Formalización del concepto de área: La integral definida.

Jose María Sánchez Reales
notodoesmatematicas@hotmail.com

17 de marzo de 2020

Resumen

La integral definida es la herramienta matemática que viene a formalizar y dar rigor al concepto clásico de área. Aunque Newton y Leibniz ponen encima de la mesa el tema, y desarrollan la teoría que permite sistematizar el cálculo de integrales, será Riemann el primero en dar una definición rigurosa del concepto. El trabajo que aquí se presenta se divide en tres partes. En la primera parte se formaliza la integral definida en el sentido de Riemann; en la segunda parte se habla sobre la integración geométrica como origen del concepto de integral; y en la tercera parte, se ilustra una de las diversas apilicaciones de esta herramienta: el cálculo de longitudes.

Índice

Introducción

Una definición, comúnmente aceptada, de área, sería la que la define como la cantidad de cuadrados de lado unidad que caben en una región dada. Esta definición, que en un principio nos resulta sencilla, o intuitiva, desde el punto de vista de la geometría clásica, presenta ciertas complicaciones a la hora de ser formalizada con el rigor que requiere cualquier concepto matemático. Por ejemplo, ¿cuál sería el sentido de decir que el área de un disco de radio 1 es π cuadrados?. O, quizá más ilustrativo, sea el ejemplo utilizado por [10]: dada una función

$f : [0, 1] \to \{1, 2\}$, con $f(x) = 1$ si $x \in \mathbb{Q}$, y $f(x) = 2$ si $x \in \mathbb{R}/\mathbb{Q}$. ¿Esta función delimita un área?, y en caso de que la respuesta sea sí, ¿el área es 1, o es 2?.

Aunque como veremos en la segunda sección de este trabajo, el concepto de integral sea quizá uno de los conceptos más antiguos de la matemática, no se conseguirá desarrollar hasta la época de Newton y Leibnitz. Desde ese momento, la integral ha sido foco principal de estudio de muchos matemáticos, en la busqueda de comprensión, rigor, interpretación, y aplicabilidad. Riemann fue el primero en dar una definición formal del concepto de integral definida[1] (área), que es a día de hoy la más extendida por su utilidad en multitutd de aplicaciones comunes (cálculo de longitudes, áreas, volúmenes, centro de gravedad, centro de masas,...).

Este trabajo desarrolla tres secciones diferenciadas. En la sección 1 se introduce el concepto de integral de Riemann, su definición formal y sus propiedades, así como las condiciones que debe cumplir una función para poder ser integrable Riemann. Se terminará demostrando el Teorema Fundamental Del Cálculo Integral que sistematiza el cálculo de primitivas. En la sección 2, dedicada al desarrollo de un tema historico relacionado con el concepto, se hablará de los orígenes de la integral como un método geométrico de aproximación exhaustiva desarrollado por Eudoxo y llevado a la fama por Arquímedes. En la sección 3, que se dedica a exponer una aplicación actual de la integral definida, ilustramos cómo esta herramienta nos permite calcular distancias sobre lineas curvas.

1. Integral de Riemann.

Supongamos una función f tal que $f(x) \geq 0$[2] dentro de un intervalo $[a, b]$. Denotaremos por $R(f, a, b)$, y lo llamaremos *integral de f sobre* $[a, b]$, al área de la región limitada por esta función f, el eje horizontal, y las rectas verticales $x = a$ y $x = b$ (Figura 1). Supongamos además, que sobre el intervalo $[a, b]$ la función f tiene un valor mínimo, m, y un valor máximo, M. Resulta intuitivo decir que

$$m(b - a) \leq R(f, a, b) \leq M(b - a),$$

puesto que, tanto $m(b - a)$ como $M(b - a)$, pueden ser interpretados como el área de dos rectánculos situados sobre el intervalo $[a, b]$, y de alturas m y M, respectivamente.

Una mejor aproximación al valor de $R(f, a, b)$ se podría conseguir al utilizar rectánculos más finos en sustitución de un único rectángulo de base $b - a$. Supongamos que, por ejemplo,

[1]Cabe decir en este punto, que el concepto de integral es un concepto en constante desarrollo, inlcuso actualmente se avanza en su estudio buscando formas alternativas, quizá más generales, de definir, aplicar e interpretar esta herramienta. La integral de Riemann es, en cierto sentido, restrictiva, y no puede, por ejemplo, resolver el problema que planteabamos en el primer párrafo para una función con infinitas discontinuidades. Para ello es necesario la definición de integral que introdujo Lebesgue unos años después de Riemann. Como ejemplo de que la integración es un tema de investicación actual, podemos referenciar a [4] donde se propone una adaptación de la integral de Riemann-Stieltjes donde la integración se realiza, no con respecto a una variable o a una función, sino con respecto a un proceso estocástico.

