Tema 9

Teorema de derivación de Lebesgue

La relación entre derivada e integral, para funciones de una variable real, se describe en el llamado *Teorema Fundamental del Cálculo*, cuyo nombre ya indica su importancia. Como paso previo para obtener la versión general de este teorema, probamos ahora otro resultado no menos importante, que se conoce como *Teorema de derivación de Lebesgue*. Afirma en principio que toda función real creciente, definida en un intervalo compacto, es derivable casi por doquier en dicho intervalo. Obviamente, el resultado también es válido para funciones decrecientes, y de hecho lo extenderemos a un tipo de funciones más general que las monótonas, que son las llamadas *funciones de variación acotada*. En realidad, la demostración se reduce al caso de una función creciente, pues como veremos, toda función de variación acotada en un intervalo compacto, se puede expresar como diferencia de dos funciones crecientes. También será fácil extender el teorema al caso de un intervalo no trivial, que puede no ser compacto.

9.1. Funciones de variación acotada

En lo que sigue, trabajamos en un intervalo compacto $[a, b] \subset \mathbb{R}$ con $a < b$. Empezamos observando una útil propiedad de las funciones crecientes.

- *Si $f : [a, b] \to \mathbb{R}$ es una función creciente, entonces f tiene límite en los puntos a y b, y tiene límites laterales en todo punto $c \in]a, b[$. Como consecuencia, f sólo puede tener un conjunto numerable de discontinuidades, todas ellas evitables o de salto.*

Fijado $c \in]a, b]$, para $a \leqslant x < c$ tenemos $f(x) \leqslant f(c)$, luego podemos escribir:

$$L(c) \stackrel{\text{def}}{=} \sup \{ f(x) \, : \, a \leqslant x < c \} \leqslant f(c)$$

Dado $\varepsilon > 0$ existirá un $y \in [a, c[$ tal que $f(y) > L(c) - \varepsilon$. Tomando $\delta = c - y > 0$, para todo $x \in]c - \delta, c[=]y, c[$ tenemos $L(c) - \varepsilon < f(y) \leqslant f(x) \leqslant L(c)$, luego $|f(x) - L(c)| < \varepsilon$. Cuando $a < c < b$, esto prueba que $L(c)$ es el límite por la izquierda de f en c, mientras que tomando $c = b$, vemos que $L(b)$ es el límite (ordinario) de f en b.

Análogamente, fijado $c \in [a, b[$ escribimos

$$R(c) \stackrel{\text{def}}{=} \text{ínf}\left\{ f(x) \, : \, c < x \leqslant b] \right\} \geqslant f(c)$$

Dado $\varepsilon > 0$ existirá un $z \in]c, b]$ tal que $f(z) < R(c) + \varepsilon$, y tomando $\delta = z - a > 0$, para todo $x \in]c, c+\delta[=]c, z[$ tenemos $R(c) \leqslant f(x) \leqslant f(z) < R(c) + \varepsilon$, luego $|f(x) - R(c)| < \varepsilon$. Cuando $a < x < b$, esto prueba que $R(c)$ es el límite por la derecha de f en c, y para $c = a$, hemos visto que $R(a)$ es el límite de f en a. En resumen, tenemos:

$$L(c) = \lim_{x \nearrow c} f(x) \leqslant f(c) \leqslant \lim_{x \searrow c} f(x) = R(c) \qquad \forall c \in]a, b[$$

mientras que

$$f(a) \leqslant R(a) = \lim_{x \to a} f(x) \qquad \text{y} \qquad L(b) = \lim_{x \to b} f(x) \leqslant f(b)$$

Sea ahora E el conjunto de las discontinuidades de f. Para cada $c \in E \cap]a, b[$, se tiene obligadamente $L(c) < R(c)$, de modo que c es una discontinuidad de salto. Además, podemos tomar $\varphi(c) \in \mathbb{Q}$ con $L(c) < \varphi(c) < R(c)$. Si $a \in E$, la discontinuidad de f en a es evitable y tomamos $\varphi(a) \in \mathbb{Q}$ con $f(a) < \varphi(a) < R(a)$. Análogamente, si $b \in E$, la discontinuidad será evitable y tomamos $\varphi(b) \in \mathbb{Q}$ con $L(b) < \varphi(b) < f(b)$. De esta forma hemos definido una aplicación $\varphi : E \to \mathbb{Q}$. Para $c, d \in E$ con $c < d$, tomando $x \in]c, d[$ tenemos claramente

$$\varphi(c) < R(c) \leqslant f(x) \leqslant L(d) < \varphi(d)$$

lo que demuestra que φ es estrictamente creciente, luego inyectiva. Por tanto, E es numerable, al ser equipotente a un subconjunto de $\mathbb{Q}$. ■

Es obvio que una función decreciente $f : [a, b] \to \mathbb{R}$ también verifica las afirmaciones del enunciado anterior, pues lo ya probado se puede usar para la función creciente $-f$, que tiene límites en los mismos puntos que f y también tiene las mismas discontinuidades. Por tanto, el resultado es válido para funciones monótonas, pero llegaremos más lejos. Si $f, g : [a, b] \to \mathbb{R}$ son funciones crecientes, la diferencia $f - g$ verifica la tesis del resultado anterior, pero es fácil ver que $f - g$ puede no ser una función monótona. Es natural por tanto, preguntarse por las funciones que pueden expresarse como diferencia de dos crecientes. Es claro que, tales funciones forman un espacio vectorial, el engendrado por las funciones crecientes. Pues bien, vamos a estudiar ahora una propiedad que caracteriza a este tipo de funciones.

