*. The American Mathematical Monthly , 970-979, 1971.

[10] M. E. C. Montoya. *El conjunto de Cantor y algunas de sus propiedades*. Universidad de Sonora, 2003.

[11] J. V. Morales. *Introducción a la medida e integración*. Universidad Autónoma de Aguascalientes, 2005.

[12] W. F. Osgood. *A geometrical method for the treatment of uniform convergence and certain double limits*. Bulletin of the american mathematical society. 3, 59-86, 1896.

[13] B. Riemann. *Ueber die Darstellbarkeit einer Function durch eine trigonometrische Reihe*. In der Dieterichschen Buchhandlung, 1867.

[14] H. L. Royden. *Real analysis*. Macmillan publishing company, 3ª edición, 1988.

[15] W. Rudin. *Principles of mathematical analysis*. McGraw Hill, 3ª edición, 1980.

[16] M. Spivak. *Calculus* . Reverté, 3ª Ed, 2012.

[17] B. S. Thomson. *Monotone Convergence Theorem for the Riemann Integral*. The American Mathematical Monthly, 547-550, 2010.