[2]Suponer que f sea una función positiva nos va a permitir simplificar todo el discurso y expresar los resultados de forma más sencilla que si esta condición no existiera. Esto no quita generalidad a los resultados desde el momento en que cualquier función se puede expresar como la diferencia de dos funciones positivas. Si f es una función cualquiera, entonces $f = f^+ - f^-$, donde $f^+ = max\{f, 0\}$ y $f^- = max\{-f, 0\}$, y ambas funciones f^+ y f^- son positivas. Intuitivamente podemos entender que el área delimitada por f podrá obtenerse como la suma de las áreas de las regiones delimitadas por f^+ y f^-. Formalmente esto se garantizará en la sección 1.2, cuando se demuestren las propiedades de la integral definida.

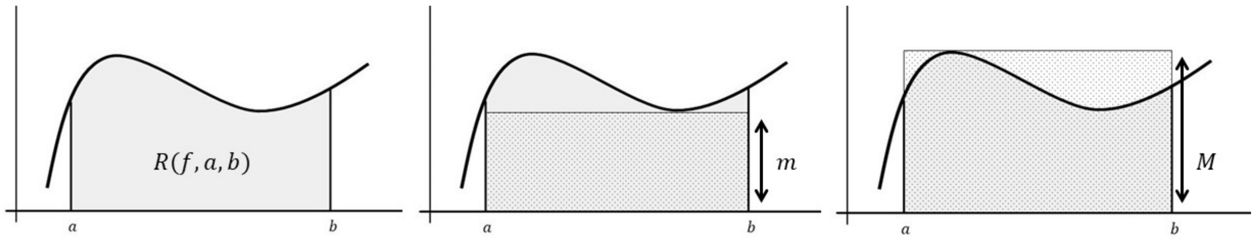

Figura 1: Área delimitada por una curva, el eje horizontal, y dos rectas verticales; y acotación robusta de dicho área a partir de dos cuadrádos.

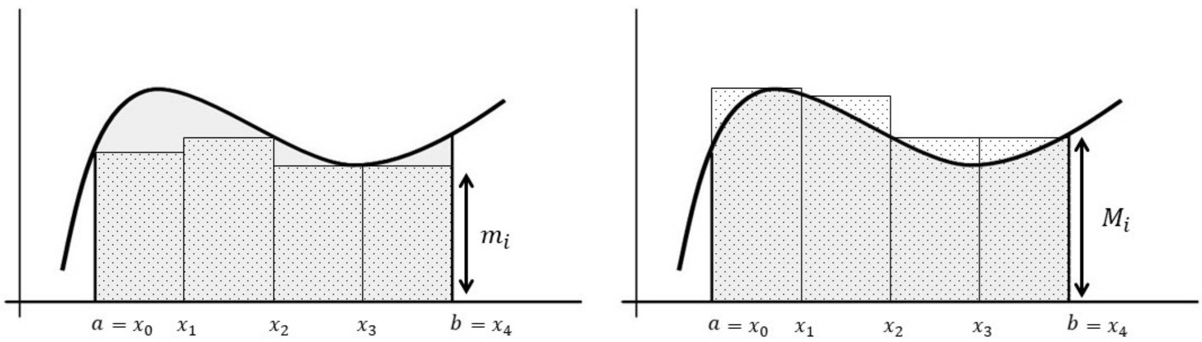

Figura 2: Acotación geométrica del área de una curva a partir de una partición del intervalo $[a, b]$.

se seleccionan los puntos $a = x_0 < x_1 < x_2 < x_3 < b = x_4$. En cada subintervalo $[x_{i-1}, x_i]$, la función f tendrá un valor mínimo, m_i, y un valor máximo, M_i (Figura 2). La suma de las áreas de los rectánculos para las alturas mínimas, resultará una cota inferior del valor de $R(f, a, b)$ -puesto que cada rectángulo estará enteramente contenido en la zona delimitada-, y de la misma manera, la suma de las áreas de los rectángulos para las alturas máximas, resultará una cota superior -puesto que cada rectángulo contiene enteramente una zona del área delimitada-. Así, podríamos decir que

$$\sum_{i=1}^{4}(x_i - x_{i-1})m_i \leq R(f, a, b) \leq \sum_{i=1}^{4}(x_i - x_{i-1})M_i,$$

y esto resulta intuitivo independientemente de cuales sean los puntos x_1, x_2, x_3 seleccionados dentro del intervalo $[a, b]$.

Definición 1.1. *Dados $a < b$, llamamos **partición** del intervalo $[a, b]$ a toda colección finita de puntos dentro del intervalo, de forma que uno de ellos es a, y otro es b. De forma común puede escribirse $a = x_0 < x_1 < x_2 < ... < x_n = b$. La partición $P = \{x_0, x_1, x_2, ..., x_n\}$ dará lugar a una división del intervalo $[a, b]$ en n subintervalos contenidos en él.*

Definición 1.2. *Sea $P = \{x_0, x_1, x_2, ..., x_n\}$ una partición del intervalo $[a, b]$, y sea t_i un punto de cada subintervalo determinado por dicha partición $t_i \in [x_{i-1}, x_i]$. Llamaremos **suma de Riemann** de la función f para la partición P al número*

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

Puesto que en cada subintervalo de una partición P existirá un valor que será mínimo, m_i, o máximo, M_i, de entre todos los posibles para la función dentro de ese intervalo, ha de suceder que $m_i \leq f(t_i) \leq M_i$, para todo $t_i \in [x_{i-1}, x_i]$. Esto nos permite definir una suma (de Riemann) superior y una inferior que serán únicas para cada partición.

