MACHETE DE IDEAS DE DEMOS

Por Teo López Puccio y Uriel Chami

Cuando uno estudia para el final de Análisis I, es normal ver resúmenes con las demostraciones enteras de los teoremas. Pero es difícil memorizar todas esas cadenas lógicas llenas de detalles y notación complicada. Si uno quiere aprender las pruebas, primero es indispensable entender la demostración, pero luego es más práctico aprender un mapita mental de cómo llegar al resultado, sin darle mucha importancia al detalle. **Es mejor aprender a buscar el camino que memorizar cada paso**.

Por eso hicimos este resumen delineando las *ideas* de algunas demostraciones, dejando de lado los formalismos. No vas a encontrar los dominios de algunas funciones, definiciones de conceptos ni notación muy rigurosa, a menos que sean centrales para entender la demostración. Pero, por supuesto, no olvides sentarte a escribir las cosas bien.

Este resumen es producto de un estudio de a dos, que ha dado muy buenos resultados. Recomendamos ampliamente estudiar con otros, porque explicar conceptos es la mejor manera de asegurar el entendimiento y no caer en auto-engaños. ¡Mucha suerte!

Límite de f via sucesiones[1]

Una f tiene límite $\iff$ las imágenes de toda sucesión de puntos del dominio tienen ese mismo límite.

Dem: Hay $\Rightarrow$ y $\Leftarrow$.

$\Rightarrow$) Supongo que vale el límite y "sea $\{P_k\}$ una sucesión $\in A$ que tiende a P". Luego uno *compone* con la definición de límite.

$\Leftarrow$) Suponemos todas las sucesiones tienden a P, pero que no existe el límite de f. Enunciamos el recíproco: para cualquier $\delta > 0$, hay un X en un entorno de P de radio δ que hace $\|f(X) - L\| \geq \varepsilon$. Si tomamos $\delta = \frac{1}{k}$, nos armamos una sucesión $\{X_k\}$ que tiende a P y contradice la hipótesis.

Fermat en $\mathbb{R}$

Si una f derivable tiene un extremo relativo en el interior de su dominio, la derivada allí se anula.

Dem: Los límites laterales de la derivada por definición resultan $\geq$ y ≤ 0, entonces son iguales. Dibujito ayuda.

Rolle en $\mathbb{R}$

En una función derivable, si $f(a) = f(b)$, la derivada se anula en algún $c \in (a, b)$.

Dem: La función alcanza max y mín absoluto en $[a, b]$, y alguno se alcanza en el abierto, entonces aplico **Fermat**. (Recordar: si f es constante, *"no hay nada que probar"*.)

Lagrange en $\mathbb{R}$

(o TVM)

Si f diferenciable, existe $c \in (a, b)$ donde $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Dem: Tomamos la recta que une los dos puntos y una funcion auxiliar $h(x) = f(x) - L(x)$. Esto cumple las hipótesis de **Rolle** y al derivarla obtenemos el resultado.

[1] Este teorema a veces se enuncia como un teorema de continuidad. Esto es idéntico, solo hay que tomar el caso en que $\lim_{x\to p} f(x) = f(p)$

Bolzano en $\mathbb{R}$

Si f es continua y $f(a) \cdot f(b) < 0$, la función se anula en algún $c \in (a, b)$.

Necesitamos dos lemas:

• **Lema copado** nos dice que si una suc. $\{X_k\} \in \mathrm{Dom} f \to P$, y $f(X_k) \leq 0 \;\Rightarrow\; f(P) \leq 0$. Se demuestra por absurdo, tomando $\varepsilon = \frac{f(P)}{2}$, y viceversa análogo tomando $\varepsilon = -\frac{f(P)}{2}$.

• **Conservación de signo** nos dice que si $f(p) > 0$, hay un entorno de "positividad" alrededor de p. Se prueba por el absurdo: si no fuera cierto, en cada bola $B_{\frac{1}{k}}(P)$, $k \in \mathbb{N}$, tendríamos un X_k de imagen ≤ 0. Entonces por el **lema copado**, tendríamos una sucesión que contradice la hipótesis.

Dem de Bolzano: Nombramos A al conjunto de positividad de f. Tiene supremo $= s$ por ser no vacío. Usamos el **lema** para ver que $f(s) \geq 0$ (porque hay una sucesión de A que tiende a s). Usamos **conservación de signo** para ver que $f(s) \not> 0$ (porque si fuera mayor que cero, habría un entorno de positividad, y habría $x_0 \in A > s$, abs.) Al final tomamos $s = c$.