Llamaremos **partición** del intervalo $[a, b]$ a todo conjunto finito $P \subset [a, b]$ con $a, b \in P$, y denotaremos por $\Pi(a, b)$ al conjunto de tales particiones. Si $P \in \Pi(a, b)$ tiene $n+1$ elementos, con $n \in \mathbb{N}$, solemos numerarlos de menor a mayor, escribiendo $P = \{ a = x_0 < x_1 \ldots < x_n = b \}$. Fijada dicha partición, a cada función $f : J \to \mathbb{R}$, definida en un intervalo $J \subset \mathbb{R}$ con $a, b \in J$, podemos asociar la suma dada por

$$\sigma(f, P) = \sum_{k=1}^{n} \left| f(x_k) - f(x_{k-1} \right|$$

Obsérvese que esta suma cuantifica la "variación" que experimenta la función f al recorrer los puntos de P, pero en términos absolutos, sin distinguir entre aumentos y disminuciones.

Llamaremos **variación total** de f en $[a,b]$ al supremo de las sumas del tipo anterior, que se obtienen para todas las particiones del intervalo $[a,b]$, a la que se denota por

$$V(f;a,b) = \sup\big\{\sigma(f,P) : P \in \Pi(a,b)\big\} \in [0,\infty]$$

Decimos que f tiene **variación acotada** en $[a,b]$ cuando $V(f;a,b) < \infty$. Resaltamos que esta definición tiene sentido para funciones definidas en cualquier intervalo J tal que $a,b \in J$. En el caso $J = [a,b]$, cuando una función $f : [a,b] \to \mathbb{R}$ tiene variación acotada en $[a,b]$, se dice simplemente que f es una **función de variación acotada**.

Dadas dos funciones $f,g : [a,b] \to \mathbb{R}$ y $\alpha \in \mathbb{R}$, para toda partición $P \in \Pi(a,b)$ se tiene claramente que $\sigma(\alpha f, P) = |\alpha|\sigma(f,P)$, mientras que $\sigma(f+g,P) \leqslant \sigma(f,P) + \sigma(g,P)$, de donde deducimos que

$$V(\alpha f;a,b) = |\alpha|V(f;a,b) \qquad \text{y} \qquad V(f+g;a,b) \leqslant V(f;a,b) + V(g;a,b)$$

Por tanto, si f y g son funciones de variación acotada, también lo son αf y $f+g$. Así pues, en el espacio vectorial de todas las funciones de $[a,b]$ en $\mathbb{R}$, las de variación acotada forman un subespacio vectorial.

Es fácil ver que toda función creciente $f : [a,b] \to \mathbb{R}$ es de variación acotada. De hecho, si una partición $P \in \Pi(a,b)$ viene dada por $P = \big\{a = x_0 < x_1 < \ldots < x_n = b\big\}$, se tiene

$$\sigma(f,P) = \sum_{k=1}^{n} \big| f(x_k) - f(x_{k-1}) \big| = \sum_{k=1}^{n} \big(f(x_k) - f(x_{k-1}) \big) = f(b) - f(a)$$

de donde deducimos que $V(f;a,b) = f(b) - f(a)$. Podemos ya caracterizar las funciones que se obtienen como diferencia de dos crecientes.

- *Una función de $[a,b]$ en $\mathbb{R}$ tiene variación acotada en $[a,b]$ si, y sólo si, se puede expresar como diferencia de dos funciones crecientes.*

Si $g,h : [a,b] \to \mathbb{R}$ son crecientes, hemos visto que ambas son de variación acotada, pero las funciones con esa propiedad forman un espacio vectorial, luego $g-h$ también es de variación acotada.

Recíprocamente, sea $f : [a,b] \to \mathbb{R}$ una función de variación acotada. Fijado $x \in]a,b[$, vamos a ver que f también tiene variación acotada en $[a,x]$. Dada una partición $P \in \Pi(a,x)$, tomamos $Q = P \cup \{b\} \in \Pi(a,b)$, y es claro que

$$\sigma(f,P) + \big| f(b) - f(x) \big| = \sigma(f,Q) \leqslant V(f;a,b)$$

Como la desigualdad anterior es válida para toda partición $P \in \Pi(a,x)$, obtenemos que

$$V(f;a,x) + \big| f(b) - f(x) \big| \leqslant V(f;a,b)$$

y en particular tenemos $V(f;a,x) \leqslant V(f;a,b) < \infty$. Para $a < x < y \leqslant b$, el razonamiento anterior puede ahora hacerse en el intervalo $[a,y]$, en vez de $[a,b]$, para obtener que

$$a < x < y \leqslant b \implies V(f;a,x) + \big| f(y) - f(x) \big| \leqslant V(f;a,y) \tag{1}$$

Pues bien, consideremos la función $g : [a, b] \to \mathbb{R}$ definida por

$$g(x) = V(f\,; a,x) \quad \forall x \in]a, b] \qquad \text{y} \qquad g(a) = 0$$

Si $a < x < y \leqslant b$, de (1) deducimos que $g(x) \leqslant g(y)$, desigualdad que es obvia para $x = a$, luego la función g es creciente.

Por otra parte, de nuevo para $a < x < y \leqslant b$, de (1) también deducimos que

$$g(x) + \big(f(y) - f(x)\big) \leqslant g(x) + \big|f(y) - f(x)\big| \leqslant g(y)$$

de donde $g(x) - f(x) \leqslant g(y) - f(y)$. Cuando $x = a$ se tiene claramente que

$$f(y) - f(a) \leqslant \big|f(y) - f(a)\big| \leqslant V(f\,; a,y) = g(y)$$

de donde $g(a) - f(a) = -f(a) \leqslant g(y) - f(y)$. Esto prueba que la función $g - f$ también es creciente, luego $f = g - (g - f)$ es diferencia de dos funciones crecientes. ■

El resultado sobre límites y continuidad de la funciones crecientes, demostrado al principio, se puede ya extender a las funciones de variación acotada:

- *Si $f : [a, b] \to \mathbb{R}$ es una función de variación acotada, entonces f tiene límite en los puntos a y b, y límites laterales en todo punto $c \in]a, b[$. Además, f sólo puede tener un conjunto numerable de discontinuidades, todas ellas evitables o de salto.*

Es claro que una función creciente puede no ser continua, luego una función de variación acotada puede no ser continua. En sentido opuesto, para una función continua $f : [a, b] \to \mathbb{R}$, cabe preguntarse si f tiene que ser de variación acotada. Como f es uniformemente continua, la siguiente observación parece apuntar hacia una respuesta afirmativa.