Definición 1.3. *Supongamos que f es una función acotada sobre el intervalo $[a, b]$, y sea $P = \{x_0, x_1, x_2, ..., x_n\}$ una partición de $[a, b]$. Sean $m_i = inf\{f(x)|x \in [x_{i-1}, x_i]\}$, y $M_i = sup\{f(x)|x \in [x_{i-1}, x_i]\}$, los valores mínimo y máximo de la función en cada uno de los subintervalos determinados por la partición P, respectivamente. Llamaremos* ***suma inferior*** *de f para P, y la denotaremos por $s(f, P)$, a la suma*

$$s(f, P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}).$$

De forma análoga, llamaremos ***suma superior*** *de f para P, y la denotaremos por $S(f, P)$, a la suma*

$$S(f, P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}).$$

Notese, que la condición de f de estar acotada es indispensable para garantizar la existencia de un valor máximo y un valor mínimo, dentro de cada subintervalo delimitado por los puntos de la partición. Además, puesto que no se está exigiendo que la función f sea una función continua, estaremos obligados a definir los valores m_i y M_i como un infimo y un supremo, pues de definirlos como un mínimo y un máximo, podrían no existir. Esta definición generaliza la idea de acotación del valor de $R(f, a, b)$ por medio de la suma de las áreas de rectángulos contenidos o que contienen la zona delimitada por la curva de f en el intervalo $[a, b]$. Es intuitivo que

$$s(f, P) \leq R(f, a, b) \leq S(f, P)$$

para cualquier partición P del intervalo $[a, b]$; sin embargo, hemos de notar aquí que hasta ahora nos hemos limitado a decir que $R(f, a, b)$ es el área de una zona concreta, pero que esta «definición» carece de significado matemático. Para poder definir de manera rigurosa qué significa realmente el valor $R(f, a, b)$, hemos de demostrar primero algunos resultados.

Proposición 1.4. *Dadas dos particiones P y P' del intervalo $[a, b]$, de forma que P' es una partición más fina que P (es decir, $P \subset P'$), entonces se cumple que*

$$s(f, P) \leq s(f, P')$$

y

$$S(f, P) \geq S(f, P').$$

Demostración. Supongamos en primer lugar que P' es una partición que contiene exactamente un punto más que P, es decir, que

$$P = \{x_0, x_1, ..., x_n\}$$

$$P' = \{x_0, x_1, ..., x_{k-1}, y, x_k, ..., x_n\}.$$

Denotemos ahora

$$m_i = inf\{f(x)|x \in [x_{i-1}, x_i]\}$$

y tomemos como casos particulares los valores

$$m' = inf\{f(x)|x \in [x_{k-1}, y]\}$$

$$m'' = inf\{f(x)|x \in [y, x_i]\}.$$

Tememos que

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

$$s(f,P') = \sum_{i=1}^{k-1} m_i(x_i - x_{i-1}) + m'(y - x_{k-1} + m''(x_k - y) + \sum_{i=k+1}^{n} m_i(x_i - x_{i-1}).$$

Notese que el conjunto $\{f(x)|x \in [x_{k-1}, x_k]\}$ contiene todos los números del conjunto $\{f(x)|x \in [x_{k-1}, y]\}$, y también todos los números del conjunto $\{f(x)|x \in [y, x_k\}$, lo que obliga a que $m_k \leq m'$ y también $m_k \leq m''$. Esto nos lleva a que,

$$m_k(x_k - x_{k-1}) = m_k(y - x_{k-1}) + m_k(x_k - y) \leq m'(y - x_{k-1}) + m''(x_k - y).$$

Esto demuestra que $s(f,P) \leq (f,P')$, y una construcción análoga demostraría que $S(f,P) \geq S(f,P')$. Para demostrar el caso general en el que P' tiene exactamente n puntos más que P es suficiente proceder de forma recursiva, creando la sucesión de particiones P_i, con $1 \leq i \leq n$ de forma que P_{i+1} contiene exactamente un punto más que P_i, y donde se tendría que

$$s(f,P) = s(f,P_1) \leq s(f,P_2) \leq ... \leq s(f,P_n) = s(f,P')$$

$$S(f,P) = S(f,P_1) \geq s(f,P_2) \geq ... \geq S(f,P_n) = S(f,P').$$

□

Proposición 1.5. *Sean P_1 y P_2 dos particiones cualquiera del intervalo $[a,b]$, y sea f una función acotada sobre el intervalo dado. Entonces sucede que*

$$s(f,P_1) \leq S(f,P_2).$$

Demostración. Dadas dos particiones cualquiera P_1 y P_2 del intervalo $[a,b]$, la partición $P = P_1 \cup P_2$ contiene a la vez a P_1 y P_2. Se sigue entonces de la proposición 1.4 que

$$s(f,P_1) \leq s(f,P) \leq S(f,P) \leq S(f,P_2)$$

□

El teorema que se acaba de demostrar garantiza que cualquier suma superior $S(f,P)$ es una cota superior del conjunto de todas las sumas inferiores $s(f,P')$. Esto nos lleva de forma natural a que