Bolzano en $\mathbb{R}^n$

Si el dominio de f continua es arcoconexo, y $f(A) < 0$, $f(B) > 0$, existe P en el dominio que anula la función.

Dem: Por ser arcoconexo, hay una curva continua del dominio, $\gamma(t)$, que une A con B. La composición $f \circ \gamma\ (t)$ es derivable, entonces aplicamos **bolzano en** $\mathbb{R}$.

Weierstrass (en $\mathbb{R}^n$)

Una f continua alcanza máximo y mínimo en un compacto.

Dem: tiene dos partes. En **1.** probamos que la imagen es acotada, en **2.** probamos que se alcanzan los extremos.

1. Suponemos que f no es acotada por arriba (y por debajo es análogo). Entonces existe una sucesión $\{X_k\}$ de puntos del dominio cuyas imágenes divergen. (Esto se dice: $f(X_k) \geq k \;\forall\; k \in \mathbb{N}$.) Como **cualquier sucesión acotada contiene una subsucesión convergente**, hay una subsuc. $\{X_{k_j}\}$ que tiende a algún P del dominio. Pero las imágenes de esta también divergen. (Se escribe $f(X_{k_j}) \geq k_j \;\forall\; k_j \in \mathbb{N}$.) Esto es absurdo por ser f continua.

> Nota: en el *Larotondum* esto se da por absurdo. Pero si no alcanza *sentir* que esto es obvio, podemos escribir que, por ser f continua, existe un k_j a partir del cual $|k_j - f(P)| \leq |f(X_{k_j}) - f(p)| < \varepsilon$. Esto sí es manifiestamente absurdo, ya que $f(P)$ tiene un valor fijo y k_j es arbitrariamente grande.

2. Usamos lo de recién, **1.**, para decir que $\mathrm{Im}(f)$ es acotada. Entonces tiene supremo s, y hay una sucesión de puntos del dominio cuyas imágenes tienden a s. Como esta suc. es acotada, existe una **subsucesión que converge** a algún P_m. Como f **es continua**, $f(P_m) = \lim_k f(X_k) = s$.

Unicidad del diferencial

y existencia de las derivadas direccionales

Si existe una transformación lineal $T_p : \mathbb{R}^n \to \mathbb{R}$ *que cumpla el límite de la definición, entonces existen todas las* $\frac{\partial f}{\partial V}(P)$ *con* $\|V\| = 1$*. También existen las derivadas parciales, la transformación es única, y resulta que*

$$\frac{\partial f}{\partial V}(P) = T_p(V) = Df_P(V) = \langle \nabla f_P, V \rangle.$$

Dem: Hay que ver que existen las $\frac{\partial f}{\partial V}(P)$, que por definición son $= \lim_{t\to 0} \frac{f(P+tV)-f(P)}{t}$. Si vale el límite del enunciado, vale también **componiendo** con cualquier curva que tenga límite en P. Entonces tomamos $X = P + tV$, y todo encaja a la perfección:

- El límite de $X \to P$ se vuelve $t \to 0$.
- $T_P(X - P) = T_P(tV)$ y la t sale afuera por ser lineal.
- $\|X - P\| = \|tV\| = |t|$

Entonces el límite que nos queda muestra que $T_P(V) = \frac{\partial f}{\partial V}(P)$. Inmediatamente deducimos que las f_{x_i} son iguales a $T_p(E_i)$, y como una transf. lineal queda completamente determinada por su valor en una base de $\mathbb{R}^n$, demostramos la *alta formuleta*[2] $f_V(P) = \langle \nabla f(P), V \rangle$

Si hubiera otra transf. lineal que cumpliera las mismas hipótesis, también cumpliría lo de recién, entonces sería necesariamente igual a T_P.

Por último, nuestra *alta formuleta* nos muestra que T_P es igual al $D_P f$, porque así habíamos definido el diferencial. Igualmente queda probado que la función es diferenciable en P.

Dirección de mayor crecimiento

Para f *diferenciable en* P*, y una dirección* $\|V\| = 1$*,* $\frac{\partial f}{\partial V}$ *se maximiza en dirección del gradiente, y su valor máximo es* $\|\nabla f_P\|$*.*

Demostración en dos pasos: Usar **desigualdad C-S** para ver que $\langle \nabla f_P, V \rangle$ es a lo sumo la norma del gradiente, luego tomar V igual al gradiente normalizado.