- *Toda función lipschitziana $f : [a, b] \to \mathbb{R}$ es de variación acotada.*

En efecto, si $M \in \mathbb{R}_0^+$ es la constante de Lipschitz de f y $P = \{a = x_0 < x_1 \ldots < x_n = b\}$ es una partición del intervalo $[a, b]$, se tiene:

$$\sum_{k=1}^{n} \big|f(x_k) - f(x_{k-1}\big| \leqslant M \sum_{k=1} < |x_k - x_{k-1}| = M \sum_{k=1} (x_k - x_{k-1}) = M(b-a)$$

luego $V(f\,; a,b) \leqslant M(b-a) < \infty$. ■

Sin embargo, la respuesta a la pregunta antes planteada es negativa:

Ejemplo. *Una función continua que no es de variación acotada.* Sea $f : [0, 1] \to \mathbb{R}$ la función definida por

$$f(x) = x \cos \frac{\pi}{x} \quad \forall x \in]0, 1] \qquad \text{y} \qquad f(0) = 0$$

Es claro que f es continua, pero vamos a comprobar que $V(f\,; 0,1) = \infty$.

Para ello, fijado $n \in \mathbb{N}$ con $n > 3$, usaremos la partición dada por

$$P = \left\{ 0 < \frac{1}{n+1} < \frac{1}{n} < \frac{1}{n-1} < \ldots < 1 \right\} \in \Pi(0,1)$$

Usando que $\cos k\pi = (-1)^k$ para todo $k \in \mathbb{N}$, vemos que, para $k \in \Delta_n$ se tiene

$$\left| f\left(\frac{1}{k}\right) - f\left(\frac{1}{k+1}\right) \right| = \left| \frac{(-1)^k}{k} - \frac{(-1)^{k+1}}{k+1} \right| = \frac{1}{k} + \frac{1}{k+1} \geqslant \frac{1}{k}$$

de donde deducimos claramente que

$$\sigma(f,P) = \left| f\left(\frac{1}{n+1}\right) - f(0) \right| + \sum_{k=1}^{n} \left| f\left(\frac{1}{k}\right) - f\left(\frac{1}{k+1}\right) \right| \geqslant \sum_{k=1}^{n} \frac{1}{k}$$

Por tanto, $V(f\,;0,1) \geqslant \sum_{k=1}^{n} \frac{1}{k}$ para todo $n \in \mathbb{N}$, luego $V(f\,;0,1) = \infty$. ■

9.2. Dos lemas previos

Nuestro objetivo es estudiar la derivabilidad de una función creciente, lo que requiere varios preparativos. En primer lugar probaremos una versión elemental de un resultado que se conoce como *lema del recubrimiento de Vitali*. Para una familia de intervalos abiertos, cuya unión sea un conjunto acotado, nos permite conseguir una subfamilia finita de intervalos dos a dos disjuntos, cuya suma de longitudes guarda una concreta relación con la medida de la unión de la familia de partida.

Resaltamos que, para una familia de conjuntos $\mathcal{U}$, denotamos simplemente por $\cup\mathcal{U}$ a la unión de todos los conjuntos que forman la familia $\mathcal{U}$, y como es natural, la unión de una familia $\mathcal{W}$ de conjuntos dos a dos disjuntos, se denota por $\uplus\mathcal{W}$. Trabajamos primero con una familia finita de intervalos abiertos.

- *Toda familia finita $\mathcal{F}$ de intervalos abiertos acotados en $\mathbb{R}$, contiene una subfamilia $\mathcal{W}$, formada por intervalos dos a dos disjuntos, que verifica:* $\lambda\big(\cup\,\mathcal{F}\big) \leqslant 3\lambda\big(\uplus\,\mathcal{W}\big)$.

Razonamos por inducción sobre el número m de elementos de $\mathcal{F}$, siendo obvio el resultado cuando $m = 1$. Dado $n \in \mathbb{N}$, suponemos que el lema es cierto para $m \leqslant n$ y lo demostramos en el caso $m = n+1$. Para ello, empezamos eligiendo un intervalo $I_0 \in \mathcal{F}$ que tenga longitud máxima, es decir, verificando que $\lambda(I_0) \geqslant \lambda(I)$ para todo $I \in \mathcal{F}$. Consideramos entonces la familia $\mathcal{F}_0 = \big\{ I \in \mathcal{F} : I \cap I_0 = \emptyset \big\}$, pudiéndose dar dos casos.

Si $\mathcal{F}_0 \neq \emptyset$, como $\mathcal{F}_0$ tiene a lo sumo n elementos, la hipótesis de inducción nos da una familia $\mathcal{W}_0$ de intervalos dos a dos disjuntos, con $\mathcal{W}_0 \subset \mathcal{F}_0$ y $\lambda\big(\cup\,\mathcal{U}_0\big) \leqslant 3\lambda\big(\uplus\,\mathcal{W}_0\big)$. Tomamos entonces $\mathcal{W} = \mathcal{W}_0 \cup \{I_0\}$. Como los intervalos de $\mathcal{W}_0$ son dos a dos disjuntos, y todos ellos son disjuntos de I_0, vemos que $\mathcal{W}$ también es una familia de intervalos dos a dos disjuntos, con $\mathcal{W} \subset \mathcal{F}$. Si $\mathcal{F}_0 = \emptyset$, tomamos simplemente $\mathcal{W} = \{I_0\}$.