$$sup\{s(f,P')\} \leq S(f,P), \tag{1}$$

independientemente de cuales sean las particiones P y P' del intervalo $[a,b]$. Un razonamiento análogo conducirá a deducir que

$$inf\{S(f,P')\} \geq s(f,P), \tag{2}$$

y que por lo tanto

$$sup\{s(f,P)\} \leq inf\{S(f,P)\}.$$

Dado un número x tal que $sup\{s(f,P)\} \leq x \leq inf\{S(f,P)\}$, entonces por (1) y (2) se cumplirá que

$$s(f,P') \leq x \leq S(f,P'),$$

para cualquier partición P' del intervalo $[a,b]$. El valor de $R(f,a,b)$, en caso de que exista, ha de ser uno de estos números. Esto nos permite definir el concepto de función integrable y, finalmente, dar una definición rigurosa de $R(f,a,b)$.

1.1. Función integrable.

Definición 1.6. *Sea f una función acotada en el intervalo $[a,b]$. Diremos que f es una **función integrable** sobre el intervalo $[a,b]$ si se cumple que*

$$sup\{s(f,P)\} = inf\{S(f,P)\},$$

para cierta partición P de $[a,b]$. En este caso, llamaremos integral de f sobre $[a,b]$, y denotaremos por $\int_a^b f$, a dicho valor único. La integral $\int_a^b f$ es por tanto, la formalización del concepto abstracto $R(f,a,b)$[3].

Supongamos que f es una función integrable, entonces sucederá que

$$sup\{s(f,P')\} \leq \int_a^b f \leq inf\{S(f,P')\},$$

para cualquier partición P' del intervalo $[a,b]$, y además $\int_a^b f$ es el único número con esa propiedad. Notese que esta definición nos dice qué es una función integrable, pero no nos dice si una función es integrable o no. Por ejemplo, supongamos la función constante $f(x) = k$ definida en un intervalo cualquiera $[a,b]$. En este caso, para cualquier partición del intervalo, y para cada subintervalo de esa partición, los valores máximo y mínimo en la definición 1.3 son constantes e igual a k. Por lo tanto, resulta evidente que

$$\int_a^b f = k(b-a).$$

Para ilustrar el extremo opuesto, podemos volver a la función que utilizabamos en la introducción para motivar la necesidad de formalizar el concepto de área, y definir $f : [0,1] \to \{1,2\}$ como

$$f(x) = \begin{cases} 1 & si \quad x \text{ es racional} \\ \\ 2 & si \quad x \text{ es irracional} \end{cases}$$

En este caso, para cualquier partición del intervalo $[0,1]$, cualquier subintervalo contendrá tanto un número racional, por lo que el mínimo en la definición 1.3 será siempre 1, como un número irracional, lo que nos llevará a que el máximo será siempre 2. Esto conduce a que

$$sup\{s(f,P)\} < inf\{S(f,P)\},$$

para toda partición P del intervalo $[a,b]$, por lo que la función no es integrable según la defininición 1.6. Un caso menos extremo que el propuesto, sería definir la función $f : [0,1] \to \{0,1\}$ como

$$f(x) = \begin{cases} 0 & si \quad x \neq 1/2 \\ \\ 1 & si \quad x = 1/2 \end{cases}$$

En este caso, para cualquier partición del intervalo $[0,1]$, contendrá un subintervalo $[x_{k-1}, x_k]$ que contenga el valor $1/2$. Esto nos lleva a que los valores mínimo y máximo de la definición 1.3, serán $m_i = M_i = 0$, para cualquier $i \neq k$; quedando $m_k = 0$, y $M_k = 1$. Esto nos lleva a que $s(f,P) = 0$ y $S(f,P) = x_k - x_{k-1}$. Cuanto más fina sea la partición P más cercano será el valor de $S(f,P)$ al valor de $s(f,P) = 0$. Esto nos lleva a que en el caso límite, cuando $n \to \infty$, entonces $S(f,P) \to s(f,P)$, y que por lo tanto, la función es integrable, y además $\int_0^1 f = 0$. Este último caso motiva la siguiete proposición.

[3] Notese que se exigió la condición de que $f(x) \geq 0$ para todo $x \in [a,b]$.

Proposición 1.7. *Sea f una función acotada sobre el intervalo $[a,b]$, entonces f es integrable sobre $[a,b]$ si y sólo si para todo $\epsilon > 0$ existe una partición P de $[a,b]$ tal que*

$$S(f,P) - s(f,P) < \epsilon. \tag{3}$$

Demostración. $\boxed{\Rightarrow}$ Supongamos en primer lugar que f es integrable en $[a,b]$. Esto implica, por definición, que $sup\{s(f,P)\} = inf\{S(f,P)\}$. Por lo tanto, dado $\epsilon > 0$ existen particiones P' y P'' tales que $S(f,P'') - s(f,P') < \epsilon$. Tomando ahora una parcitión que contenga a P' y P'' y aplicando las desigualdades de la proposición 1.4 tenemos que