Demostración en un paso: Usar la **propiedad** del producto escalar $\langle \nabla f_P, V \rangle = \|\nabla f_P\| \|V\| \cos\theta$ (con θ el ángulo entre ambos vectores). El coseno se maximiza si $\theta = 0$, o sea, cuando apuntan en la misma dirección.

Fermat en $\mathbb{R}^n$

Si f *es diferenciable en un abierto y* P *un extremo local,* $\nabla f_P = \mathbb{O}$*.*

Dem: Suponemos que P es máximo local (y el mínimo es análogo). Decimos que entonces es máximo absoluto en una bola de radio r. Si **componemos** f con la recta de dirección E_i que pasa por P, la composición es derivable, también tiene un extremo allí, y **Fermat** nos dice que su derivada en P se anula. Por la definición de derivada direccional, y porque pasa para cualquier E_i, queda probado.

[2] "*Alta formuleta*" es un término técnico extremadamente avanzado que puede generar confusiones entre matemáticos no especializados. Por esta única razón recomendamos evitar su uso en el examen final, ya que el docente a cargo puede no ser versado en *teoría de altas formuletas*.

Diferenciable ⇒ Continua

Dem: Queremos probar $\lim_{X\to P} f(X) = f(P)$. Es lo mismo probar que $\lim_{X\to P} |f(X) - f(P)| = 0$.

Escribimos $|f(X) - f(P)|$, y queremos llegar a una expresión mayor o igual, que contenga a la definición de diferenciable. Entonces habremos ganado. La receta para esto es:

1. Sumar y restar $D_P f(X-P)$ dentro del módulo.
2. Aplicar desigualdad **triangular** para separar el término $+|D_P f(X-P)|$
3. Escribir $D_P f(X-P)$ como $\langle \nabla f(P), (X-P)\rangle$
4. A eso aplicarle desigualdad de **Cauchy-Schwartz**: $\langle \nabla f(P), (X-P)\rangle \le \|\nabla f(P)\| \cdot \|X-P\|$
5. Al otro término multiplicarlo y dividirlo por $\|X-P\|$

Y ahora, si f es diferenciable, todo esto tiende a cero si $X \to P$, y al ser $\ge |f(X) - f(P)|$ alcanzamos la victoria prometida.

Lagrange en $\mathbb{R}^n$

Si f es diferenciable en un dominio convexo, $P, Q \in dom(f)$, hay un P_0 que cumple

$$f(Q) - f(R) = \langle \nabla f_{P_0}, (Q-R)\rangle$$

Dem: Se parametriza la recta que une ambos puntos, se compone con f, se apela a la **regla de la cadena** para afirmar que la composición es derivable, continua, y que entonces en ella vale **Lagrange** en $\mathbb{R}$. El resultado se obtiene aplicando **regla de la cadena** a la composición. (¡OJO ESA CADENA!: el diferencial de f es el gradiente evaluado en el punto que terminamos llamando P_0, y la derivada de la paramétrica es $Q-R$.)

C¹ ⇒ diferenciable (en $\mathbb{R}^2$)

Dem: Suponemos que las derivadas parciales de f existen en todos lados. Tomamos dos vectores: $P = (a,b)$ y $X = (a+h, b+k)$. Queremos armar el “límite definición” de diferenciabilidad acotado por las derivadas parciales. (ATENCIÓN: MEGA RESUMEN) La receta es:

1. Escribir $f(X) - f(P) =$
2. Restar y sumar $f(a, b+k)$ para aplicar **Lagrange** en $\mathbb{R}$ a las funciones $f(x, b+k)$ y $f(a,y)$ (que se puede por hipótesis). Reescribir los dos términos.
3. Restar de ambos lados $\langle \nabla f_P, (X-P)\rangle$. Sacar factor común h y k en el lado derecho.
4. Aplicar módulo a ambos lados y desigualdad triangular.
5. Acotar $|h|, |k| \le \|X-P\|$, pasar esa norma dividiendo.

Ya nos armamos el “límite definición” de diferenciabilidad del lado izquierdo. Entonces cuando hacemos que $X - P = (h,k) \to 0$, lo de la derecha tiende a 0, y listo.