Para probar que en ambos casos se tiene $\lambda\big(\cup\mathcal{F}\big) \leqslant 3\lambda\big(\uplus\,\mathcal{W}\big)$, nos basamos en una sencilla observación: dados $a\in\mathbb{R}$ y $r\in\mathbb{R}^+$, si un intervalo abierto I verifica que $I\cap\,]a-r,a+r[\,\neq\emptyset$, y $\lambda(I)\leqslant 2r$, entonces $I\subset\,]a-3r,a+3r[$. En efecto, tomando $x_0\in I\cap\,]a-r,a+r[$, para todo $x\in I$ se tiene $|x-a|\leqslant|x-x_0|+|x_0-a|<2r+r=3r$.

Escribimos ahora $I_0=\,]a-r,a+r[$ con $a\in\mathbb{R}$ y $r\in\mathbb{R}^+$. Dado un intervalo $I\in\mathcal{F}\setminus\mathcal{F}_0$, la definición de $\mathcal{F}_0$ nos dice que $I\cap I_0\neq\emptyset$, y por la elección de I_0 tenemos $\lambda(I)\leqslant\lambda(I_0)=2r$, luego $I\subset\,]a-3r,a+3r[$. Deducimos que

$$\cup(\mathcal{F}\setminus\mathcal{F}_0)\subset\,]a-3r,a+3r[,\qquad\text{luego}\qquad \lambda\big(\cup(\mathcal{F}\setminus\mathcal{F}_0)\big)\leqslant 6r=3\,\lambda(I_0)$$

En el caso $\mathcal{F}_0=\emptyset$ tenemos $\lambda\big(\cup\mathcal{F}\big)\leqslant 3\lambda(I_0)=3\lambda\big(\uplus\mathcal{W}\big)$, lo que concluye la demostración. Pero si $\mathcal{F}_0\neq\emptyset$, teniendo en cuenta que $I_0\cap\big(\uplus\,\mathcal{W}_0\big)=\emptyset$, obtenemos

$$\lambda\big(\cup\,\mathcal{F}\big)\leqslant\lambda\big(\cup(\mathcal{F}\setminus F_0)\big)+\lambda\big(\cup\,\mathcal{F}_0\big)\leqslant 3\lambda(I_0)+3\lambda\big(\uplus\,\mathcal{W}_0\big)=3\lambda\big(\uplus\,\mathcal{W}\big)$$

concluyendo igualmente la demostración. ■

Pasamos ahora al caso en que la familia de partida es arbitraria, salvo que su unión sigue siendo un conjunto acotado.

Lema 1. *Supongamos que un abierto acotado $G\subset\mathbb{R}$ es la unión de una familia $\mathcal{U}$ de intervalos abiertos. Entonces existe una subfamilia finita $\mathcal{W}\subset\mathcal{U}$, formada por intervalos dos a dos disjuntos, tal que $\lambda(G)\leqslant 4\lambda\big(\uplus\,\mathcal{W}\big)$.*

Demostración. Por la regularidad interior de la medida de Lebesgue, existe un conjunto compacto $K\subset G$, tal que $\lambda(K)\geqslant(3/4)\lambda(G)$. Entonces $\mathcal{U}$ es un recubrimiento de K por conjuntos abiertos, del que se podrá extraer un subrecubrimiento finito $\mathcal{F}\subset\mathcal{U}$. El resultado probado previamente nos da ahora una subfamilia $\mathcal{W}\subset\mathcal{F}\subset\mathcal{U}$, formada por intervalos dos a dos disjuntos, tal que $\lambda\big(\cup\mathcal{F}\big)\leqslant 3\lambda\big(\uplus\mathcal{W}\big)$. Esto concluye la demostración, ya que

$$\lambda(G)\leqslant(4/3)\lambda(K)\leqslant(4/3)\lambda\big(\cup\mathcal{F}\big)\leqslant 4\lambda\big(\uplus\,\mathcal{W}\big)$$ ■

Nuestro segundo lema previo cuantifica la variación de una función al recorrer una partición de un intervalo, conociendo la pendiente media de su gráfica en ciertos subintervalos.

Lema 2. *Dados $c,d\in\mathbb{R}$ con $c<d$, sea $P=\{c=y_0<y_1<\ldots y_m=d\}$ una partición del intervalo $[c,d]$, y fijemos un conjunto no vacío $S\subset\Delta_m$ y un $\delta\in\mathbb{R}^+$. Supongamos ahora que una función $g:[c,d]\to\mathbb{R}$ verifica una de las dos afirmaciones siguientes:*

$$\begin{array}{llcl}
(a) & g(c)\leqslant g(d) & y & g(y_j)-g(y_{j-1})<-\delta\,(y_j-y_{j-1})\quad\forall j\in S\\
(b) & g(c)\geqslant g(d) & y & g(y_j)-g(y_{j-1})>\delta\,(y_j-y_{j-1})\quad\forall j\in S
\end{array}$$

Se tiene entonces que

$$\sum_{j=1}^{n}\big|g(y_j)-g(y_{j-1})\big|\geqslant\big|g(d)-g(c)\big|+\delta\sum_{j\in S}(y_j-y_{j-1})\tag{2}$$

Demostración. Suponiendo que se verifica *(a)*, tenemos

$$\begin{aligned}\big|g(d)-g(c)\big| &= g(d)-g(c) = \sum_{k=1}^{n}\big(g(y_j)-g(x_{j-1})\big)\\ &= \sum_{j\in S}\big(g(y_j)-g(y_{j-1})\big) + \sum_{j\in\Delta_n\setminus S}\big(g(y_j)-g(y_{j-1})\big)\\ &< -\delta\sum_{j\in S}\big(y_j-y_{j-1}\big) + \sum_{j=1}^{n}\big|g(y_j)-g(y_{j-1})\big|\end{aligned}$$

de donde se obtiene claramente la desigualdad (2).

