Tema 8

Espacios de Lebesgue

Vamos a estudiar una amplia gama de espacios de Banach, directamente relacionados con la teoría de la integración. Para ello, empezamos probando las versiones para integrales de dos resultados clásicos con gran utilidad: las *desigualdades de Hölder y Minkowski*. Para cada valor de un parámetro $p \in \mathbb{R}$ con $p \geqslant 1$, la segunda de estas desigualdades permite obtener una seminorma en el espacio vectorial de las funciones que llamaremos *p-integrables*. Mediante un paso a cociente que consiste en identificar dos funciones que coincidan casi por doquier, obtendremos los espacios normados que nos interesan. La complitud de todos estos espacios se conoce como *teorema de Riesz-Fischer*, y en el caso $p = 1$ puede considerarse como un nuevo teorema de convergencia.

Estudiaremos algunos subespacios destacados de los espacios de Lebesgue, que permiten aproximar cada función integrable, o p-integrable con $p > 1$, mediante funciones del mismo tipo, que sean más sencillas, en algún sentido. En términos de clases de equivalencia, lo que hacemos es obtener varios subespacios densos en los espacios de Lebesgue, por lo que este tipo de resultados se conocen como *teoremas de densidad*.

El primero de estos teoremas utiliza las que llamaremos *funciones simples*, porque tienen imagen finita, pero pueden tomar valores negativos. En el contexto de los espacios de Lebesgue, este resultado se puede considerar como la versión definitiva del teorema de aproximación de Lebesgue. Deducimos otros dos teoremas de densidad, que son los más útiles en la práctica. Uno, porque aprovecha el tipo más sencillo de funciones simples, las *funciones escalonadas*. El otro, porque usa funciones más regulares, las *funciones continuas de soporte compacto*.

La integral de una función continua de soporte compacto puede definirse elementalmente como hizo Cauchy, sin usar ninguna medida. Su definición de integral fue evolucionando a lo largo del siglo XIX, hasta culminar en la que se conoce como *integral de Riemann*. Probaremos que esta integral no es más que un caso muy particular de la de Lebesgue, al mismo tiempo que obtenemos una caracterización, curiosamente debida al propio Lebesgue, de las funciones que son integrables en el sentido de Riemann.

8.1. Desigualdad de Hölder

A diferencia de la suma de funciones, que hemos manejado con frecuencia en el estudio de la integración, poco hemos dicho sobre el producto de funciones. Sabemos que el producto de dos funciones reales medibles es medible, pero más adelante veremos que el producto de dos funciones integrables puede no ser integrable, así que es natural buscar condiciones suficientes para que el producto de dos funciones sea integrable, independientemente de que los factores lo sean o no. Como siempre, seguimos trabajando en un conjunto medible $\Omega \subset \mathbb{R}^N$.

Para motivar la condición suficiente que vamos a obtener, merece la pena comentar un caso particular muy sencillo. Para $a,b \in \mathbb{R}_0^+$ es evidente que $a^2 + b^2 - 2ab = (a-b)^2 \geqslant 0$, de donde deducimos que

$$ab \leqslant \frac{a^2}{2} + \frac{b^2}{2} \qquad \forall a,b \in \mathbb{R}_0^+ \tag{1}$$

Para dos funciones $f,g : \Omega \to \mathbb{R}$, tenemos por tanto que $2|fg| \leqslant f^2 + g^2$, de donde deducimos que fg es integrable siempre que lo sean f^2 y g^2. Pues bien, vamos a generalizar esta sencilla observación, sustituyendo f^2 y g^2 por dos potencias de $|f|$ y $|g|$, cuyos exponentes son conjugados, en el sentido que pasamos a explicar.

Dado $p \in \mathbb{R}$ con $p > 1$, definimos su **exponente conjugado** $p^* \in \mathbb{R}$, mediante la igualdad

$$\frac{1}{p} + \frac{1}{p^*} = 1$$

Vemos que también $p^* > 1$ y que la relación entre p y p^* es simétrica: $\big(p^*\big)^* = p$. Nótese que, para $p = 2$ se tiene también $p^* = 2$. El primer paso hacia la desigualdad que buscamos es una generalización de (1) que se deduce fácilmente de la concavidad del logaritmo:

- **Desigualdad de Young.** *Para $a,b \in \mathbb{R}_0^+$ y $p \in \mathbb{R}$ con $p > 1$, se tiene:*

$$ab \leqslant \frac{a^p}{p} + \frac{b^{p^*}}{p^*} \tag{2}$$

Si $ab = 0$ la desigualdad es evidente. En otro caso, es decir, para $a,b \in \mathbb{R}^+$, la concavidad del logaritmo nos dice que

$$\log\left(\frac{a^p}{p} + \frac{b^{p^*}}{p^*}\right) \geqslant \frac{1}{p}\log\left(a^p\right) + \frac{1}{p^*}\log\left(b^{p^*}\right) = \log\left(ab\right)$$

y basta usar que la exponencial es una función creciente. ■

Nuestro objetivo es usar (2) en vez de (1), obteniendo así una condición suficiente más general, para que el producto de dos funciones sea integrable. Antes de enunciarla, introducimos la noción que será clave en lo que sigue.

Fijado $p \in \mathbb{R}$ con $p > 1$ y dada una función medible $f \in \mathcal{L}(\Omega)$, observamos que $|f|^p$ es una función medible positiva, por ser la composición de $|f|$ con la función $t \mapsto t^p$, de $\mathbb{R}_0^+$ en sí mismo, que es continua.

Pues bien, diremos que $f \in \mathcal{L}(\Omega)$ es una **función p-integrable** cuando $\displaystyle\int_\Omega |f|^p < \infty$, y denotaremos por $\mathcal{L}_p(\Omega)$ al conjunto de las funciones p-integrables, es decir:

$$\mathcal{L}_p(\Omega) = \left\{ f \in \mathcal{L}(\Omega) \, : \, \int_\Omega |f|^p < \infty \right\}$$

Obsérvese que, para $p = 1$ recuperaríamos la definición de función integrable. Esto explica el subíndice 1 que venimos usando para el espacio $\mathcal{L}_1(\Omega)$.

Desigualdad integral de Hölder. *Dado $p \in \mathbb{R}$ con $p > 1$, si $f \in \mathcal{L}_p(\Omega)$ y $g \in \mathcal{L}_{p^*}(\Omega)$, entonces la función producto fg es integrable y de hecho se tiene que*

$$\int_\Omega |fg| \leqslant \left(\int_\Omega |f|^p \right)^{1/p} \left(\int_\Omega |g|^{p^*} \right)^{1/p^*} \tag{3}$$

Demostración. Si α y β son los dos factores del segundo miembro, la desigualdad es inmediata cuando $\alpha\beta = 0$. En efecto, si $\alpha = 0$, tenemos $|f|^p = 0$ c.p.d., o lo que es lo mismo, $f = 0$ c.p.d., de donde deducimos que $|fg| = 0$ c.p.d., luego el primer miembro de la desigualdad buscada se anula. En el caso $\beta = 0$ se razona análogamente, con g en vez de f.

Suponiendo ahora que $\alpha, \beta \in \mathbb{R}^+$, la desigualdad de Young nos dice que

$$\frac{|f(x)|\,|g(x)|}{\alpha\beta} \leqslant \frac{|f(x)|^p}{p\,\alpha^p} + \frac{|g(x)|^{p^*}}{p^*\,\beta^{\,p^*}} \qquad \forall x \in \Omega$$

y el crecimiento de la integral nos permite deducir que

$$\frac{1}{\alpha\,\beta} \int_\Omega |fg| \leqslant \frac{1}{p\,\alpha^p} \int_\Omega |f|^p + \frac{1}{p^*\,\beta^{\,p^*}} \int_\Omega |g|^{p^*} = \frac{1}{p} + \frac{1}{p^*} = 1$$

Esto prueba que $fg \in \mathcal{L}_1(\Omega)$ y nos da la desigualdad buscada. ■

Aunque no la vamos a usar en lo que sigue, conviene conocer la versión más elemental de la desigualdad anterior, que en vez de integrales, involucra sumas finitas:

Desigualdad de Hölder para sumas finitas. *Sean $n \in \mathbb{N}$ y $(a_1, a_2, \ldots, a_n)$, $(b_1, b_2, \ldots, b_n)$ dos n-uplas de números reales no negativos. Entonces, para $p \in \mathbb{R}$ con $p > 1$, se tiene:*

$$\sum_{k=1}^n a_k b_k \leqslant \left(\sum_{k=1}^n a^p \right)^{1/p} \left(\sum_{k=1}^n b^{p^*} \right)^{1/p^*} \tag{4}$$

Se puede hacer un razonamiento similar al usado para la versión integral de esta desigualdad, pero merece la pena ver que, de hecho, (4) es caso particular de (3). Para ello fijamos n intervalos dos a dos disjuntos de longitud 1. Por ejemplo, podemos tomar $J_k =]k, k+1[$ para todo $k \in \Delta_n$. Consideramos entonces las funciones simples positivas dadas por

$$f = \sum_{k=1}^n a_k \chi_{J_k} \qquad \text{y} \qquad g = \sum_{k=1}^n b_k \chi_{J_k}$$

Es obvio que $f \in \mathcal{L}_p(\Omega)$ y $g \in \mathcal{L}_{p^*}(\Omega)$, así como que $f g \in \mathcal{L}_1(\Omega)$, verificándose que

$$\int_\Omega |f|^p = \sum_{k=1}^n a^p, \qquad \int_\Omega |g|^{p^*} = \sum_{k=1}^n b^{p^*}, \qquad \int_\Omega |f g| = \sum_{k=1}^n a_k b_k$$

Por tanto, usando (3) obtenemos directamente (4). ■

8.2. Desigualdad de Minkowski

Antes de obtener esta otra desigualdad importante, probamos una propiedad de $\mathcal{L}_p(\Omega)$ que ya conocemos en el caso $p=1$:

- *Para $p \in \mathbb{R}$ con $p > 1$, se tiene que $\mathcal{L}_p(\Omega)$ es un subespacio vectorial de $\mathcal{L}(\Omega)$.*

Para $f \in \mathcal{L}_p(\Omega)$ y $\alpha \in \mathbb{R}$, es evidente que $\alpha f \in \mathcal{L}_p(\Omega)$, ya que

$$\int_\Omega |\alpha f|^p = |\alpha|^p \int_\Omega |f|^p < \infty$$

Fijamos ahora $f, g \in \mathcal{L}_p(\Omega)$ para probar que $f+g$ es una función p-integrable. Para ello, escribimos $\Omega = A \uplus B$ donde A y B son los conjuntos medibles definidos por

$$A = \big\{ x \in \Omega \, : \, |f(x)| \leqslant |g(x)| \big\} \qquad \text{y} \qquad B = \big\{ x \in \Omega \, : \, |f(x)| > |g(x)| \big\}$$

Para todo $x \in A$ tenemos que

$$|f(x)+g(x)|^p \leqslant \big(|f(x)| + |g(x)| \big)^p \leqslant 2^p \, |g(x)|^p$$

y análogamente, $|f(x)+g(x)|^p \leqslant 2^p |f(x)|^p$ para todo $x \in B$. Deducimos claramente que

$$\begin{aligned} \int_\Omega |f+g|^p &= \int_A |f+g|^p + \int_B |f+g|^p \leqslant 2^p \int_A |g|^p + 2^p \int_B |f|^p \\ &\leqslant 2^p \int_\Omega |g|^p + 2^p \int_\Omega |f|^p < \infty \end{aligned}$$

y esto prueba que $f+g \in \mathcal{L}_p(\Omega)$, como se quería. ■

Desigualdad integral de Minkowski. *Para $p \in \mathbb{R}$ con $p \geqslant 1$ y $f, g \in \mathcal{L}_p(\Omega)$, se tiene:*

$$\left(\int_\Omega |f+g|^p \right)^{1/p} \leqslant \left(\int_\Omega |f|^p \right)^{1/p} + \left(\int_\Omega |g|^p \right)^{1/p}$$

Demostración. En el caso $p=1$ basta tener en cuenta que $|f+g| \leqslant |f| + |g|$. Para $p>1$, sea γ el primer miembro de la desigualdad buscada y llamemos α y β a los dos sumandos del segundo miembro. Por el resultado previo sabemos que $\gamma < \infty$ y la desigualdad buscada es obvia si $\gamma = 0$, así que suponemos $\gamma \in \mathbb{R}^+$.

La función medible $h = |f+g|^{p-1}$, verifica que $h^{p^*} = |f+g|^{(p-1)p^*} = |f+g|^p$, de donde deducimos que
$$\int_\Omega h^{p^*} = \int_\Omega |f+g|^p = \gamma^p < \infty$$
luego $h \in \mathcal{L}_{p^*}(\Omega)$. Observamos que $|f+g|^p = |f+g|h \leqslant |f|h + |g|h$ y podemos usar la desigualdad de Hölder para los dos productos que han aparecido, que tienen h como factor común, obteniendo
$$\gamma^p \leqslant \int_\Omega |f|h + \int_\Omega |g|h \leqslant (\alpha+\beta)\left(\int_\Omega h^{p^*}\right)^{1/p^*} = (\alpha+\beta)\,\gamma^{p/p^*}$$
Teniendo en cuenta que $p-(p/p^*)=1$, basta dividir en ambos miembros por $\gamma^{p/p^*} \in \mathbb{R}^+$, para obtener que $\gamma \leqslant \alpha+\beta$, como se quería. ■

Razonando exactamente igual que se hizo con la desigualdad de Hölder, obtenemos:

Desigualdad de Minkowski para sumas finitas. *Fijado $p \in \mathbb{R}$ con $p \geqslant 1$, para $n \in \mathbb{N}$ y cualesquiera $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n \in \mathbb{R}_0^+$, se tiene:*
$$\left(\sum_{k=1}^n (a_k+b_k)^p\right)^{1/p} \leqslant \left(\sum_{k=1}^n a^p\right)^{1/p} + \left(\sum_{k=1}^n b^p\right)^{1/p}$$

8.3. Definición de los espacios de Lebesgue

Todo está ya preparado para definir una gama de espacios normados, que en cierto modo protagonizan la teoría de la integración. Fijado el exponente $p \in \mathbb{R}$ con $p \geqslant 1$, consideremos la función $\varphi_p : \mathcal{L}_p(\Omega) \to \mathbb{R}_0^+$ definida por:
$$\varphi_p(f) = \left(\int_\Omega |f|^p\right)^{1/p} \qquad \forall f \in \mathcal{L}_p(\Omega)$$
La desigualdad integral de Minkowski nos dice que φ_p verifica la desigualdad triangular:
$$\varphi_p(f+g) \leqslant \varphi_p(f) + \varphi_p(g) \qquad \forall f,g \in \mathcal{L}(\Omega)$$
Por otra parte, de la definición de φ_p se deduce claramente que
$$\varphi_p(\alpha f) = |\alpha|\,\varphi_p(f) \qquad \forall \alpha \in \mathbb{R},\ \forall f \in \mathcal{L}_p(\Omega)$$
Las dos propiedades obtenidas se resumen diciendo que φ_p es una **seminorma** en el espacio vectorial $\mathcal{L}_p(\Omega)$. No llega a ser una norma, pues para $f \in \mathcal{L}_p(\Omega)$, vemos que
$$\varphi_p(f) = 0 \iff \int_\Omega |f|^p = 0 \iff f = 0 \ \text{c.p.d.}$$
y esto no implica que f sea idénticamente nula. Denotaremos por $\mathcal{N}(\Omega)$ al conjunto de todas las funciones de Ω en $\mathbb{R}$ que se anulan c.p.d. y conviene aclarar que todas estas funciones son medibles.

- *Toda función* $f \in \mathcal{N}(\Omega)$ *es medible*

Por hipótesis, el conjunto $E = \{x \in \Omega : f(x) \neq 0\}$ tiene medida nula, luego es medible, e igual le ocurre a $\Omega \setminus E$. Para $H \subset \mathbb{R}$ con $0 \notin H$, se tiene $f^{-1}(H) \subset E$, luego $f^{-1}(H) \in \mathcal{M}$, pero si $0 \in H$, escribimos $f^{-1}(H) = f^{-1}\big(H \setminus \{0\}\big) \biguplus (\Omega \setminus E)$ y vemos que $f^{-1}(H)$ es medible, como unión de dos conjuntos medibles. Así pues, $f^{-1}(H)$ es medible para todo $H \in \mathcal{P}(\mathbb{R})$ y, en particular, f es medible. ■

Deducimos claramente que si dos funciones $f, g : \Omega \to \mathbb{R}$ coinciden c.p.d., entonces f es medible si, y sólo si lo es g. Esta observación tendrá bastante utilidad más adelante.

Está claro ahora que $\mathcal{N}(\Omega)$ es un subespacio de $\mathcal{L}_p(\Omega)$, que según hemos visto, coincide con el núcleo de la seminorma φ_p, aunque conviene observar que $\mathcal{N}(\Omega)$ no depende de p. Se define ahora el **espacio de Lebesgue** $L_p(\Omega)$ como el espacio vectorial cociente

$$L_p(\Omega) = \mathcal{L}_p(\Omega) / \mathcal{N}(\Omega) = \big\{ f + \mathcal{N}(\Omega) \, : \, f \in \mathcal{L}_p(\Omega) \big\}$$

al que enseguida convertiremos en un espacio normado. Fijada $f \in \mathcal{L}_p(\Omega)$ para cualquier otra función $g \in f + \mathcal{N}(\Omega)$ de la misma clase de equivalencia, tenemos $|g|^p = |f|^p$ c.p.d., luego ambas funciones tienen la misma integral. Por tanto, escribiendo

$$\big\| f + \mathcal{N}(\Omega) \big\|_p = \varphi_p(f) = \left(\int_\Omega |f|^p \right)^{1/p}$$

vemos que $f + \mathcal{N}(\Omega) \mapsto \big\| f + \mathcal{N}(\Omega) \big\|_p$ es una aplicación bien definida $\| \cdot \|_p : L_p(\Omega) \to \mathbb{R}_0^+$. Al ser φ_p una seminorma, se comprueba rutinariamente que $\| \cdot \|_p$ también lo es, pero ahora, de $\big\| f + \mathcal{N}(\Omega) \big\|_p = 0$ se deduce que $f \in \mathcal{N}(\Omega)$, es decir, que $f + \mathcal{N}(\Omega)$ es el vector 0 del espacio cociente $L_p(\Omega)$. Por tanto $\| \cdot \|_p$ es una norma, la que siempre se usa en $L_p(\Omega)$.