$C^2 \Rightarrow$ Las derivadas cruzadas coinciden

Dem: Se demuestra sobre una función de $\mathbb{R}^2 \to \mathbb{R}$, sin pérdida de generalidad. Primero consideramos un punto (a, b) del dominio, y pedimos auxilio a la función

$$g(t) = f(a+t, b+t) - f(a+t, b) - f(a, b+t) + f(a, b).$$

Con esta vamos a encontrar las dos derivadas cruzadas. La receta es:

1. Deφnimos $\varphi(x) = f(x, b+t) - f(x, b)$, para escribir $g(t) = \varphi(a+t) - \varphi(a)$

2. Notamos que φ es dos veces derivable. Entonces escribimos a $g(t)$ como su polinomio de **Taylor** de primer orden, centrado en 0 y en función de φ. Lo planteamos como igualdad, sin olvidar el término del **resto de Lagrange**, con un c que la hace cumplir.

3. Reescribimos las derivadas de φ como las derivadas parciales de f, evaluadas en los puntos que correspondan.

4. Dividimos ambos lados de la igualdad por t^2, tomamos límite cuando $t \to 0$ y notamos la magia. Un término tiende a cero porque $f \in C^2$. El otro es, por definición, una derivada cruzada. Nos queda
$$\lim_{t\to 0} \frac{g(t)}{t^2} = \frac{\partial^2 f}{\partial y \partial x}(a, b)$$

5. Ahora volvemos a 1. y tomamos $\psi(y) = f(a+t, y) - f(a, y)$. Argumentamos igual que antes, escribiendo el Taylor de $g(t)$ con nuestra nueva ψ. Llegamos al mismo límite con la otra derivada cruzada, y terminamos (si se quiere, por *unicidad del límite*).

Polinomio de Taylor

de orden dos en $\mathbb{R}^n$, con resto de Lagrange

Si f tiene un dominio convexo y es C^3, para todo $P \in dom(f)$ se tiene

$$f(X) = f(P) + \langle \nabla f_P, X - P\rangle + \tfrac{1}{2}\langle Hf_P \cdot (X-P), (X-P)\rangle + R_P(X-P)$$

$$\textit{donde} \quad R_P(X-P) = \tfrac{1}{6}\langle D^3 f_c(X-P)\cdot(X-P), X-P\rangle$$

con c en el segmento que une X con P, y además

$$\lim_{X\to P} \frac{R_P(X-P)}{\|X-P\|^2} = 0.$$

Dem: Primero vemos la expresión del polinomio, después mostramos el límite del resto. Primero definimos $g(t)$ como $f(P + t(X-P))$, la composición de f con la recta del dominio que une a X con P. Armamos el polinomio de Taylor de orden 2 de g, centrado en 0 (y con su resto para tener igualdad). Si calculamos $g'(t)$, $g''(t)$, $g'''(t)$ con un par de cadenitas, y evaluamos en $t = 1$, tenemos nuestra expresión para $f(x)$. *Solo usamos Taylor en una variable, pero con notación que da más miedo.*

Para ver el límite, usamos dos desigualdades:

$$|\langle D^3 f_c(X-P)\cdot(X-P), (X-P)\rangle| \leqslant \|D^3 f_c(X-P)\|_\infty \|X-P\|^2 \leqslant \sum_{i=1}^{n}\left(\left\|H\frac{\partial f}{\partial x_i}(c)\right\|_\infty^2\right)^{\frac{1}{2}} \|X-P\|^3$$

Como las derivadas terceras son continuas por ser $f \in C^3$, la suma de todas ellas está acotada por un M en un entorno de P. Entonces vale $\frac{R_P(X-P)}{\|X-P\|^2} \leq M\|X-P\|$ si X está suficientemente cerca de P. Esto prueba el límite.

Criterio del Hessiano

• definida positiva $\Rightarrow$ mínimo

Si $f : A \subset \mathbb{R}^n \to \mathbb{R}$ con A abierto es C^3, y en un punto crítico P sucede que Hf_P es definida positiva, entonces P es mínimo local. Análogamente, definida negativa $\Rightarrow$ máximo.

Dem:

1. Escribimos $f(X)$ como su **polinomio de Taylor** de segundo orden, centrado en P. Por hipótesis P es punto crítico, así que se anula su ∇, y con él se va todo el término lineal.