Si g verifica *(b)*, es claro que la función $-g$ cumple la condición *(a)*, luego por lo ya demostrado, $-g$ verifica la desigualdad (2), lo que equivale a que g la verifique. ■

9.3. Derivación de funciones crecientes

Para estudiar la derivabilidad de una función, usaremos una técnica similar a la empleada en otra ocasión para la continuidad. Para un intervalo no trivial $J\subset\mathbb{R}$, trabajamos con una función creciente $f: J\to\mathbb{R}$.

Fijados $x\in J$ y $r\in\mathbb{R}^+$, vemos que, para todo $y\in J\setminus\{x\}$, se tiene $(f(y)-f(x))/(y-x)\geqslant 0$, lo que nos permite definir

$$\begin{aligned}\underline{D}_r f(x) &= \inf\left\{\frac{f(y)-f(x)}{y-x} : y\in J,\ 0<|y-x|<r\right\}\in\mathbb{R}_0^+\\ \overline{D}_r f(x) &= \sup\left\{\frac{f(y)-f(x)}{y-x} : y\in J,\ 0<|y-x|<r\right\}\in[0,\infty]\end{aligned}$$

Para $0<r<s$ se tiene claramente que $\underline{D}_s f(x)\leqslant\underline{D}_r f(x)\leqslant\overline{D}_r f(x)\leqslant\overline{D}_s f(x)$, lo que nos permite ahora definir $\underline{D} f(x),\overline{D} f(x)\in[0,\infty]$, escribiendo

$$\begin{aligned}\underline{D} f(x) &= \lim_{r\to 0}\underline{D}_r f(x) = \sup\left\{\underline{D}_r f(x) : r\in\mathbb{R}^+\right\}\\ \overline{D} f(x) &= \lim_{r\to 0}\overline{D}_r f(x) = \inf\left\{\overline{D}_r f(x) : r\in\mathbb{R}^+\right\}\end{aligned}$$

Lo anterior puede hacerse para todo $x\in J$, obteniendo dos funciones $\underline{D} f,\overline{D} f: J\to[0,\infty]$, que caracterizan la derivabilidad de f, como vamos a comprobar.

- *Dado un intervalo no trivial $J\subset\mathbb{R}$, una función creciente $f: J\to\mathbb{R}$ es derivable en un punto $x\in J$ si, y sólo si, $\underline{D} f(x)=\overline{D} f(x)=\alpha<\infty$, en cuyo caso, $f'(x)=\alpha$.*

Si f es derivable en el punto x y escribimos $\alpha=f'(x)$, para todo $\varepsilon>0$ existe $\delta>0$ tal que

$$y\in J,\ 0<|y-x|<\delta \implies \alpha-\varepsilon<\frac{f(y)-f(x)}{y-x}<\alpha+\varepsilon$$

Entonces, para $0 < r < \delta$ se tiene claramente que $\alpha - \varepsilon \leqslant \underline{D}_r f(x) \leqslant \overline{D}_r f(x) \leqslant \alpha + \varepsilon$, de donde deducimos que $\alpha - \varepsilon \leqslant \underline{D} f(x) \leqslant \overline{D} f(x) \leqslant \alpha + \varepsilon$. Como $\varepsilon > 0$ era arbitrario, concluimos que $\underline{D} f(x) = \overline{D} f(x) = \alpha < \infty$.

Recíprocamente, supongamos que $\underline{D} f(x) = \overline{D} f(x) = \alpha \in \mathbb{R}_0^+$, para ver que $f'(x) = \alpha$. Dado $\varepsilon > 0$ podemos encontrar $\delta > 0$ de forma que, para $0 < r < \delta$ se tenga

$$\alpha - \varepsilon = \underline{D} f(x) - \varepsilon < \underline{D}_r f(x) \leqslant \overline{D}_r f(x) < \overline{D} f(x) + \varepsilon = \alpha + \varepsilon$$

Entonces, para $y \in J$ con $0 < |y - x| < \delta$, tomamos r con $|y - x| < r < \delta$ para obtener que

$$\alpha - \varepsilon < \underline{D}_r f(x) \leqslant \frac{f(y) - f(x)}{y - x} \leqslant \overline{D}_r f(x) < \alpha + \varepsilon$$

Esto prueba que $\lim\limits_{y \to x} \dfrac{f(y) - f(x)}{y - x} = \alpha$, como queríamos demostrar. ■

Entramos ya en la demostración del teorema principal que vamos buscando. La dividimos en dos etapas, que enunciaremos en forma de proposiciones.

Proposición 1. *Dados $a, b \in \mathbb{R}$ con $a < b$, toda función creciente $f : [a, b] \to \mathbb{R}$ verifica que $\underline{D} f(x) < \infty$ para casi todo $x \in [a, b]$.*

Demostración. Excluyendo por comodidad los puntos a y b, consideramos el conjunto

$$B = \{ x \in]a, b[\, : \, \underline{D} f(x) = \infty \}$$

y se trata de probar que B tiene medida nula.