Es una sana costumbre manejar los elementos de $L_p(\Omega)$ como si fuesen funciones, en vez de clases de equivalencia, y hablar de una "función" $f \in L_p(\Omega)$ para referirse en realidad a la clase de equivalencia $f + \mathcal{N}(\Omega)$. Es un abuso de lenguaje que, si se maneja con la debida precaución, no da lugar a errores, y que a veces llega a ser imprescindible, porque la reiterada distinción entre funciones y clases de equivalencia acaba siendo demasiado tediosa. Preferimos mantener aquí esa distinción, pero sí usaremos una notación abreviada. A partir de ahora, $\widetilde{f}$ será la clase de equivalencia en $L_p(\Omega)$ a la que pertenece una función $f \in \mathcal{L}_p(\Omega)$, es decir,

$$\widetilde{f} = f + \mathcal{N}(\Omega) \qquad \forall f \in \mathcal{L}_p(\Omega)$$

8.4. Complitud de los espacios de Lebesgue

Vamos a probar ahora la propiedad fundamental de los espacios de Lebesgue:

Teorema de Riesz-Fischer. *Para todo conjunto medible* $\Omega \subset \mathbb{R}^N$ *y todo* $p \in \mathbb{R}$ *con* $p \geqslant 1$, *el espacio normado* $L_p(\Omega)$ *es completo, es decir, es un espacio de Banach.*

Demostración. Sea $f_n \in \mathcal{L}_p(\Omega)$ para todo $n \in \mathbb{N}$, y suponiendo que $\{\widetilde{f_n}\}$ es una sucesión de Cauchy en $L_p(\Omega)$, probaremos que converge. Para cada $n \in \mathbb{N}$, existe $k_n \in \mathbb{N}$ tal que

$$i, j \in \mathbb{N}, \; i, j \geqslant k_n \implies \varphi_p(f_i - f_j) = \|\widetilde{f_i} - \widetilde{f_j}\|_p < \frac{1}{2^n} \tag{5}$$

Tomando por inducción $\sigma(1) = k_1$ y $\sigma(n+1) = \max\{k_{n+1}, \sigma(n)+1\}$ para todo $n \in \mathbb{N}$, obtenemos una aplicación estrictamente creciente $\sigma : \mathbb{N} \to \mathbb{N}$, con lo que $\{\widetilde{f}_{\sigma(n)}\}$ es una sucesión parcial de $\{\widetilde{f_n}\}$. Es sabido que, en cualquier espacio métrico, toda sucesión de Cauchy que admita una sucesión parcial convergente, es convergente. Por tanto, escribiendo $g_n = f_{\sigma(n)}$ para todo $n \in \mathbb{N}$, bastará probar que la sucesión $\{\widetilde{g_n}\}$ es convergente. La ventaja de esta sucesión radica en que controlamos mejor la distancia entre sus términos. Concretamente, para todo $n \in \mathbb{N}$ se tiene $\sigma(n+1) > \sigma(n) \geqslant k_n$, y vemos en (5) que

$$\varphi_p(g_{n+1} - g_n) = \varphi_p(f_{\sigma(n+1)} - f_{\sigma(n)}) < \frac{1}{2^n} \qquad \forall n \in \mathbb{N} \tag{6}$$

Conviene expresar la sucesión $\{g_n\}$ como una serie, para usar la noción de convergencia absoluta. Concretamente, tomando por comodidad $g_0 = 0$, tenemos $\{g_n\} = \sum_{n \geqslant 1} (g_n - g_{n-1})$. Con el fin de estudiar la convergencia absoluta de esta serie, escribimos

$$\rho_n = \sum_{k=1}^{n} |g_k - g_{k-1}| \quad \forall n \in \mathbb{N} \qquad \text{y} \qquad \rho = \sum_{n=1}^{\infty} |g_n - g_{n-1}|$$

Entonces $\{\rho_n\}$ es una sucesión creciente de funciones medibles positivas, tal que $\{\rho_n\} \nearrow \rho$, con lo que la función $\rho : \Omega \to [0, \infty]$ también es medible.

Para todo $k \in \mathbb{N}$, tenemos claramente que $|g_k - g_{k-1}| \in \mathcal{L}_p(\Omega)$, y si $k \geqslant 2$, vemos en (6) que $\varphi_p(|g_k - g_{k-1}|) = \varphi_p(g_k - g_{k-1}) < 1/2^{k-1}$. Por tanto, $\rho_n \in \mathcal{L}_p(\Omega)$ para todo $n \in \mathbb{N}$ y, también para $n \geqslant 2$, la desigualdad triangular nos permite obtener que

$$\varphi_p(\rho_n) \leqslant \sum_{k=1}^{n} \varphi_p(|g_k - g_{k-1}|) < \varphi_p(g_1) + \sum_{k=2}^{n} \frac{1}{2^{k-1}} < \varphi_p(g_1) + 1$$

Esto también es cierto para $n = 1$, y tomando $M = (\varphi_p(g_1) + 1)^p$ tenemos

$$\int_\Omega \rho_n^p = \varphi_p(\rho_n)^p \leqslant M \qquad \forall n \in \mathbb{N}$$

Observamos ahora que $\{\rho_n^p\} \nearrow \rho^p$, donde se entiende que $\infty^p = \infty$. El teorema de la convergencia monótona nos dice entonces que

$$\int_\Omega \rho^p = \lim_{n \to \infty} \int_\Omega \rho_n^p = \lim_{n \to \infty} \varphi_p(\rho_n) \leqslant M < \infty$$

Así pues, la función medible positiva ρ^p tiene integral finita, luego $\rho(x)^p < \infty$ p.c.t. $x \in \Omega$. Equivalentemente, el conjunto medible $A = \{x \in \Omega : \rho(x) < \infty\}$ verifica que $\lambda(\Omega \setminus A) = 0$.

Para $x \in A$ tenemos $\sum_{n=1}^{\infty} |g_n(x) - g_{n-1}(x)| = \rho(x) < \infty$, luego la serie $\sum_{n \geqslant 1} \big(g_n(x) - g_{n-1}(x)\big)$ es convergente en $\mathbb{R}$, es decir, la sucesión de números reales $\{g_n(x)\}$ converge. Para cada $n \in \mathbb{N}$ definimos ahora $h_n = \chi_A\, g_n$, con lo que h_n es medible y $h_n = g_n$ c.p.d., luego $h_n \in \mathcal{L}_p(\Omega)$ y vemos que $\widetilde{h_n} = \widetilde{g_n}$.

Además, la sucesión $\{h_n\}$ converge puntualmente en Ω, pues para $x \in A$, hemos visto que la sucesión $\{h_n(x)\} = \{g_n(x)\}$ es convergente, mientras que si $x \in \Omega \setminus A$, se tiene $h_n(x) = 0$ para todo $n \in \mathbb{N}$. Tenemos por tanto una función medible $h : \Omega \to \mathbb{R}$ que verifica:

$$h(x) = \lim_{n \to \infty} h_n(x) \qquad \forall x \in \Omega$$

La demostración se concluirá viendo que $h \in \mathcal{L}_p(\Omega)$ y que $\{\widetilde{g_n}\} = \{\widetilde{h_n}\}$ converge a $\widetilde{h}$ en $L_p(\Omega)$. Resaltamos que $\{h_n\}$ verifica la condición análoga a (6), o más concretamente, usaremos que

$$\varphi_p\big(h_k - h_{k+1}\big) = \varphi_p\big(g_k - g_{k+1}\big) < \frac{1}{2^k} \qquad \forall k \in \mathbb{N}$$

Fijados $m, n \in \mathbb{N}$ la desigualdad triangular nos dice entonces que

$$\varphi_p\big(h_m - h_{m+n}\big) \leqslant \sum_{k=m}^{m+n-1} \varphi_p\big(h_k - h_{k+1}\big) < \sum_{k=m}^{m+n-1} \frac{1}{2^k} < \frac{1}{2^{m-1}}$$

Manteniendo por ahora $m \in \mathbb{N}$ fijo, vemos que $\{|h_m - h_{m+n}|^p\}$ es una sucesión de funciones medibles positivas, que converge puntualmente en Ω a la función $|h_m - h|^p$. El lema de Fatou nos dice entonces que

$$\begin{aligned}\int_\Omega \big|h_m - h\big|^p &\leqslant \liminf_{n \to \infty} \int_\Omega \big|h_m - h_{m+n}\big|^p \\ &= \liminf_{n \to \infty} \varphi_p\big(\big|h_m - h_{m+n}\big|\big)^p \leqslant \left(\frac{1}{2^{m-1}}\right)^p < \infty\end{aligned}$$

Así pues, $h_m - h \in \mathcal{L}_p(\Omega)$, luego $h = h_m - (h_m - h) \in \mathcal{L}_p(\Omega)$. Pero además, la desigualdad recién probada, válida para todo $m \in \mathbb{N}$, nos dice ahora que

$$\big\|\, \widetilde{h_m} - \widetilde{h} \,\big\|_p = \varphi_p\big(h_m - h\big) \leqslant \frac{1}{2^{m-1}} \qquad \forall m \in \mathbb{N}$$

Esto prueba que la sucesión $\{\widetilde{h_n}\}$ converge a $\widetilde{h}$ en $L_p(\Omega)$, como se quería. ■

El nombre del teorema anterior es algo discutible. Lo que en 1907 prueban, por separado, el matemático húngaro F. Riesz (1880-1956) y el austriaco E. Fischer (1875-1954), es un teorema sobre series de Fourier, que pone de manifiesto la gran utilidad de la integral de Lebesgue en el estudio de tales series. Para obtenerlo, prueban la complitud de $L_2(I)$, para un intervalo compacto $I \subset \mathbb{R}$, de la que se deduce inmediatamente el resultado buscado. Puede entenderse por tanto, que su teorema equivale a la complitud de $L_2(I)$. La demostración es prácticamente la misma que acabamos de hacer, y de ahí que por extensión, se le dé el mismo nombre al teorema anterior y a otros resultados aún más generales.

Por otra parte, conviene resaltar la relación entre dos tipos de convergencia que aparecen en la demostración anterior. Con la notación usada, hemos probado que la sucesión $\{\widetilde{h_n}\}$ converge a $\widetilde{h}$ en $L_p(\Omega)$, pero también se tiene que $\{h_n\}$ converge a h puntualmente en Ω. Esta información se aprovecha en el siguiente enunciado para obtener una relación entre ambos tipos de convergencia.