$$S(f,P) - s(f,P) \leq S(f,P'') - s(f,P') < \epsilon.$$

$\boxed{\Leftarrow}$ Supongamos ahora que dado $\epsilon > 0$, existe un apartición P, tal que

$$S(f,P) - s(f,P) < \epsilon.$$

Sucede que por definición, $inf\{S(f,P')\} \leq S(f,P)$ y que $sup\{s(f,P'\} \geq s(f,P)$, siendo P' una partición cualquiera del intervalo dado. Esto implica que

$$inf\{S(f,P')\} - sup\{s(f,P')\} < \epsilon$$

y puesto que esta desigualdad es cierta para todo $\epsilon > 0$, ha de seguirse que

$$inf\{S(f,P')\} = sup\{s(f,P')\}.$$

$\square$

Esta proposición supone una reformulación de la definición 1.6, y este paso al límite va a permitir determinar las integrales de funciones que, aun siendo continuas, van a presentar más dificultades prácticas que las que presentaban los ejemplos anteriores.

Supongamos la función $f(x) = x$ definida en el intervalo $[0,b]$ ($b > 0$), y sea $P = \{x_0, x_1, ..., x_n\}$ una partición de dicho intervalo. Entonces, los valores mínimo y máximo de la definición 1.3 quedarán determinados por $m_i = x_{i-1}$ y $M_i = x_i$, para todo i. De esta manera,

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

$$S(f,P) = \sum_{i=1}^{n} x_i(x_i - x_{i-1}), \tag{4}$$

que no tienen una simplificación aparente. La proposición 1.7 nos garantiza que una función es integrable si existe una partición para la cual se cumpla (3). Podemos entonces particularizar la partición P elegida, para obligar a que todos los subintervalos que determina sean del mismo tamaño. Para ello, podemos dividir el intervalo $[0,b]$ en n subintervalos iguales, dando lugar a una partición $P_n = \{x_0, x_1, ..., x_n\}$, donde la amplitud de cada subintervalo es b/n, y $x_i = ib/n$. Esto nos permite simplificar las expresiones en (4) de forma que

$$s(f,P_n) = \sum_{i=1}^{n} \frac{(i-1)b}{n}\frac{b}{n} = \frac{b^2}{n^2}\sum_{i=1}^{n}(i-1) = \frac{b^2}{n^2}\frac{(n-1)n}{2} = b^2\frac{n-1}{2n},$$

$$S(f,P_n) = \sum_{i=1}^{n} \frac{ib}{n}\frac{b}{n} = \frac{b^2}{n^2}\sum_{i=1}^{n} i = \frac{b^2}{n^2}\frac{(1+n)n}{2} = b^2\frac{n+1}{2n}, \tag{5}$$

y de estas expresiones obtenemos que

$$S(f,P_n)-s(f,P_n)=b^2(\frac{n+1}{2n}-\frac{n-1}{2n})=\frac{b^2}{n},$$

por lo que dado $\epsilon>0$, bastará tomar un valor de n suficientemente grande para que $\frac{b^2}{n}<\epsilon$ para que se cumpla la condición (3). Por lo tanto la función es integrable. Además, haciendo que $n\to\infty$ en (5), tendremos que tanto $s(f,P_n)$, como $S(f,P_n)$, se acercan indefinidamente a $\frac{b^2}{2}$. Esto nos lleva a que se cumple que

$$s(f,P_n)\le\frac{b^2}{2}\le S(f,P_n),$$

para todo n y que, por lo tanto, $\int_0^b f=\frac{b^2}{2}$.

Conviene en este punto que formalicemos un resultado, cuya demostración resulta evidente desde todo lo descrito hasta ahora, pero que nos será de gran utilidad para simplificar futuras demostarciones.

Teorema 1.8. *Sea f una función integrable en el intervalo $[a,b]$, y sea P_n una sucesión de particiones del intervalo $[a,b]$, tal que la mayor de las longitudes de los subintervalos de P_n tiende a 0 cuando n se hace indefinidamente grande, se verifica que*

$$lim_{n\to\infty}s(f,P_n)=lim_{n\to\infty}\sigma(f,P_n)=lim_{n\to\infty}S(f,P_n)=\int_a^b f \qquad (6)$$

El cálculo de integrales, tal cual se ha presentado hasta ahora, hace que en la práctica sea imposible de determinar para la mayoría de las funciones. Este problema se resolverá gracias a los dos Teoremas Fundamentales del Cálculo, que nos permiten sistematizar el cálculo de integrales definidas a partir de funciones primitivas (funciones cuya derivada es la función integrable). Sin embargo, antes de desarrollar esos apartados, parece oportuno estudiar las condiciones más elementales que garanticen que una función sea integrable, así como algunas de las propiedades más importantes de esta integral.