Fijado $M \in \mathbb{R}^+$, para cada $x \in B$ tenemos $\lim\limits_{r \searrow 0} \underline{D}_r f(x) = \infty$, luego podemos tomar $r \in \mathbb{R}^+$, con $a < x - r < x + r < b$, verificando que $\underline{D}_r f(x) > M$, con lo que

$$x - r < y < x + r \quad \Longrightarrow \quad \frac{f(y) - f(x)}{y - x} > M$$

Tomando ahora $a_x, b_x \in \mathbb{R}$ con $x - r < a_x < x < b_x < x + r$, tenemos

$$\begin{aligned} \frac{f(b_x) - f(a_x)}{b_x - a_x} &= \frac{f(b_x) - f(x)}{b_x - x} \left(\frac{b_x - x}{b_x - a_x} \right) + \frac{f(x) - f(a_x)}{x - a_x} \left(\frac{x - a_x}{b_x - a_x} \right) \\ &> M \, \frac{b_x - x}{b_x - a_x} + M \, \frac{x - a_x}{b_x - a_x} = M \end{aligned} \tag{3}$$

desigualdad que es válida para todo $x \in B$.

Abreviamos escribiendo $I_x =]a_x, b_x[$ para cada $x \in B$, con lo que $\mathcal{U} = \{I_x : x \in B\}$ es una familia de intervalos abiertos, cuya unión es un abierto $G \subset]a, b[$. Usamos ahora el lema 1, para obtener una subfamilia finita $\mathcal{W} = \{I_x : x \in B_0\}$, donde B_0 es un subconjunto finito de B, que está formada por intervalos dos a dos disjuntos y verifica que $\lambda(G) \leqslant 4\lambda(\biguplus \mathcal{W})$. Puesto que $B \subset G$, podemos escribir:

$$\lambda^*(B) \leqslant \lambda(G) \leqslant 4 \, \lambda \left(\biguplus_{x \in B_0} I_x \right) = 4 \sum_{x \in B_0} (b_x - a_x) \tag{4}$$

Para cada $x \in B_0$ consideramos ahora el intervalo abierto $J_x =]\, f(a_x)\,,\, f(b_x)\,[$ que en vista de (3) no es vacío, y vemos que los intervalos así definidos también son dos a dos disjuntos. En efecto, dados $x, y \in B_0$ con $x \neq y$, podemos suponer sin perder generalidad que $a_x < a_y$, con lo que, por ser $I_x \cap I_y = \emptyset$ se tiene $b_x \leqslant a_y$. Como f es creciente, deducimos que $f(b_x) \leqslant f(a_y)$, con lo que $J_x \cap J_y = \emptyset$. También por ser f creciente, tenemos claramente que $J_x \subset [\, f(b)\,,\, f(a)\,]$ para todo $x \in B_0$. Usando entonces (3) y (4) obtenemos que

$$\begin{aligned} f(b) - f(a) &= \lambda\big([\, f(a)\,,\, f(b)\,]\big) \geqslant \lambda\left(\biguplus_{x \in B_0} J_x\right) \\ &= \sum_{x \in B_0} \big(f(b_x) - f(a_x)\big) > M \sum_{x \in B_0} \big(b_x - a_x\big) \geqslant M \, \frac{\lambda^*(B)}{4} \end{aligned}$$

Así pues, tenemos $\lambda^*(B) < 4\big(f(b) - f(a)\big)/M$, pero usando que $M \in \mathbb{R}^+$ era arbitrario, concluimos que $\lambda^*(B) = 0$, como se quería. ■

Pasamos a la segunda etapa de la demostración, que será más laboriosa.

Proposición 2. *Dados $a, b \in \mathbb{R}$ con $a < b$, toda función creciente $f : [\, a\,,\, b\,] \to \mathbb{R}$ verifica que $\underline{D} f(x) = \overline{D} f(x)$ para casi todo $x \in [\, a\,,\, b\,]$.*

Demostración. Sea A el conjunto formado por las discontinuidades de f en $]\, a\,,\, b\,[$, junto con los puntos a y b. Por un resultado previo, vemos que A es numerable. Consideramos ahora el conjunto definido por $C = \big\{x \in [\, a\,,\, b\,] \setminus A \,:\, D^- f(x) < D^+ f(x)\big\}$.

Para $x \in [\, a\,,\, b\,] \setminus (A \cup C)$ tenemos claramente que $\underline{D} f(x) = \overline{D} f(x)$, luego bastará probar que $A \cup C$ tiene medida nula. Como A es numerable, tenemos $\lambda(A) = 0$, y queda probar que también C tiene medida nula. Por reducción al absurdo, suponemos que $\lambda^*(C) > 0$, para llegar a una contradicción.

Para $\alpha, \delta \in \mathbb{Q}^+$ consideramos el conjunto

$$C(\alpha, \delta) = \big\{x \in C \,:\, \underline{D} f(x) < \alpha - \delta < \alpha + \delta < \overline{D} f(x)\big\}$$

Dado $x \in C$, tenemos $0 \leqslant \underline{D} f(x) < \overline{D} f(x)$, luego es claro que podemos encontrar $\alpha, \delta \in \mathbb{Q}^+$ tales que $x \in C(\alpha, \delta)$. Por tanto, podemos expresar C como unión numerable de conjuntos:

$$C = \bigcup_{\alpha, \delta \in \mathbb{Q}^+} C(\alpha, \delta)$$

Como hemos supuesto que $\lambda^*(C) > 0$, han de existir $\alpha, \delta \in \mathbb{Q}^+$ tales que $\lambda^*\big(C(\alpha, \delta)\big) > 0$. Abreviamos la notación escribiendo $E = C(\alpha, \delta)$ y repasamos las propiedades del conjunto E, que nos deben llevar a una contradicción. Por una parte tenemos que $\lambda^*(E) > 0$, y por otra, para cada $x \in E$ tenemos que $\underline{D} f(x) < \alpha - \delta < \alpha + \delta < \overline{D} f(x)$, y como $x \in C$, también sabemos que $a < x < b$ y que f es continua en el punto x.