- *Dadas $f \in \mathcal{L}_p(\Omega)$ y $f_n \in \mathcal{L}_p(\Omega)$ para todo $n \in \mathbb{N}$, supongamos que la sucesión $\{\widetilde{f_n}\}$ converge a $\widetilde{f}$ en el espacio de Banach $L_p(\Omega)$. Entonces la sucesión de funciones $\{f_n\}$ tiene una sucesión parcial $\{f_{\sigma(n)}\}$ tal que:*

$$f(x) = \lim_{n\to\infty} f_{\sigma(n)}(x) \qquad \text{p.c.t. } x \in \Omega \tag{7}$$

Para obtener este resultado, basta analizar un poco la demostración del teorema anterior. Como $\{\widetilde{f_n}\}$ es una sucesión de Cauchy en $L_p(\Omega)$, dicho razonamiento nos da una sucesión parcial $\{g_n\} = \{f_{\sigma(n)}\}$, de la que obtenemos un conjunto $A \subset \Omega$, con $\lambda(\Omega \setminus A) = 0$, tal que la sucesión de funciones $\{h_n\} = \{\chi_A\, g_n\}$ converge puntualmente en Ω a una función $h \in \mathcal{L}_p(\Omega)$, para obtener finalmente que $\{\widetilde{h_n}\}$ converge a $\widetilde{h}$ en $L_p(\Omega)$. Observamos ahora que, si $x \in A$, se tiene $h_n(x) = f_{\sigma(n)}(x)$ para todo $n \in \mathbb{N}$, luego $\{f_{\sigma(n)}(x)\} \to h(x)$. Pero por otra parte, también tenemos que $\widetilde{h_n} = \widetilde{g_n} = \widetilde{f}_{\sigma(n)}$ para todo $n \in \mathbb{N}$, y como ahora estamos suponiendo que $\{\widetilde{f_n}\}$ converge a $\widetilde{f}$, deducimos que la sucesión parcial $\{\widetilde{h_n}\}$ también converge a $\widetilde{f}$ en $L_p(\Omega)$. Esto implica que $\widetilde{h} = \widetilde{f}$, es decir, existe otro conjunto medible $B \subset \Omega$, también con $\lambda(\Omega \setminus B) = 0$, tal que $h(x) = f(x)$ para todo $x \in B$. Se tiene por tanto que

$$\lim_{n\to\infty} f_{\sigma(n)}(x) = h(x) = f(x) \qquad \forall x \in A \cap B$$

Como $\lambda\big(\Omega \setminus (A \cap B)\big) \leqslant \lambda(\Omega \setminus A) + \lambda(\Omega \setminus B) = 0$, hemos probado (7). ■

Merece la pena comentar el tipo de convergencia que ha aparecido en (7). Es natural decir que una sucesión $\{f_n\}$ de funciones de Ω en $\mathbb{R}$, **converge casi por doquier** (abreviado c.p.d.) a una función $f : \Omega \to \mathbb{R}$, cuando $\{f_n(x)\} \to f(x)$ p.c.t. $x \in \Omega$, es decir, cuando existe un conjunto de medida nula $E \subset \Omega$, tal que $\{f_n(x)\} \to f(x)$ para todo $x \in \Omega \setminus E$. Nótese que la función f no queda determinada de manera única. De hecho, vemos que $\{f_n\}$ converge c.p.d. a otra función $g : \Omega \to \mathbb{R}$ si, y sólo si, $g = f$ c.p.d. Esto hace que la convergencia c.p.d. esté más indicada para trabajar con clases de equivalencia en los espacios de Lebesgue, cuestión en la que no vamos a entrar.

8.5. Funciones simples integrables

Extendiendo la noción de función simple positiva, es natural llamar función simple a toda combinación lineal de funciones características de conjuntos medibles. Razonando de forma análoga a como lo hicimos en el caso de funciones positivas, es fácil ver que las funciones simples no son, ni más ni menos, que las funciones medibles de $\mathbb{R}^N$ en $\mathbb{R}$ con imagen finita.

De manera más general, para poder trabajar como siempre en un conjunto medible $\Omega \subset \mathbb{R}^N$, diremos que $s : \Omega \to \mathbb{R}$ es una **función simple** en Ω cuando s sea medible y $s(\Omega)$ sea un conjunto finito.

El ejemplo más sencillo es la restricción a Ω de χ_E donde E es un subconjunto medible de Ω. Vemos que, para $p \in \mathbb{R}$ con $p \geqslant 1$, se tiene $|\chi_E|^p = \chi_E$, mientras que $\int_\Omega \chi_E = \lambda(E)$. Por tanto, χ_E es p-integrable en Ω para todo $p \in \mathbb{R}$ con $p \geqslant 1$ si, y sólo si, $\lambda(E) < \infty$. A continuación obtenemos un resultado análogo para funciones simples.

- *Para una función simple $s : \Omega \to \mathbb{R}$, las siguientes afirmaciones son equivalentes:*

  (i) *Existe $p \in \mathbb{R}$ con $p \geqslant 1$ tal que $s \in \mathcal{L}_p(\Omega)$*

  (ii) *La función s es la restricción a Ω de una combinación lineal de funciones características de subconjuntos medibles de Ω, con medida finita, es decir:*

$$s(x) = \sum_{k=1}^{n} \alpha_k \chi_{A_k}(x) \qquad \forall x \in \Omega \tag{8}$$

  donde $n \in \mathbb{N}$ y, para $k \in \Delta_n$, se tiene $\alpha_k \in \mathbb{R}$ y $A_k \in \mathcal{M} \cap \mathcal{P}(\Omega)$ con $\lambda(A_k) < \infty$

  (iii) *Se tiene $s \in \mathcal{L}_p(\Omega)$, para todo $p \in \mathbb{R}$ con $p \geqslant 1$.*

De la observación hecha previamente, para el caso de una función característica, se deduce claramente que $(ii) \Rightarrow (iii)$, y es obvio que $(iii) \Rightarrow (i)$, luego basta probar que $(i) \Rightarrow (ii)$. Sea pues $p \in \mathbb{R}$ con $p \geqslant 1$, tal que $s \in \mathcal{L}_p(\Omega)$ y, suponiendo $s \neq 0$, usaremos algo parecido a la descomposición canónica de una función simple positiva.

Como $s(\Omega)$ es finito, pongamos $s(\Omega) \setminus \{0\} = \{\alpha_k : k \in \Delta_n\}$, para conveniente $n \in \mathbb{N}$, sin repetir valores de s, es decir, $\alpha_k \neq \alpha_j$ para $k, j \in \Delta_n$ con $k \neq j$. Entonces, para cada $k \in \Delta_n$ tomamos $A_k = \{x \in \Omega : s(x) = \alpha_k\} \in \mathcal{M} \cap \mathcal{P}(\Omega)$, con lo que se verifica (8), pero queda comprobar que $\lambda(A_k) < \infty$ para todo $k \in \Delta_n$. Para ello observamos que $\Omega = \biguplus_{k=0}^{n} A_k$, donde hemos escrito $A_0 = \{x \in \Omega : s(x) = 0\} \in \mathcal{M}$. Deducimos claramente que

$$\sum_{k=1}^{n} |\alpha_k|^p \lambda(A_k) = \sum_{k=1}^{n} \int_{A_k} |s|^p = \sum_{k=0}^{n} \int_{A_k} |s|^p = \int_\Omega |s|^p < \infty$$

Por tanto, para cada $k \in \Delta_n$ se tiene $|\alpha_k|^p \lambda(A_k) < \infty$ con $\alpha_k \neq 0$, luego $\lambda(A_k) < \infty$. ■

Naturalmente, cuando una función simple $s : \Omega \to \mathbb{R}$ verifica las afirmaciones equivalentes del enunciado anterior, decimos que s es una **función simple integrable** en Ω. Denotaremos por $\mathcal{S}(\Omega)$ al conjunto de tales funciones y, fijado $p \in \mathbb{R}$ con $p \geqslant 1$, hemos visto que $\mathcal{S}(\Omega)$ es un subespacio vectorial de $\mathcal{L}_p(\Omega)$. En el cociente, tenemos el conjunto $S(\Omega) = \{\widetilde{s} : s \in \mathcal{S}(\Omega)\}$, de las clases de equivalencia que contienen una función simple integrable en Ω, que es un subespacio vectorial de $L_p(\Omega)$. Todo está ya preparado para un teorema que permite aproximar funciones p-integrables por funciones simples, obteniendo un subespacio denso en $L_p(\Omega)$.

Primer teorema de densidad. *Dado un conjunto medible $\Omega \subset \mathbb{R}^N$, y $p \in \mathbb{R}$ con $p \geqslant 1$, para cada $f \in \mathcal{L}_p(\Omega)$, existe una sucesión $\{s_n\}$ de funciones simples integrables en Ω, que converge puntualmente a f en Ω y verifica:*

$$\lim_{n\to\infty} \int_\Omega |f - s_n|^p = 0 \tag{9}$$

Como consecuencia, el conjunto $S(\Omega)$ de las clases de equivalencia que contienen una función simple integrable en Ω, es denso en $L_p(\Omega)$.

Demostración. Fijada una $f \in \mathcal{L}_p(\Omega)$, como f^+ y f^- son funciones medibles positivas, el teorema de aproximación de Lebesgue nos da dos sucesiones crecientes $\{u_n\}$ y $\{v_n\}$ de funciones simples positivas, que convergen puntualmente en Ω a f^+ y f^- respectivamente. Conviene además resaltar que, para cada $n \in \mathbb{N}$, las funciones u_n y v_n son combinaciones lineales de funciones características de subconjuntos medibles de Ω.