Teorema 1.9. *Si f es continua en el intervalo $[a,b]$, entonces f es integrable en $[a,b]$.*

Demostración. Toda función continua en un intervalo cerrado y acotado, es una función uniformemente continua[4]. Esto es, dado $\epsilon>0$, exite un $\delta>0$, tal que para cualquiera $x,y\in[a,b]$ con $|x-y|<\delta$, entonces $|f(x)-f(y)|<\epsilon/(b-a)$. Sea $P=\{x_0,x_1,...,x_n\}$ una partición del intervalo $[a,b]$ tal que $|x_k-x_{k-1}|<\delta$. Existen puntos $z_k,t_k\in[x_{k-1},x_k]$ en los cuales f alcanzará su máximo y su mínimo dentro del intervalo $[x_{k-1},x_k]$. De esta manera,

$$S(f,P)-s(f,P)=\sum_{k=0}^{n}(f(z_k)-f(y_k))(x_k-x_{k-1})<\frac{\epsilon}{b-a}\sum_{k=0}^{n}(x_k-x_{k-1})=\epsilon,$$

por lo que, por la proposición 1.7, demuestra que f es integrable en $[a,b]$.

□

Teorema 1.10. *Si f tiene una cantidad finita de discontinuidades en el intervalo $[a,b]$, entonces f es integrable en $[a,b]$.*

[4]Una demostración de este resultado (conocido algunas veces como Teorema de Heine) puede consultarse en [2], p. 191, Teorema 1.

Demostración. Supongamos que f es continua en el intervalo abierto (a,b) y que está acotada en el intervalo cerrado $[a,b]$. Puesto que f es una función acotada, existe $M>0$ de forma que $f(x)<M$ para todo $x\in[a,b]$. Dado $\epsilon>0$, definimos el intervalo $[c,d]$ de forma que $a<c<d<b$, $c-a<\epsilon/(3M)$ y $b-d<\epsilon/(3M)$. Por el teorema 1.9, f es integrable en $[c,d]$, por lo que existe una partición P' de dicho intervalo de forma que $S(f,P')-s(f,P')<\epsilon/3$. Sea P la partición P' ampliada con los puntos a y b, entonces se tiene que

$$S(f,P)-s(f,P)\leq(c-a)M+S(f,P')-I(f,P')+(b-d)M<\epsilon,$$

lo que demuestra, por la proposición 1.7, que f es integrable en $[a,b]$.

Si suponemos ahora que f está acotada en $[a,b]$ y que tiene una cantidad finita de discontinuidades $c_1<c_2<...<c_p$ dentro del intervalo, entonces, f es integrable en cada uno de los intervalos $[a,c_1]$, $[c_k,c_{k+1}]$ $(k=1,2,...,p)$, y $[c_p,b]$. Por lo tanto es integrable en la unión de todos ellos, el intervalo $[a,b]$.

$\square$

Un resultado más general que el expuesto en el teorema 1.10, es el conocido como *criterio de Lebesgue* para la integrabilidad de Riemann, que nos garantizará que una función es integrable Riemann sí y sólo sí el conjunto de las discontinuidades de la función tiene medida cero. La demostración de este resultado requerirá de conceptos más avanzados relativos a la teoría de la medida, y queda, por lo tanto, fuera del alcance de este trabajo. Se recomienda al lector interesado consultar ([9], p. 15, Teorema 1.10) para una aproximación más formal. Aquí nos conformaremos con estos resultados que nos garantizan condiciones suficientes de integrabilidad, y dotan de utilidad a la integral de Riemann para un amplio abanico de funciones.

Corolario 1.11. *Si f y g son dos funciones que coinciden en todos los puntos de un intervalo $[a,b]$ excepto en una cantidad finita de ellos, entonces se cumple que f es integrable en $[a,b]$ si y sólo si g es integrable en $[a,b]$. En dicho caso, se cumple que*

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

Demostración. Sea $h=f-g$. Puesto que por hipótesis $f=g$ en todos los puntos del intervalo $[a,b]$ excepto en una cantidad finita de ellos, necesariamente $h=0$ en todos los puntos del intervalo $[a,b]$ excepto en una cantidad finita de ellos. Esto implica que h es continua en $[a,b]$ excepto en una cantidad finita de puntos, y que, por el teorema 1.10, es por lo tanto integrable en $[a,b]$. Puesto que h es identicamente nula en todos los puntos de $[a,b]$ excepto en una cantidad finita de ellos, sucederá que $\int_a^b h=0$.

Demostrar que $\int_a^b f=\int_a^b g$ se sigue inmediatamente de la propiedad de linealidad de la integral definida en la proposición 1.12.

$\square$

Este resultado es de gran importancia, puesto que nos permite relajar las condiciones sobre el dominio de una función a la hora de estudiar su integrabilidad. Si una función no existe en una cantidad finita de puntos de un intervalo dado, esto no afectará a que sea o no sea una función integrable. En particular, garantiza que cualquier función continua y acotada en $[a,b]/\{a_1,a_2,....,a_n\}$, con $a_1,a_2,...,a_n\in[a,b]$, será integrable en $[a,b]$[5].

[5] Por ejemplo, la función $f(x)=sin(1/x)$ no está definida para $x=0$. Para poder estudiar la integrabilidad de esta función en el intervalo $[0,1]$ nos bastará definir $f(0)=k$, lo que hará de f una función continua en $(0,1)$ y acotada en $[0,1]$, y por lo tanto integrable.