A partir de ahora usaremos la función $g : [\, a\,,\, b\,] \to \mathbb{R}$ definida por $g(x) = f(x) - \alpha x$ para todo $x \in [\, a\,,\, b\,]$, que es claramente una función de variación acotada. La forma de llegar a una contradicción consistirá en construir una partición $Q \in \Pi(a, b)$ tal que $\sigma(g, Q) > V(g\,; a, b)$.

Empezamos usando la definición de $V(g\,;a,b)$ para encontrar $P \in \Pi(a,b)$ que verifique:

$$\sigma(g,P) + \frac{\delta\lambda^*(E)}{4} > V(g\,;a,b) \tag{5}$$

Escribimos $P = \{a = x_0 < x_1 < \ldots < x_n = b\}$ y, fijado $k \in \Delta_n$, trabajamos por ahora en el intervalo $]x_{k-1}, x_k[$. Abreviamos escribiendo $E_k = E \cap]x_{k-1}, x_k[$, y suponiendo que $E_k \neq \emptyset$, consideramos los dos casos que pueden darse, con vistas a usar el lema 2.

(a). Caso en que se tenga $g(x_{k-1}) \leqslant g(x_k)$. Para cada $x \in E_k$, la definición de $\underline{D} f(x)$ nos permite encontrar $r \in \mathbb{R}^+$, con $x_{k-1} < x - r < x + r < x_k$, tal que $\underline{D}_r f(x) < \alpha - \delta$. Por tanto, existe $z \in \mathbb{R}$, con $0 < |z - x| < r$, verificando que $\dfrac{f(x) - f(z)}{x - z} < \alpha - \delta$.

Si $z < x$, tomamos $a_x = z$ y usamos la continuidad de f en x para encontrar un $b_x \in \mathbb{R}$, con $x < b_x < x + r$, que verifique: $\dfrac{f(b_x) - f(a_x)}{b_x - a_x} < \alpha - \delta$.

Si por el contrario $z > x$, tomamos $b_x = z$ y usamos también la continuidad de f en x para obtener $a_x \in \mathbb{R}$ con $x - r < a_x < x$, que de nuevo verifique: $\dfrac{f(b_x) - f(a_x)}{b_x - a_x} < \alpha - \delta$.

En resumen, a cada $x \in E_k$ hemos asociado un intervalo abierto $I_x =]a_x, b_x[\subset]x_{k-1}, x_k[$, verificando que

$$\frac{g(b_x) - g(a_x)}{b_x - a_x} = \frac{f(b_x) - f(a_x) - \alpha(b_x - a_x)}{b_x - a_x} < \alpha - \delta - \alpha = -\delta \tag{6}$$

(b). Cuando $g(x_{k-1}) \geqslant g(x_k)$, razonamos de manera similar. Para cada $x \in E_k$, la definición de $\overline{D} f(x)$ nos da un $r \in \mathbb{R}^+$ verificando que $x_{k-1} < x - r < x + r < x_k$ y $\overline{D}_r f(x) > \alpha + \delta$, luego existe un $w \in \mathbb{R}$, con $0 < |w - x| < r$, tal que $\dfrac{f(x) - f(w)}{x - w} > \alpha + \delta$.

Razonando igual que en el caso *(a)*, encontramos $a_x, b_x \in \mathbb{R}$ con $x - r < a_x < x < b_x < x + r$, tales que $\dfrac{f(b_x) - f(a_x)}{b_x - a_x} > \alpha + \delta$.

De nuevo, a cada $x \in E_k$ hemos asociado un intervalo abierto $I_x =]a_x, b_x[\subset]x_{k-1}, x_k[$, que ahora verifica:

$$\frac{g(b_x) - g(a_x)}{b_x - a_x} = \frac{f(b_x) - f(a_x) - \alpha(b_x - a_x)}{b_x - a_x} > \alpha + \delta - \alpha = \delta \tag{7}$$

El siguiente razonamiento es análogo al usado en la Proposición 1. Tenemos una familia de intervalos abiertos $\mathcal{U} = \{I_x : x \in E_k\}$ cuya unión es un abierto acotado $G_k \subset]x_{k-1}, x_k[$. Usando el lema 1, conseguimos una subfamilia finita $\mathcal{W} = \{I_x : x \in F_k\}$, donde F_k es un subconjunto finito de E_k, que está formada por intervalos dos a dos disjuntos y verifica que

$$\lambda^*(E_k) \leqslant \lambda(G_k) \leqslant 4\,\lambda\left(\biguplus_{x \in F_k} I_x\right) = 4 \sum_{x \in F_k} \big(b_x - a_x\big) \tag{8}$$

Consideramos ahora una partición del intervalo $[x_{k-1}, x_k]$, que incluye los extremos de los intervalos I_x con $x \in F_k$. Concretamente: $P_k = \{x_{k-1}, x_k\} \cup \{a_x : x \in F_k\} \cup \{b_x : x \in F_k\}$.