2. Juntamos al término del resto y al término cuadrático en un solo paréntesis. Lo multiplicamos y dividimos por $\|X-P\|^2$. Por las propiedades del producto escalar, esta norma entra en el producto y queda así:

$$f(X) = f(P) + \|X-P\|^2 \left[\frac{1}{2} \left\langle Hf_P \frac{X-P}{\|X-P\|}, \frac{X-P}{\|X-P\|} \right\rangle + \frac{R(X-P)}{\|X-P\|^2} \right]$$

Hay una notación cómoda para las formas cuadráticas, que es $\frac{1}{2}\langle Hf_P V, V\rangle = Q_P(V)$. Usándola, esto nos queda

$$f(X) = f(P) + \|X-P\|^2 \left[Q_P\left(\frac{X-P}{\|X-P\|} \right) + \frac{R(X-P)}{\|X-P\|^2} \right] \qquad (*)$$

3. Estamos evaluando la forma cuadrática en un vector de norma unitaria. Este vector se encuentra en la superficie de la esfera unitaria de $\mathbb{R}^n$, que es **compacta**. Y un lema nos dice que la forma cuadrática es una función **continua**. Entonces por **Weierstrass** hay un vector unitario V' para el cual $Q_P(V')$ es mínimo. A este mínimo le decimos m. También sabemos que $Q_P(X)$ es definida positiva. Esto significa que $Q_P(X) > 0 \ \forall \ X \neq \mathbb{O}$. Y como $V' \neq \mathbb{O}$, encontramos que $m > 0$.

4. Ahora a acotar. Sabemos por el **Resto de Taylor** que $\frac{R(X-P)}{\|X-P\|^2} \xrightarrow{X \to P} 0$. Significa que en un entorno $B_\delta(P)$ tenemos $\left| \frac{R(X-P)}{\|X-P\|^2} \right| < \varepsilon$. Como en 3. encontramos que $m > 0$, podemos tomar $\varepsilon = \frac{m}{2}$. La **cota central** es:

$$\left[Q_P\left(\frac{X-P}{\|X-P\|} \right) + \frac{R(X-P)}{\|X-P\|^2} \right] > m - \frac{m}{2} > 0.$$

Entonces vemos que $(*)$ nos queda $f(X) = f(P) + \text{ALGO POSITIVO}$ para $X \in B_\delta$. Esto muestra que P es mínimo local de f.

• indefinido $\Rightarrow$ punto silla

Esto se demuestra con un lema previo:

Si $f\, C^3, \nabla f_P = \mathbb{O}$ y $\exists\ V \ / \ Q_P(V) < 0$, entonces a lo largo de $P + tV$ hay un máximo local de f.

Por definición, si $Hf(P)$ es indefinida, hay dos vectores en donde la forma cuadrática es positiva y negativa respectivamente. Entonces del lema se deduce que f tiene un punto silla.

Este lema misterioso se prueba exactamente igual que lo que hicimos más arriba, pero a través de una recta. Se escribe el **Taylor**, se toma $X = P + tV$, se **multiplica y divide** por $\|tV\|^2$ y se **acota** el corchete para mostrar su positividad (o negatividad), que implica que P es extremo de f a través de la recta. (Recordar que $Q_P(tV) = t^2 Q_P(V)$. ¡Por eso es la forma cuadrática!).

Multiplicadores de Lagrange en $\mathbb{R}^n$

Sean las funciones $f, g : \mathbb{R}^n \to \mathbb{R}$ y la superficie de nivel $Niv = \{X \in \mathbb{R}^n : g(X) = c \in \mathbb{R}\}$. Si P es un extremo de f restringida a Niv, y $\nabla g_P \neq \mathbb{O}$, entonces $\nabla f_P = \lambda \nabla g_P$ para algún $\lambda \in \mathbb{R}$.

Dem: Como Niv es una superficie de $\mathbb{R}^n$ definida implícitamente, y $\nabla g_p \neq \mathbb{O}$, se usa el **Teorema de la Función Implícita**: existe una bola $B \subset \mathbb{R}^{n-1}$ y una función $\varphi : B \to \mathbb{R}$, cuyo gráfico es Niv en un entorno de P. Digamos que la bola está centrada en un $Z_0 \in \mathbb{R}^{n-1}$ y, por simplicidad, que la derivada de g que no se anula es la última. Todo esto significa que

$$P = (\underset{\substack{\downarrow \\ \textit{n-1 coordenadas}}}{Z_0}, \varphi(Z_0)) \quad \text{y además} \quad (Z, \varphi(Z)) \in Niv \ \forall Z \in B.$$