1.2. Propiedades de las integrales definidas.

A continuación nos limitaremos a listar una serie de propiedades de las integrales definidas que son de gran utilidad en el cálculo efectivo. Algunas de estas demostraciones pueden resultar evidentes desde el resultado expuesto en el teorema 6 y otras necesitaran un poco más de trabajo. Demostraciones de estos resultados se pueden encontrar en cualquier guía básica de cálculo, en particular podría encontrarlas en [1] (Capítulo 7) ó en [2] (Capítulo 13).

Proposición 1.12. (La integral es una operación lineal) Si f y g son dos funciones integrables en el intervalo $[a, b]$ y α y β son números reales, se cumple que la función $\alpha f + \beta g$ también es integrable en el intervalo $[a, b]$, y que

$$\int_a^b \alpha f + \beta g = \alpha \int_a^b f + \beta \int_a^b g.$$

Proposición 1.13. (La integral conserva el orden) Si f y g son dos funciones integrables en el intervalo $[a, b]$ y $f(x) \leq g(x)$ para todo $x \in [a, b]$, entonces se cumple que

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

En particular, si f es integrable en el intervalo $[a, b]$, y $m \leq f(x) \leq M$ para todo $x \in [a, b]$, entonces se cumple que

$$m(b-a) \leq \int_a^b f(x)dx \leq M(b-a).$$

Proposición 1.14. (Valor absoluto) Si f es una función integrables en el intervalo $[a, b]$, entonces $|f|$ también es integrable en el intervalo $[a, b]$ y se cumple que

$$\left| \int_a^b f \right| \leq \int_a^b |f|.$$

Proposición 1.15. (Producto de funciones) Si f y g son dos funciones integrables en el intervalo $[a, b]$, entonces $f \cdot g$ es una función integrable en el intervalo $[a, b]$.

Proposición 1.16. (La integral es aditiva en intervalos consecutivos) Sea $c \in [a, b]$. Una función f es integrable en el intervalo $[a, b]$ si y sólo si es integrable en los intervalos $[a, c]$ y $[c, b]$, y en ese caso se cumple que

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

1.3. Teoremas fundamentales del cálculo.

Sea f una función integrable sobre el intervalo $[a, b]$. Definimos una nueva función F, y la llamamos función área de f en $[a, b]$, de la forma

$$F(x) = \int_a^x f(t)dt, \quad x \in [a, b]. \tag{7}$$

Proposición 1.17. *La función área, F, de una función f integrable en el intervalo $[a, b]$, es una función continua.*

Demostración. Puesto que f es integrable en el intervalo $[a, b]$, ha de estar acotada en dicho intervalo, esto es, ha de existir $M > 0$ de forma que $|f(x)| \leq M$, para todo $x \in [a, b]$. Sean $x, y \in [a, b]$. Suponiendo que $x < y$, entonces

$$|F(y) - F(x)| = \left| \int_x^y f(t)dt \right| \leq^{6} \int_x^y |f(t)|dt \leq M(y - x),$$

y de forma análoga, suponiendo que $y < x$, se obtiene que $|F(x) - F(y)| \leq M(x - y)$, lo que resulta en que

$$|F(x) - F(y)| = M|x - y|,$$

para todo $x, y \in [a, b]$, por lo que F es continua en dicho intervalo.

□

Hasta ahora hemos supuesto que f es integrable sobre un intervalo dado. Si además, la función f es continua, entonces tenemos uno de los teoremas más importantes del cálulo, el Primer Teorema del Cálculo integral, que nos garantiza que, bajo esta condición, la función F es derivable, y que su derivada es precisamente la función f.

Teorema 1.18. ***(Teorema Fundamental del Cálculo)*** *Si f es una función continua en el intervalo $[a, b]$, entonces la función de área*

$$F(x) = \int_a^x f$$

con $x \in [a, b]$ es derivable y además

$$F'(x) = f(x),$$

para todo $x \in [a, b]$.

Demostración. Sea c un punto del intervalo $[a, b]$ donde f es continua. Sea

$$\frac{F(x) - F(c)}{x - c} - f(c) = \frac{F(x) - F(c) - (x - c)f(c)}{x - c} = \frac{\int_c^x (f(t) - f(c))dt}{x - c}. \tag{8}$$

Puesto que f es continua en c, dado $\epsilon > 0$, existe $\delta > p$ tal que para cualquier $t \in [a, b]$ con $|t - c| < \delta$, entonces $|f(t) - f(c)| < \epsilon$. Sea $x \in [a, b]$ tal que $|x - c| < \delta$, entonces, para todo t entre x y c, sucede que $|t - c| < \delta$, lo que conlleva que $|f(t) - f(c)| < \epsilon$. De esta manera,

$$\left| \int_c^x (f(t) - f(c))dt \right| \leq \epsilon |x - c|.$$

Tomando ahora valores absolutos en (8), tendremos

$$\left| \frac{F(x) - F(c)}{x - c} - f(c) \right| = \frac{|\int_c^x (f(t) - f(c))dt|}{|x - c|} \leq \frac{\epsilon |x - c|}{|x - c|} = \epsilon,$$

lo que es equivalente a decir que $F'(c) = f(c)$.