Si para $n \in \mathbb{N}$ llamamos s_n a la restricción a Ω de $u_n - v_n$, obtenemos una sucesión $\{s_n\}$ de funciones simples en Ω, que converge puntualmente en Ω a $f^+ - f^- = f$. Además, para cada $n \in \mathbb{N}$ se tiene que

$$|s_n(x)| \leqslant u_n(x) + v_n(x) \leqslant f^+(x) + f^-(x) = |f(x)| \qquad \forall x \in \Omega$$

y como f es p-integrable en Ω, deducimos que s_n también lo es, con lo que $s_n \in \mathcal{S}(\Omega)$. Para probar (9) vemos que $\{|f - s_n|^p\}$ es una sucesión de funciones medibles, que converge puntualmente a cero en Ω y verifica que

$$|f(x) - s_n(x)|^p \leqslant \big(|s_n(x)| + |f(x)|\big)^p \leqslant 2^p |f(x)|^p \qquad \forall x \in \Omega, \ \forall n \in \mathbb{N}$$

Como la función $2^p |f|^p$ es integrable en Ω, podemos usar el teorema de la convergencia dominada para obtener (9).

Por otra parte, cualquier clase de equivalencia en $L_p(\Omega)$ se escribe como $\widetilde{f}$ con $f \in \mathcal{L}_p(\Omega)$. Usando lo que ya hemos demostrado, obtenemos una sucesión $\{\widetilde{s}_n\}$ en $S(\Omega)$ que, en vista de la igualdad (9) verifica: $\lim\limits_{n\to\infty} \|\widetilde{f} - \widetilde{s}_n\|_p = 0$. Así pues, $\{\widetilde{s}_n\}$ converge a $\widetilde{f}$ en $L_p(\Omega)$, y esto prueba que $S(\Omega)$ es denso en $L_p(\Omega)$. ■

En realidad, el resultado anterior es un paso intermedio hacia otros teoremas de densidad más útiles, que vamos a obtener en lo que sigue.

8.6. Funciones escalonadas

Las funciones simples toman sólo un conjunto finito de valores, pero pueden llegar a ser muy complicadas desde un punto vista constructivo, tanto como pueden serlo los conjuntos medibles. En el caso $\Omega = \mathbb{R}^N$, que es el más interesante, vamos a mejorar el teorema anterior, aproximando por funciones simples que sólo involucran los conjuntos medibles más sencillos, que son los intervalos acotados en $\mathbb{R}^N$. Recordemos que $\mathcal{J}$ es la familia de dichos intervalos.

Llamaremos **función escalonada** a toda combinación lineal de funciones características de intervalos acotados, es decir, a toda función $h:\mathbb{R}^N \to \mathbb{R}$ de la forma

$$h = \sum_{k=1}^{n} \alpha_k \, \chi_{J_k} \qquad \text{donde} \qquad n \in \mathbb{N}, \ \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}, \ J_1, J_2, \ldots, J_n \in \mathcal{J}$$

Es claro que entonces h es una función simple integrable en $\mathbb{R}^N$, con

$$\int_{\mathbb{R}^N} h \, = \, \sum_{k=1}^{n} \alpha_k \, \lambda(J_k)$$

Para trabajar como siempre en un conjunto medible $\Omega \subset \mathbb{R}^N$, denotaremos por $\mathcal{E}(\Omega)$ al conjunto de las restricciones a Ω de funciones escalonadas, a las que podemos referirnos como funciones escalonadas *en* Ω. Hemos visto que se trata de funciones simples integrables en Ω, pero de un tipo muy particular, pues se obtienen usando solamente funciones características de intervalos acotados, en vez de conjuntos medibles más generales.

Fijado $p \in \mathbb{R}$ con $p \geqslant 1$, vemos que $\mathcal{E}(\Omega)$ es un subespacio vectorial de $\mathcal{L}_p(\Omega)$, que está contenido en $\mathcal{S}(\Omega)$. Al pasar a cociente, el conjunto $E(\Omega) = \big\{ \widetilde{g} : g \in \mathcal{E}(\Omega) \big\}$ es a su vez un subespacio vectorial de $L_p(\Omega)$, contenido en $S(\Omega)$. El primer teorema de densidad nos dice que $S(\Omega)$ es denso en $L_p(\Omega)$, pero vamos a probar que de hecho $E(\Omega)$ ya lo es. Previamente hacemos una observación sobre el cierre de un subespacio en un espacio normado:

- *Si Y es un subespacio vectorial de un espacio normado X, entonces el cierre de Y también es un subespacio vectorial de X.*

Dados $u, v \in \overline{Y}$ y $\alpha \in \mathbb{R}$, se trata de probar que $\alpha u + v \in \overline{Y}$, pero esto es inmediato. Tomamos dos sucesiones convergentes $\{y_n\} \to u$ y $\{z_n\} \to v$, con $y_n, z_n \in Y$ para todo $n \in \mathbb{N}$. Entonces también $\alpha y_n + z_n \in Y$ para todo $n \in \mathbb{N}$, y tenemos

$$\| (\alpha y_n + z_n) - (\alpha u + v) \| \leqslant |\alpha| \, \| y_n - u \| + \| z_n - v \| \qquad \forall n \in \mathbb{N}$$

Deducimos claramente que $\{\alpha y_n + z_n\} \to \alpha u + v$, luego $\alpha u + v \in \overline{Y}$ como se quería. ■

Densidad de las funciones escalonadas. *Fijados un conjunto medible $\Omega \subset \mathbb{R}^N$, y $p \in \mathbb{R}$ con $p \geqslant 1$, el conjunto $E(\Omega)$, formado por todas las clases de equivalencia que contienen una función escalonada en Ω, es denso en $L_p(\Omega)$. Como consecuencia, para cada $f \in \mathcal{L}_p(\Omega)$, existe una sucesión $\{g_n\}$ de funciones escalonadas que verifica:*

$$\{g_n(x)\} \to f(x) \ \text{ p.c.t. } x \in \Omega \qquad \text{y} \qquad \lim_{n\to\infty} \int_{\Omega} |f - g_n|^p = 0 \tag{10}$$

Demostración. Gran parte del razonamiento se hará en el caso $\Omega = \mathbb{R}^N$, para luego extender el resultado al caso general. Si X_p es el cierre de $E(\mathbb{R}^N)$ en $L_p(\mathbb{R}^N)$, sabemos que X_p es un subespacio vectorial cerrado de $L_p(\mathbb{R}^N)$, pero queremos probar que $X_p = L_p(\mathbb{R}^N)$.

Empezamos con un conjunto compacto $K \subset \mathbb{R}^N$, recordando que la medida exterior de Lebesgue puede calcularse usando solamente recubrimientos por intervalos abiertos acotados. Fijado $\varepsilon > 0$, existe por tanto una sucesión $\{I_j\}$ de intervalos abiertos acotados, tal que

$$K \subset \bigcup_{j=1}^{\infty} I_j \qquad \text{y} \qquad \sum_{j=1}^{\infty} \lambda(I_j) < \lambda(K) + \varepsilon^p$$

Por ser K compacto, existe $m \in \mathbb{N}$ tal que $K \subset A = \bigcup_{j=1}^{m} I_j$. Vemos que A es una figura elemental en $\mathbb{R}^N$, que se expresa como unión finita de intervalos acotados dos a dos disjuntos:

$$K \subset A = \biguplus_{k=1}^{n} J_k\,, \qquad \text{con} \qquad n \in \mathbb{N}, \quad J_1, J_2, \ldots, J_n \in \mathcal{J}$$

Entonces, por una parte, $\chi_A = \sum_{k=1}^{n} \chi_{J_k}$ es una función escalonada, y por otra,

$$\lambda(A) \leqslant \sum_{j=1}^{m} \lambda(I_j) \leqslant \sum_{j=1}^{\infty} \lambda(I_j) < \lambda(K) + \varepsilon^p$$

de donde deducimos claramente que

$$\int_{\mathbb{R}^N} |\chi_A - \chi_K|^p = \lambda(A \setminus K) < \varepsilon^p$$

Pasando a clases de equivalencia en $L_p(\mathbb{R}^N)$, tenemos $\widetilde{\chi}_A \in E(\mathbb{R}^N)$, y la desigualdad anterior nos dice que $\|\widetilde{\chi}_A - \widetilde{\chi}_K\|_p < \varepsilon$. Como $\varepsilon > 0$ era arbitrario, hemos probado que $\widetilde{\chi}_K \in X_p$.

Sea ahora $E \in \mathcal{M}$ con $\lambda(E) < \infty$, y fijemos de nuevo $\varepsilon > 0$. La regularidad interior de la medida de Lebesgue nos da un conjunto compacto $K \subset E$ tal que $\lambda(E \setminus K) < \varepsilon^p$, con lo que

$$\int_{\mathbb{R}^N} |\chi_E - \chi_K|^p = \lambda(E \setminus K) < \varepsilon^p$$

En el cociente tenemos $\|\widetilde{\chi}_E - \widetilde{\chi}_K\|_p < \varepsilon$. Ciertamente K depende de ε, pero siempre sabemos que $\widetilde{\chi}_K \in X_p$, y siendo X_p cerrado, la arbitrariedad de ε nos permite concluir que $\widetilde{\chi}_E \in X_p$.

Como X_p es subespacio vectorial de $L_p(\mathbb{R}^N)$, de $\{\widetilde{\chi}_E : E \in \mathcal{M},\, \lambda(E) < \infty\} \subset X_p$ se deduce claramente que $S(\mathbb{R}^N) \subset X_p$. Por el primer teorema de densidad, $S(\mathbb{R}^N)$ es denso en $L_p(\mathbb{R}^N)$, luego X_p también lo es, pero X_p es cerrado, así que $X_p = L_p(\mathbb{R}^N)$ como queríamos.