Como P es un extremo de f restringida a Niv, $f(P) \geq f(X)\ \forall X \in Niv$ cerca de P. Escribiendo esos puntos como recién, $f(Z_0, \varphi(Z_0)) \geq f(Z, \varphi(Z))$. O sea que podemos ver a Z_0 como un máximo local de la función

$$h(Z) : B \subset \mathbb{R}^{n-1} \to \mathbb{R}, \quad h(Z) := f(Z, \varphi(Z)),$$

y entonces el gradiente de h se anula en Z_0. Esto es, cada $\frac{\partial h}{\partial z_i}$ se anula:

$$\begin{aligned}
0 &= \frac{\partial}{\partial z_i}\, h(Z)\bigg|_{Z_0} \\
&= \frac{\partial}{\partial z_i}\, f(Z, \varphi(Z))_{Z_0} \\
\textit{por } \textbf{\textit{regla de la cadena}} \rightarrow &= \left\langle \nabla f_{(Z_0, \varphi(Z_0))}, \frac{\partial}{\partial z_i}(Z_0, \varphi(Z_0)) \right\rangle \\
0 &= \left\langle \nabla f_P, \left(E_i, \frac{\partial}{\partial z_i}\varphi(Z_0)\right) \right\rangle
\end{aligned}$$

Para pensar esa última derivada sirve ver que $(Z, \varphi(Z))$ es en sí una función de $\mathbb{R}^{n-1}$ en $\mathbb{R}$.

Solo falta recordar que $(E_i, \frac{\partial}{\partial z_i}\varphi(Z))$ son los vectores generadores del plano tangente a Niv, cuya normal es ∇g_P. Encontramos que ∇f_P es $\perp$ a todos ellos, y la normal del plano es ∇g_P, así que en P ambos gradientes son paralelos.

Regla de la cadena en $\mathbb{R}^n$

Si el codominio de F es el dominio de G, si F es dif. en P y G es dif. en $Q = F(P)$, entonces $G \circ F$ es dif. en P y

$$D(G \circ F)_P = DG_{F(P)} DF_P \ \leftarrow \textit{composición de transf. lineales.}$$

Dem: Escribimos el límite de la definición de diferenciabilidad para $G \circ F$ con el diferencial propuesto, y queremos ver que tiende a cero. Arriba, se resta y se suma $DG_{F(P)}(F(X) - F(P))$. El que resta se separa a la izquierda con $G(F(X)) - G(F(P))$, por desigualdad triangular. Ahora numerador del término de la derecha es $\leq \|DG_Q\|_\infty \|F(X) - F(P) - DF_P(X - P)\|$ (como si sacáramos la matriz factor común. La cota vale porque $\|\cdot\|_\infty$ de la matriz se define como cota superior del producto matriz-vector). Por hipótesis ese término tiende a cero. Luego se multiplica y divide el término de la izquierda por $\|F(X) - F(P)\|$. Una parte del producto es la definición de diferenciabilidad de G en Q, y la otra se puede acotar así:

$$\frac{\|F(X) - F(P)\|}{\|X - P\|} \leq \frac{\|F(X) - F(P) - DF_P(X - P)\|}{\|X - P\|} + \frac{\|DF_P(X - P)\|}{\|X - P\|} \leq 1 + \|DF_P\|_\infty$$

por ser F diferenciable en P. Entonces nuestro cociente tiende a cero.

Criterio de integrabilidad Riemann

Una $f : [a, b] \to \mathbb{R}$ es Riemann Integrable (es decir, el supremo de las sumas inferiores es igual al ínfimo de las sumas superiores) $\iff \forall\ \varepsilon > 0\ \exists\ P$ partición de $[a, b]$ tal que $S(f, P) - I(f, P) < \varepsilon$.

Dem: hay $\Rightarrow$ y $\Leftarrow$.

$\Rightarrow$) Asumimos que f es integrable Riemman. Esto es, que la integral superior (ínfimo de las sumas superiores) es igual a la integral inferior (supremo de sumas inferiores). Esto significa $\int_a^b f = I^* = I_*$. Entonces $\int_a^b f - I(f, P)$ es tan chico como queramos, eligiendo un P adecuado. Lo mismo vale con $S(f, P) - \int_a^b f$. Llevamos esto a notación épsilon, sumamos las desigualdades, y llegamos.
Si querés quedar bien, elegí $\varepsilon/2$ antes de sumar.