Numeramos como es habitual, $P_k = \{x_{k-1} = y_0 < y_1 < \ldots < y_m = x_k\}$, y consideramos el conjunto $S \subset \Delta_m$ formado por los índices $j \in \Delta_m$ tales que $y_{j-1} = a_x$ para algún $x \in F_k$. Por ejemplo, es claro que $y_0 = x_{k-1} < y_1 = \min\{a_x : x \in F_k\}$, luego $2 \in S$ pero $1 \notin S$. Observamos también que, si $j \in S$ porque $y_{j-1} = a_x$ con $x \in F_k$, se tiene obligadamente $y_j = b_x$. En efecto, en otro caso se tendría $b_x > y_j \geqslant a_z$, y por tanto $I_x \cap I_z \neq \emptyset$, para algún $z \in F_k$ con $z \neq x$, lo que no es posible. Así pues, $\big\{]y_{j-1}, y_j[\,:\, j \in S \big\} = \big\{ I_x : x \in F_k \big\}$, y usando (8) obtenemos

$$\sum_{j\in S} \big(y_j - y_{j-1}\big) = \sum_{x\in F_k} \big(b_x - a_x\big) \geqslant (1/4)\,\lambda^*(E_k) \tag{9}$$

Además, usando (6) o (7) según el caso, se verifica una de las siguientes afirmaciones:

(a) $g(x_{k-1}) \leqslant g(x_k)$ y $g(y_j) - g(y_{j-1}) < -\delta(y_j - y_{j-1})$ para todo $j \in S$

(b) $g(x_{k-1}) \geqslant g(x_k)$ y $g(y_j) - g(y_{j-1}) > \delta(y_j - y_{j-1})$ para todo $j \in S$

Esto nos permite usar el lema 2 para obtener que

$$\sigma(g,P_k) = \sum_{j=1}^{m} \big|g(y_j) - g(y_{j-1}\big| \geqslant \big|g(x_k) - g(x_{k-1})\big| + \delta \sum_{j\in S} \big(y_j - y_{j-1}\big)$$

y teniendo en cuenta (9) concluimos que

$$\sigma(g,P_k) \geqslant \big|g(x_k) - g(x_{k-1})\big| + \frac{\delta\,\lambda^*(E_k)}{4} \tag{10}$$

Recordemos que el razonamiento anterior se ha hecho suponiendo que $E_k \neq \emptyset$. En otro caso tenemos $\lambda^*(E_k) = 0$, y tomando $P_k = \{x_{k-1}, x_k\}$ se tiene obviamente la desigualdad (10).

Tomamos finalmente $Q = \bigcup_{k=1}^{n} P_k$, que es una partición del intervalo $[a, b]$. Usando las n desigualdades que aparecen en (10) obtenemos claramente que

$$\sigma(g,Q) = \sum_{k=1}^{n} \sigma(g,P_k) \geqslant \sum_{k=1}^{n} \big|g(x_k) - g(x_{k-1})\big| + \frac{\delta}{4}\sum_{k=1}^{n} \lambda^*(E_k) \tag{11}$$

Recordando que $E_k = E \cap]x_{k-1}, x_k[$ para todo $k \in \Delta_n$, tenemos que $E \subset P \bigcup \left(\bigcup_{k=1}^{n} E_k \right)$, y por ser $\lambda^*(P) = 0$, obtenemos que $\lambda^*(E) \leqslant \sum_{k=1}^{n} \lambda^*(E_k)$. Por último, recordando también la definición de $\sigma(g,P)$, y usando (5), de (11) deducimos que

$$\sigma(g,Q) \geqslant \sigma(g,P) + \frac{\delta\,\lambda^*(E)}{4} > V(g\,;a,b)$$

Según lo previsto, esto contradice la definición de $V(g\,;a,b)$, concluyendo la demostración. ■

La parte difícil del trabajo está ya hecha. Las dos proposiciones anteriores prueban que toda función creciente en un intervalo compacto es derivable casi por doquier.

9.4. Forma general del teorema

Mediante una fácil generalización de lo recién probado, llegamos a uno de los resultados más importantes en el estudio de las funciones reales de una variable real.

Teorema de derivación de Lebesgue. *Dado un intervalo no trivial $J \subset \mathbb{R}$, supongamos que una función $f : J \to \mathbb{R}$ tiene variación acotada en cada intervalo compacto $K \subset J$. Entonces f es derivable casi por doquier en J.*

Demostración. Si $J = K$ es un intervalo compacto, como f es una función de variación acotada, podemos escribir $f = g - h$, donde $g, h : K \to \mathbb{R}$ son funciones crecientes. Usando las proposiciones 1 y 2, obtenemos que

$$\overline{D}\, g(x) = \underline{D} g(x) < \infty \quad \text{p.c.t. } x \in K \qquad \text{y} \qquad \overline{D}\, h(x) = \underline{D} h(x) < \infty \quad \text{p.c.t. } x \in K$$

Por tanto existen dos conjuntos $A, B \subset K$, con $\lambda(K \setminus A) = \lambda(K \setminus B) = 0$, tales que g es derivable en A y h es derivable en B. Tomando $C = A \cap B$, tenemos que $\lambda(K \setminus C) = 0$ y f es derivable en C, luego f es derivable c.p.d. en K.

En el caso general, es claro que podemos escribir $J = \bigcup_{n=1}^{\infty} K_n$ donde K_n es un intervalo compacto para todo $n \in \mathbb{N}$. Fijado $n \in \mathbb{N}$, tenemos por hipótesis que $f\big|_{K_n}$ es una función de variación acotada, luego existe un conjunto $C_n \subset K_n$, con $\lambda(K_n \setminus C_n) = 0$, tal que $f\big|_{K_n}$ es derivable en C_n. Suprimiendo a lo sumo dos puntos del conjunto C_n, podemos claramente suponer que $C_n \subset K_n^{\circ}$. Esto permite usar el carácter local de la derivada, para obtener que f es derivable en todo punto de C_n.

Tomando ahora $E = \bigcup_{n=1}^{\infty} C_n$, es claro que $\lambda(J \setminus E) = 0$ y, para cada $x \in J \setminus E$, existe $n \in \mathbb{N}$ tal que $x \in C_n$, luego f es derivable en el punto x. Esto prueba que f es derivable casi por doquier en J. ■