Si ahora $\varphi \in \mathcal{L}_p(\mathbb{R}^N)$, tenemos una sucesión $\{h_n\}$ de funciones escalonadas tal que $\{\widetilde{h}_n\}$ converge a $\widetilde{\varphi}$ en $L_p(\mathbb{R}^N)$. Pero entonces, sabemos que $\{h_n\}$ tiene una sucesión parcial $\{g_n\}$ que converge a φ c.p.d. Como $\{\widetilde{g}_n\}$ también converge a $\widetilde{\varphi}$ en $L_p(\mathbb{R}^N)$, concluimos que

$$\lim_{n\to\infty} \int_{\mathbb{R}^N} |\varphi - g_n|^p = \lim_{n\to\infty} \left(\|\widetilde{\varphi} - \widetilde{g}_n\|_p\right)^p = 0$$

Demostrado ya el teorema en el caso $\Omega = \mathbb{R}^N$, lo probaremos ahora en el caso general, extendiendo las funciones de la manera más sencilla que se nos puede ocurrir. Fijada una clase de equivalencia en $L_p(\Omega)$, que será $\widetilde{f}$ con $f \in \mathcal{L}_p(\Omega)$, definimos una función $\varphi : \mathbb{R}^N \to \mathbb{R}$ escribiendo

$$\varphi(x) = f(x) \quad \forall x \in \Omega \qquad \text{y} \qquad \varphi(x) = 0 \quad \forall x \in \mathbb{R}^N \setminus \Omega$$

Vemos enseguida que φ es medible, pues dado un abierto $G \subset \mathbb{R}$, se tiene $\varphi^{-1}(G) = f^{-1}(G)$, cuando $0 \notin G$ mientras que $\varphi^{-1}(G) = f^{-1}(G) \bigcup (\mathbb{R}^N \setminus \Omega)$ cuando $0 \in G$, luego en ambos casos $\varphi^{-1}(G)$ es un conjunto medible.

Observamos ahora que $\varphi \in \mathcal{L}_p(\mathbb{R}^N)$, ya que

$$\int_{\mathbb{R}^N} |\varphi|^p = \int_\Omega |\varphi|^p = \int_\Omega |f|^p < \infty$$

La primera parte de la demostración nos da una sucesión $\{g_n\}$ de funciones escalonadas que converge a φ c.p.d. en $\mathbb{R}^N$, y tal que $\{\widetilde{g_n}\}$ también converge a $\widetilde{\varphi}$ en $L_p(\mathbb{R}^N)$. Por una parte, como $f = \varphi\big|_\Omega$, es claro que $\{g_n\}$ converge a f c.p.d. en Ω, así como que

$$0 \leqslant \int_\Omega \big|f - g_n\big|^p = \int_\Omega \big|\varphi - g_n\big|^p \leqslant \int_{\mathbb{R}^N} \big|\varphi - g_n\big|^p \qquad \forall n \in \mathbb{N}$$

de donde deducimos claramente que $\lim\limits_{n\to\infty} \displaystyle\int_\Omega \big|f - g_n\big|^p = 0$, y hemos probado (10). Por otra, si para cada $n \in \mathbb{N}$ tomamos $u_n = g_n\big|_\Omega \in \mathcal{E}(\Omega)$, vemos que $\{\widetilde{u_n}\}$ converge a $\widetilde{f}$ en $L_p(\Omega)$. Esto prueba que $E(\Omega)$ es denso en $L_p(\Omega)$, concluyendo la demostración. ■

Comparando con el primer teorema de densidad, cabría preguntarse si se puede conseguir que la sucesión $\{g_n\}$ del enunciado converja a f puntualmente en Ω, y no sólo c.p.d. en dicho conjunto. Aunque no vamos a demostrarlo, podemos decir que, incluso en el caso $\Omega = \mathbb{R}^N$ la respuesta puede ser negativa. De hecho, si $E \in \mathcal{M}$ con $\lambda(E) < \infty$, pero E no es un conjunto de Borel, entonces $\chi_E \in \mathcal{L}_1(\mathbb{R}^N)$, pero no existe una sucesión de funciones escalonadas que converja a χ_E puntualmente en $\mathbb{R}^N$.

Conviene resaltar que una función escalonada en un conjunto medible $\Omega \subset \mathbb{R}^N$ involucra intervalos acotados que pueden no estar contenidos en Ω. Para usar sólo subconjuntos de Ω, podemos pensar que, para $J \in \mathcal{J}$, las funciones características de J y $J \cap \Omega$ coinciden en Ω. Por tanto, toda función escalonada en Ω es restricción a Ω de una combinación lineal de funciones características de conjuntos de la forma $J \cap \Omega$ con $J \in \mathcal{J}$. Dependiendo de Ω, el conjunto $J \cap \Omega$ puede no ser tan sencillo como J. Así pues, en general, las funciones escalonadas en Ω no son tan intuitivas como en el caso $\Omega = \mathbb{R}^N$.

En ciertos casos, el problema anterior no se presenta. Por ejemplo, si $\Omega = I$ es un intervalo, no necesariamente acotado, está claro que $J \cap I \in \mathcal{J}$ para todo $J \in \mathcal{J}$, con lo que toda función escalonada en I es restricción a I de una combinación lineal de funciones características de intervalos acotados *contenidos* en I.

Cabe también mencionar el caso en que $\Omega = U$ es un abierto de $\mathbb{R}^N$. El razonamiento usado en el teorema anterior para el caso $\Omega = \mathbb{R}^N$ puede hacerse para U, de forma que las funciones escalonadas que aparecen sean combinaciones lineales de funciones características de intervalos acotados *contenidos* en U. Se obtiene que las clases de equivalencia, que contienen una función escalonada de este tipo, forman ya un subespacio denso en $L_p(\Omega)$.

8.7. Funciones continuas de soporte compacto

Las funciones escalonadas son ciertamente muy sencillas, pero carecen de otras propiedades deseables. Por ejemplo, una función escalonada, no idénticamente nula, nunca es continua. Vamos a probar un último teorema de densidad, cuyo interés radica precisamente en que permite aproximar por funciones continuas. En vez de trabajar en un conjunto medible arbitrario Ω, esta vez nos limitamos al caso en que el teorema tiene más interés, suponiendo que Ω es abierto. Recordemos que toda función continua $f : \Omega \to \mathbb{R}$ es medible, y enseguida presentamos una condición natural, de la que se deduce que $f \in \mathcal{L}_p(\Omega)$, para todo $p \in \mathbb{R}$ con $p \geqslant 1$.

Se define el **soporte** de una función $f : \Omega \to \mathbb{R}$ como el cierre, relativo a Ω, del conjunto de puntos en los que f no se anula:

$$\operatorname{sop} f = \overline{\{x \in \Omega : f(x) \neq 0\}} \subset \Omega$$

Como su nombre indica, una **función continua de soporte compacto**, no es más que una función continua $f : \Omega \to \mathbb{R}$, tal que $\operatorname{sop} f$ es un subconjunto compacto de Ω. Denotaremos por $C_c(\Omega)$ al conjunto de tales funciones que, como se ha dicho, está contenido en $\mathcal{L}(\Omega)$. De hecho, para $g, h \in C_c(\Omega)$ y $\alpha \in \mathbb{R}$, la función $\alpha g + h$ es continua, y su soporte es un conjunto cerrado, contenido en el conjunto compacto $\operatorname{sop} g \bigcup \operatorname{sop} h$, luego $\alpha g + h \in C_c(\Omega)$. Vemos por tanto que $C_c(\Omega)$ es subespacio vectorial de $\mathcal{L}(\Omega)$, pero podemos decir mucho más.

Fijado $p \in \mathbb{R}$ con $p \geqslant 1$, si $f \in C_c(\Omega)$, $K = \operatorname{sop} f$ y $M = \max\{|f(x)| : x \in K\}$, se tiene:

$$\int_\Omega |f|^p = \int_K |f|^p \leqslant M^p \lambda(K) < \infty$$

Así pues, $C_c(\Omega)$ es de hecho subespacio vectorial de $\mathcal{L}_p(\Omega)$. Además, si $f, g : \Omega \to \mathbb{R}$ son dos funciones continuas, tales que $f = g$ c.p.d., se ha de tener $f = g$. En efecto, como el conjunto abierto $E = \{x \in \Omega : f(x) \neq g(x)\}$ verifica que $\lambda(E) = 0$, deducimos que $E = \emptyset$. Esto implica que la aplicación cociente $f \mapsto \widetilde{f}$, de $\mathcal{L}_p(\Omega)$ sobre $L_p(\Omega)$, es inyectiva en $C_c(\Omega)$, lo que permite identificar $C_c(\Omega)$ con su imagen por dicha aplicación, para considerarlo como subespacio de $L_p(\Omega)$. El teorema de densidad que buscamos se puede ya adivinar, pero para obtenerlo necesitamos algunos preparativos. En primer lugar, probaremos la versión elemental, para espacios métricos, de un resultado topológico, que se conoce como lema de Urysohn:

- *Si A_0 y A_1 son subconjuntos cerrados, no vacíos y disjuntos de un espacio métrico X, existe una función continua $h : X \to [0,1]$, tal que $h(x) = 0$ para todo $x \in A_0$ y $h(x) = 1$ para todo $x \in A_1$.*

Si d es la distancia del espacio X, dado un conjunto no vacío $A \subset X$, para cada $x \in X$, sea también $d(x,A) = \inf\{d(x,a) : a \in A\}$, la distancia del punto x al conjunto A. Es sabido que la función $x \mapsto d(x,A)$ es continua, y de hecho es fácil comprobar que, para cualesquiera $x, y \in X$, se tiene: $\big|d(x,A) - d(y,A)\big| \leqslant d(x,y)$. También es claro que $\overline{A} = \{x \in X : d(x,A) = 0\}$.

Como A_0 y A_1 son cerrados y disjuntos, tenemos $d(x,A_0) + d(x,A_1) > 0$ para todo $x \in X$, lo que nos permite definir: $h(x) = \dfrac{d(x,A_0)}{d(x,A_0) + d(x,A_1)}$ para todo $x \in X$. Es evidente que h cumple todas las condiciones requeridas. ■

En el caso de $\mathbb{R}^N$, el resultado anterior permite obtener abundantes funciones continuas de soporte compacto, como veremos enseguida.