$\Leftarrow$) Asumimos que dado $\varepsilon > 0$ siempre existe P que hace valer la desigualdad. Queremos llegar a que $I^* = I_*$. Es trivialmente cierto que $I(f, P) \leq S(f, P)\ \forall\ P$, entonces ya tenemos $I_* \leq I^*$. Para encontrar la otra desigualdad, escribimos $S(f, P) - I(f, P) < \varepsilon$ y acotamos por izquierda $I^* - I_*$. Obtenemos $I^* < I_* + \varepsilon \Rightarrow I^* \leq I_*$. ¡Bravo!

Continua $\Rightarrow$ integrable (en $\mathbb{R}$)

$f : [a, b] \to \mathbb{R}$ continua es integrable en $[a, b]$.

Dem: Queremos demostrar que f es integrable. Usamos lo anterior: queremos encontrar una partición P que, dado $\varepsilon > 0$, verifique $S(f, P) - I(f, P) < \varepsilon$.

Notamos que [a,b] es compacto (cerrado y acotado). Entonces f es uniformemente continua, y enunciamos lo que esto significa:

$$\forall\ \varepsilon_1 > 0\ \exists\ \delta > 0\ \big/\ 0 < |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon_1.$$

Ahora tomamos P una partición uniforme de norma $\frac{\delta}{2}$. Entonces para dos valores $x_i, y_i \in (t_i, t_{i+1})$, su distancia es menor que δ, y cumplen $|f(x_i) - f(y_i)| < \varepsilon_1$. Entonces tomo los x_i, y_i que hacen respectivamente máxima y mínima a la función en (t_i, t_{i+1}). Ya podemos acotar por ε_1:

$$\begin{aligned}
S(f, P) - I(f, P) &= \sum_i \Big(f(x_i) - f(y_i)\Big)(t_{i+1} - t_i) \\
&= \sum_i \Big|f(x_i) - f(y_i)\Big|(t_{i+1} - t_i) < \sum_i \varepsilon_1(t_{i+1} - t_i) = \varepsilon_1(b - a) \\
S(f, P) - I(f, P) &< \varepsilon_1(b - a)
\end{aligned}$$

Tomando $\varepsilon_1 = \dfrac{\varepsilon}{b - a}$ ganamos.

Teorema Fundamental del Cálculo

Sea $f : [a, b] \to \mathbb{R}$ continua, $F : [a, b] \to \mathbb{R}\ \big/\ F(x) := \int_a^x f(t)dt \Rightarrow F$ es continua en $[a, b]$, derivable en (a, b) y además se tiene que $F'(x) = f(x)$.

Dem: Tiene dos partes en **1.** probamos que F es continua en a y en b, y en **2.** encontramos la expresión de la derivada (así probando que es derviable en (a, b)).

1. Para ver que es continua en a queremos llegar a que $\lim_{x \to a^+} |F(x) - F(a)| = 0$. Primero, vemos que $F(a) = 0$ y ahora acotamos $|F(x)| \leq \int_a^x |f(t)| \leq S(|f|, P)$. Esta última cota es válida porque al ser f integrable, su integral es el ínfimo de las sumas superiores. Escribimos la definición de $S(|f|, P)$, acotamos cada M_i por el máximo de $|f|$ en $[a, b]$. Lo que queda tiende trivialmente a 0 cuando $x \to a^+$.

Para ver que es continua en b hay que hacer el mismo truco, escribiendo $|F(b) - F(x)|$ e invirtiendo los limites de integración del segundo término (notar que se juntan en una sola integral).

2. Tomamos x_0 un punto cualquiera de (a, b). Si queremos encontrar la derivada de F deberíamos llegar a:

$$\frac{F(x_0 + h) - F(x_0)}{h} = \frac{1}{h} \int_{x_0}^{x_0+h} f$$

Tomamos M y m, supremo e ínfimo de $f(t)$ para un $t \in (x_0, x_0 + h)$, con $h > 0$. Escribimos esa desigualdad, $f(t)$ ensandwichado entre m y M. Ahora integramos los tres miembros entre x_0 y $x_0 + h$. Por último, como $h \neq 0$, dividimos por h cada miembro. Ahora cuando $h \to 0^+$, M y m de f tienden a $f(x_0)$ por continuidad, y en el medio de la desigualdad queda escrito un límite lateral de la derivada de F. Para encontrar el otro, se toma $h < 0$, y se sigue el camino análogo (solo que al dividir por h, se invierten las desigualdades).