- *Si U es un abierto de $\mathbb{R}^N$ y K un subconjunto compacto no vacío de U, entonces existe una función $h \in C_c(\mathbb{R}^N)$ verificando que*

$$h(\mathbb{R}^N) \subset [0,1], \quad h(x) = 1 \ \ \forall x \in K, \quad \operatorname{sop} h \subset U \tag{11}$$

Podemos suponer que $U \neq \mathbb{R}^N$, pues en otro caso la inclusión $\operatorname{sop} h \subset U$ es obvia. Usando la distancia euclídea en $\mathbb{R}^N$, como $\mathbb{R}^N \setminus U$ es cerrado y disjunto de K, se tiene $d(x, \mathbb{R}^N \setminus U) > 0$ para todo $x \in K$. Por ser K compacto, la función continua $x \to d(x, \mathbb{R}^N \setminus U)$ tiene un valor mínimo en K, que será un número positivo δ, y fijamos $\rho \in \mathbb{R}$ con $0 < \rho < \delta$. Entonces, para cualesquiera $z \in \mathbb{R}^N \setminus U$ y $x \in K$, se tiene $d(z,x) \geqslant \delta$, luego $d(z,K) \geqslant \delta > \rho$.

Los conjuntos $A_0 = \{z \in \mathbb{R}^N : d(z,K) \geqslant \rho\}$ y $A_1 = K$, son cerrados, no vacíos y disjuntos, luego existe una función continua $h : \mathbb{R}^N \to [0,1]$, con $h(z) = 0$ para todo $z \in A_0$ y $h(x) = 1$ para todo $x \in K$. Si $y \in \mathbb{R}^N$ verifica que $h(y) \neq 0$, se tiene $d(y,K) < \rho$, de donde deducimos que $d(u,K) \leqslant \rho$ para todo $u \in \operatorname{sop} h$. Vemos así que el conjunto cerrado $\operatorname{sop} h$ también está acotado, luego es compacto, así que $h \in C_c(\mathbb{R}^N)$. Finalmente, para $z \in \mathbb{R}^N \setminus U$ habíamos visto que $d(z,K) > \rho$, de donde ahora obtenemos que $z \notin \operatorname{sop} h$. Tenemos así que $\operatorname{sop} h \subset U$, con lo que h verifica todo lo requerido en (11). ■

Estamos ya preparados para un último teorema de densidad.

Densidad de las funciones continuas de soporte compacto. *Si Ω es un subconjunto abierto no vacío de $\mathbb{R}^N$, y $p \in \mathbb{R}$ con $p \geqslant 1$, entonces $C_c(\Omega)$ es denso en $L_p(\Omega)$. Como consecuencia, para cada $f \in \mathcal{L}_p(\Omega)$, existe una sucesión $\{g_n\}$ en $C_c(\Omega)$, que verifica:*

$$\{g_n(x)\} \to f(x) \ \text{ p.c.t. } x \in \Omega \qquad \text{y} \qquad \lim_{n\to\infty} \int_\Omega |f - g_n|^p = 0$$

Demostración. Como $C_c(\Omega)$ es subespacio vectorial de $L_p(\Omega)$, igual le ocurre a su cierre, que denotamos por X_p, para probar que $X_p = L_p(\Omega)$. Dado $E \in \mathcal{M} \cap \mathcal{P}(\Omega)$ con $\lambda(E) < \infty$, consideremos la restricción a Ω de la función característica χ_E, a la que denotamos por χ. Está claro que $\chi \in \mathcal{L}_p(\Omega)$, y la clave de la demostración consiste en comprobar que $\widetilde{\chi} \in X_p$. Para ello, fijado $\varepsilon > 0$, debemos encontrar $g \in C_c(\Omega)$ tal que $\|\widetilde{\chi} - \widetilde{g}\|_p < \varepsilon$.

Por la regularidad interior y exterior de la medida de Lebesgue tenemos $K \subset E \subset V \subset \mathbb{R}^N$, donde K es compacto, V es abierto y $\lambda(V \setminus K) < \varepsilon^p$. Tomando $U = V \cap \Omega$, tenemos otro abierto $U \subset \mathbb{R}^N$ con $E \subset U$ y $\lambda(U \setminus K) \leqslant \lambda(V \setminus K) < \varepsilon^p$, pero ahora con $U \subset \Omega$. El resultado anterior nos da una función $h \in C_c(\mathbb{R}^N)$ que verifica (11). Tomando $g = h\big|_\Omega$, vemos que g es continua con $\operatorname{sop} g = \operatorname{sop} h \subset U \subset \Omega$, luego $g \in C_c(\Omega)$, y vamos a comparar g con χ.

Para $x \in K$ se tiene $g(x) = 1 = \chi(x)$, mientras si $x \in \Omega \setminus U$, vemos que $g(x) = 0 = \chi(x)$.Por último, para $x \in U \setminus K$ se tiene $g(x), \chi(x) \in [0,1]$, luego $|\chi(x) - g(x)| \leqslant 1$. Deducimos que

$$\int_\Omega |\chi - g|^p = \int_{U\setminus K} |\chi - g|^p \leqslant \lambda(U \setminus K) < \varepsilon^p$$

En términos de clases de equivalencia, tenemos $\|\widetilde{\chi} - \widetilde{g}\|_p < \varepsilon$, y hemos probado que $\widetilde{\chi} \in X_p$.

Toda función simple integrable $\varphi \in \mathcal{S}(\Omega)$ es combinación lineal de funciones del mismo tipo que χ, luego al ser X_p un subespacio vectorial de $L_p(\Omega)$, de lo ya demostrado deducimos que $\widetilde{\varphi} \in X_p$. Tenemos por tanto $S(\Omega) \subset X_p$. El primer teorema de densidad nos dice que $S(\Omega)$ es denso en $L_p(\Omega)$, luego X_p también lo es, pero X_p es cerrado, así que $X_p = L_p(\Omega)$.

Finalmente, dada $f \in \mathcal{L}_p(\Omega)$, existe una sucesión $\{h_n\}$ en $C_c(\Omega)$ tal que $\{\widetilde{h_n}\}$ converge a $\widetilde{f}$ en $L_p(\Omega)$. Entonces $\{h_n\}$ tiene una sucesión parcial $\{g_n\}$ que converge a f c.p.d. en Ω. Como $\{\widetilde{g_n}\}$ también converge a $\widetilde{f}$ en $L_p(\Omega)$, concluimos que

$$\lim_{n\to\infty} \int_{\mathbb{R}^N} |f - g_n|^p = \lim_{n\to\infty} \left(\| \widetilde{f} - \widetilde{g_n} \|_p \right)^p = 0 \qquad \blacksquare$$

8.8. La integral de Riemann

Vamos a comprobar que la integral de Lebesgue engloba y generaliza varias nociones de integral que se habían estudiado previamente, a todo lo largo del siglo XIX, partiendo de la integral de Cauchy para funciones continuas en un intervalo compacto, y culminando con la de Riemann, para ciertas funciones acotadas en un intervalo compacto.

Usaremos una definición de la integral de Riemann, equivalente a la original, que hace más fácil relacionarla con la de Lebesgue, así como caracterizar las funciones integrables en el sentido de Riemann. Fijamos un intervalo compacto $I \subset \mathbb{R}^N$ y, para evitar casos triviales, suponemos que I tiene interior no vacío: $I^\circ \neq \emptyset$.

Consideremos una **subdivisión** del intervalo I, esto es, una familia finita P de intervalos compactos, cuya unión es I, y cuyos interiores son dos a dos disjuntos. Usando en $\mathbb{R}^N$ la norma euclídea, definimos $\delta(P) \in \mathbb{R}^+$ como el máximo de los diámetros de los intervalos que forman la subdivisión P, es decir,

$$\delta(P) = \max\{\operatorname{diam} J : J \in P\} = \max\{\max\{\|x-y\| : x,y \in J\} : J \in P\}$$

Es claro que, para cada $\varepsilon > 0$, podemos tomar P de forma que $\delta(P) < \varepsilon$.

Si ahora $f : I \to \mathbb{R}$ es una función acotada, definimos la **suma inferior** y la **suma superior** de f, para cada subdivisión P del intervalo I, mediante las igualdades

$$I(f,P) = \sum_{J\in P} \big(\inf f(J)\big)\,\lambda(J) \qquad \text{y} \qquad S(f,P) = \sum_{J\in P} \big(\sup f(J)\big)\,\lambda(J)$$

donde sólo aparece la medida elemental de los intervalos acotados.

Pues bien, la función f es **Riemann-integrable**, cuando existe $\mathcal{R}(f) \in \mathbb{R}$ con la siguiente propiedad: para toda sucesión $\{P_n\}$ de subdivisiones del intervalo I, tal que $\{\delta(P_n)\} \to 0$, se verifica que

$$\lim_{n\to\infty} I(f,P_n) = \mathcal{R}(f) = \lim_{n\to\infty} S(f,P_n)$$

Como tales sucesiones $\{P_n\}$ existen, $\mathcal{R}(f)$ es único y se le llama **integral de Riemann** de f. Veremos que esta integral es caso particular de la de Lebesgue, pero probando mucho más.

Teorema. *Sea I un intervalo compacto en $\mathbb{R}^N$ con interior no vacío, $f : I \to \mathbb{R}$ una función acotada y $D(f)$ el conjunto de las discontinuidades de f, es decir, el conjunto de puntos de I en los que f no es continua. Entonces $D(f) \in \mathcal{M}$ y las siguientes afirmaciones son equivalentes:*

(i) *f es Riemann-integrable*

(ii) $\lambda\big(D(f)\big) = 0$

Si se cumplen (i) y (ii), entonces $f \in \mathcal{L}_1(I)$, con

$$\int_I f = \mathcal{R}(f)$$

Demostración. Usaremos una técnica muy habitual cuando se quiere estudiar el conjunto de las discontinuidades de una función real.

Para cada $x \in I$ y cada $\delta \in \mathbb{R}^+$, escribimos

$$m(\delta,x) = \inf f\big(B(x,\delta)\cap I\big) \qquad \text{y} \qquad M(\delta,x) = \sup f\big(B(x,\delta)\cap I\big)$$

donde $B(x,\delta)$ es la bola abierta euclídea, de centro x y radio δ. Para $0 < \rho < \delta$, se tiene

$$m(\delta,x) \leqslant m(\rho,x) \leqslant f(x) \leqslant M(\rho,x) \leqslant M(\delta,x)$$

Las desigualdades anteriores permiten claramente definir $m(x), M(x) \in \mathbb{R}$, escribiendo

$$m(x) = \sup\big\{ m(\delta,x) : \delta \in \mathbb{R}^+ \big\} = \lim_{\delta\to 0} m(\delta,x) \leqslant f(x)$$
$$M(x) = \inf\big\{ M(\delta,x) : \delta \in \mathbb{R}^+ \big\} = \lim_{\delta\to 0} M(\delta,x) \geqslant f(x)$$

y vamos a comprobar que f es continua en el punto x si, y sólo si, $m(x) = M(x)$.

Si f es continua en x y $\varepsilon > 0$, existe $\delta > 0$ tal que $|f(y) - f(x)| < \varepsilon$ para todo $y \in B(x,\delta)$, de donde $M(x) - m(x) \leqslant M(\delta,x) - m(\delta,x) \leqslant 2\varepsilon$, y la arbitrariedad de ε nos da $m(x) = M(x)$. Recíprocamente, dado $\varepsilon > 0$, la igualdad $m(x) = f(x) = M(x)$ nos permite encontrar $\delta > 0$ verificando que $m(\delta,x) > f(x) - \varepsilon$ y $M(\delta,x) < f(x) + \varepsilon$. Para todo $y \in B(x,\delta)$ se tiene entonces que $f(x) - \varepsilon < m(\delta,x) \leqslant f(y) \leqslant M(\delta,x) < f(x) + \varepsilon$, luego f es continua en x.

Sigamos ahora los pasos que llevaron a la integral de Riemann. Conviene aclarar que sólo usaremos funciones definidas en I, por lo que todas las funciones características que aparezcan se entenderán restringidas a I. A cada sucesión $\{P_n\}$ de subdivisiones de I, con $\{\delta(P_n)\} \to 0$, asociamos dos sucesiones $\{\varphi_n\}$ y $\{\psi_n\}$ de funciones escalonadas en I. Concretamente, para cada $n \in \mathbb{N}$, definimos:

$$\varphi_n = \sum_{J\in P_n} \big(\inf f(J)\big)\,\chi_J \qquad \text{y} \qquad \psi_n = \sum_{J\in P_n} \big(\sup f(J)\big)\,\chi_J$$

Para todo $n \in \mathbb{N}$ se tiene evidentemente

$$\int_I \varphi_n = I(f,P_n) \qquad \text{y} \qquad \int_I \psi_n = S(f,P_n) \tag{12}$$

Por tanto, el hecho de que f sea Riemann-integrable debe guardar relación con las dos sucesiones de funciones escalonadas, recién definidas. La clave para establecer dicha relación estará en probar que

$$\{\varphi_n(x)\} \to m(x) \quad \text{p.c.t. } x \in I \qquad \text{y} \qquad \{\psi_n(x)\} \to M(x) \quad \text{p.c.t. } x \in I$$

Para cada $n \in \mathbb{N}$, sea $A_n = \bigcup\limits_{J \in P_n} (J \setminus J^\circ)$, un conjunto de medida nula. Entonces $A = \bigcup\limits_{n=1}^{\infty} A_n$ también tiene medida nula. Tomando $B = I \setminus A \in \mathcal{M}$, veremos que

$$\{\varphi_n(x)\} \to m(x) \qquad \text{y} \qquad \{\psi_n(x)\} \to M(x) \qquad \forall x \in B \tag{13}$$

y de paso probaremos una útil acotación: si $\alpha = \sup\{|f(x)| : x \in I\}$, se tiene

$$|\varphi_n(x)| \leqslant \alpha \qquad \text{y} \qquad |\psi_n(x)| \leqslant \alpha \qquad \forall x \in B, \ \forall n \in \mathbb{N} \tag{14}$$

Fijados $x \in B$ y $n \in \mathbb{N}$, existe un intervalo $J_n \in P_n$ tal que $x \in J_n$, y como $x \notin A_n$, se ha de tener $x \in J_n^\circ$. Para $J \in P_n$ con $J \neq J_n$, se tiene que $x \notin J$, pues en otro caso, usando otra vez que $x \notin A_n$, se tendría $x \in J^\circ \cap J_n^\circ = \emptyset$, una clara contradicción. Vemos ahora claramente que

$$\varphi_n(x) = \inf f(J_n) \qquad \text{y} \qquad \psi_n(x) = \sup f(J_n) \tag{15}$$

y tenemos (14), ya que $-\alpha \leqslant \inf f(J_n) \leqslant \sup f(J_n) \leqslant \alpha$.

Con vistas a (13), tomamos $\delta > 0$ de forma que

$$m(x) - \varepsilon < m(\delta, x) \leqslant M(\delta, x) < M(x) + \varepsilon$$

Existe $n_0 \in \mathbb{N}$ tal que, para $n \geqslant n_0$ se tiene $\delta(P_n) < \delta$. Entonces, para $n \geqslant n_0$ tenemos como antes $J_n \in P_n$, tal que $x \in J_n^\circ$ y se verifica (15). Tomamos $\rho > 0$ con $B(x,\rho) \subset J_n$ y, como J_n tiene diámetro menor que δ, también tenemos $J_n \subset B(x,\delta)$. Deducimos claramente que

$$\begin{aligned} &m(x) - \varepsilon < m(\delta,x) \leqslant \inf f(J_n) = \varphi_n(x) \leqslant m(\rho,x) \leqslant m(x) \qquad \text{y} \\ &M(x) \leqslant M(\rho,x) \leqslant \sup f(J_n) = \psi_n(x) \leqslant M(\delta,x) < M(x) + \varepsilon \end{aligned}$$

Por tanto, $|\varphi_n(x) - m(x)| < \varepsilon$ y $|\psi_n(x) - M(x)| < \varepsilon$ para $n \geqslant n_0$, lo que demuestra (13).

Seguidamente probamos que las funciones m y M son medibles, y de hecho $m, M \in \mathcal{L}_1(I)$. En (13) vemos que $\{\chi_B \psi_n\}$ converge a $\chi_B M$ puntualmente en I, luego $\chi_B M$ es medible, pero es claro que $M = \chi_B M$ c.p.d. en I, luego M también es medible. Con análogo razonamiento, vemos que m es medible. De hecho, como m y M son funciones acotadas y $\lambda(I) < \infty$, vemos que $m, M \in \mathcal{L}_1(I)$, y vamos a calcular las correspondientes integrales.

La sucesión de funciones medibles $\{\chi_B \psi_n\}$ converge a $\chi_B M$ puntualmente en I, y en (14) tenemos $|\chi_B(x)\psi_n(x)| \leqslant \alpha$ para cualesquiera $x \in I$ y $n \in \mathbb{N}$. Al ser $\lambda(I) < \infty$, podemos usar el teorema de la convergencia dominada, y como además $\chi_B(x) = 1$ p.c.t. $x \in I$, obtenemos:

$$\int_I M = \int_I \chi_B M = \lim_{n\to\infty} \int_I \chi_B \psi_n = \lim_{n\to\infty} \int_I \psi_n$$

Análogo razonamiento al usado con M puede emplearse con m para obtener que

$$\int_I m = \lim_{n\to\infty} \int_I \psi_n$$

Recordando ahora las igualdades (12), tenemos finalmente que

$$\int_I m = \lim_{n\to\infty} I(f,P_n) \qquad \text{y} \qquad \int_I M = \lim_{n\to\infty} S(f,P_n) \tag{16}$$

igualdades válidas para toda sucesión $\{P_n\}$ de subdivisiones de I, con $\{\delta(P_n)\} \to 0$.

Las afirmaciones del enunciado se obtienen ya con facilidad. Como m y M son funciones medibles, vemos que el conjunto de las discontinuidades de f es medible:

$$D(f) = \{x \in I : M(x) - m(x) > 0\} \in \mathcal{M}$$

Si f es Riemann-integrable, vemos en (16) que la integral sobre I de la función medible positiva $M - m$ se anula, luego $m = M$ c.p.d. en I, es decir: $\lambda\big(D(f)\big) = 0$. Recíprocamente, de $\lambda\big(D(f)\big) = 0$, deducimos que $m = M$ c.p.d. en I, con lo que (16) nos dice que f es Riemann-integrable con

$$\int_I m = \mathcal{R}(f) = \int_I M$$

Tenemos así la equivalencia entre las afirmaciones (i) y (ii) del enunciado. Cuando ambas se cumplen, usamos por ejemplo que $f = m$ c.p.d. Como m es medible, deducimos que f también lo es, y de hecho tenemos $f \in \mathcal{L}_1(I)$, con

$$\int_I f = \int_I m = \mathcal{R}(f) \qquad \blacksquare$$

Obviamente, si en el teorema anterior suponemos que f es continua, obtenemos que f es Riemann-integrable. En este caso particular, la integral de Riemann es la que inicialmente había definido Cauchy. En realidad, la aportación de Riemann consistió principalmente en considerar la familia de funciones a las que se puede aplicar el mismo método que Cauchy había usado. Caracterizar estas funciones, en términos del conjunto de sus discontinuidades, era el problema abierto que resolvió Lebesgue con el teorema anterior. Pero al mismo tiempo, este resultado hizo que la integral de Riemann dejase de tener interés, al quedar como caso muy particular de la integral de Lebesgue.