□

Inmediatamente se deduce el segundo Teorema Fundamental del Cálculo integral, comunmente conocido como la Regla de Barrow, que sistematiza el cálculo del área tal cuál se prometía en una sección anterior.

[6] Por la proposición 1.14

Teorema 1.19. ***(Segundo Teorema Fundamental del Cálculo)*** *Si f es una función continua en el intervalo $[a, b]$, y $f = g'$ para alguna función g, entonces*

$$\int_a^b f = F(b) - F(a).$$

Demostración. Definimos $F(x)$ como en (7). Entonces $F' = f$ por el teorema 1.18, de donde tenemos que $F = g + c$, siendo c una constante que puede deducirse de forma sencilla de la igualdad $0 = F(a) = g(a) + c$. De esta manera, podemos decir que

$$F(x) = g(x) - g(a),$$

para cualquier x en el intervalo $[a, b]$, en particular, para $x = b$.

□

1.4. Conclusión de la parte teórica.

En este trabajo se ha presentado la teoría que formaliza el concepto clásico y algo difuso de área. Esta formalización se lleva a cabo a través de una herramienta fundamental en el cálculo, y sin miedo a ser exagerados podriamos decir que en toda la matemática moderna, como es la integral definida. Se ha presentado aquí la formalización de área que desarrollo Riemann (cierto que con algunos toques de Darboux, que hacen más sencilla la exposición) como caso límite para la suma de áreas geométricas que acotan la región limitada por una curva cualquiera. Se han visto algunas condiciones, que aun siendo parciales resultan suficientes para la mayoría de aplicaciones, que garantizan la integrabilidad de una función. Se han enunciado las propiedades más importantes de la integral definida de Riemann y se han terminado demostrando los Teoremas Fundamental del Cálculo, gracias a los cuales sistematiza el cálculo de primitivas y el cálculo de áreas.

La forma en que el trabajo queda expuesto demuestra sobradamente que la integral es un concepto y una herramienta con identidad propia e independiente de cualquier otra. Sin embargo, la forma en que se presenta generalmente al estudiante de secundaria, hace que ésta sea vista como una mera operación inversa de la derivada. Para ilustrar lo equivocado de esta idea, en la siguiente sección se presenta el marco histórico en el cuál se basa la idea de Riemann para formalizar el concepto de área. Hemos llamado a este apartado *la integral geométrica*, para distinguirlo del concepto de integral que tenemos desde el cálculo mecánizado de primitivas. Si pensamos en la integral como una suma infinita de áreas de figuras geométricas conocidas, entonces la integral no es una herramienta que nació en 1800, sino en la antigua Grecia, unos 400 años antes de nuestra era.

2. La integral geométrica.

La derivada y la integral son las dos herramientas sobre las que descansa todo un campo de conocimiento -uno con mucho peso específico- de la matemática: el cálculo. En contra de lo que pueda intuir el principiante, estos dos conceptos no surgen motivados por un problema puro del análisis matemático, zona en la que encasillariamos el cálculo de forma natural, sino que son consecuencia de problemas geométricos clásicos y, en un principio, sin relación aparente entre si.

La forma en la que se presenta el cálculo de integrales en la educación secundaria, conduce a una concepción general del concepto de integral como el de mera «operación inversa», es decir, una integral no sería más que la consecuencia esperable de la existencia de la derivada. De la misma manera que concebimos la resta como operación inversa de la suma, o la división como la operación inversa del producto, parece que el cálculo integral surge de la pregunta evidente sobre qué función habría de derivar para obtener otra cierta función dada. Esta concepción es erronea y mal infundada. Los problemas que motivan el cálculo diferencial (la derivada) y el cálculo integral (la integral), son problemas geométricos y de distinta naturaleza: el cálculo de la tangente a una curva cualquiera, y el área delimitada por una curva en un intervalo dado.

La idea clave es, hablando de manera más intuitiva que formal, que la integral indefinida sistematiza el cálculo de primitivas. La primitiva, a veces llamada anti-derivada, es el objeto matemático necesario para generalizar la utilidad de la integración, es decir, a la hora de aplicar el concepto a la resolución práctica de problemas a través de la integral definida. Sin embargo, no podemos decir que una integral, en su sentido más amplio, es una primitiva. Una integral es, definitivamente, una forma de medir, medir un área, un volumen, o la longitud de una curva. Por cómo se presenta, esto parece una consecuencia, una aplicación del cálculo integral, sin embargo, la realidad no es exactamente así. Lo que parece una consecuencia, resulta que es su motivación principal.

2.1. La integral de Riemann desde un punto de vista intuitivo.

Bernhard Riemann (1826-1866), grande entre los grandes, fué el primero en formalizar de manera rigurosa el concepto de integral, y curiosamente, esto no sucede hasta casi 200 años después de que Isaac Newton (1642-1727), Gottfried Leibnitz (1646-1716), Isaac Barrow (1630-1677), y no deberíamos olvidar a Pierre Fermat (1601-1665), desarrollaran la idea intuitiva de integral y sistematizaran el cálculo de primitivas. Recordemos que el Teorema Fundamental del Cálculo, que establece que efectivamente la derivada y la integral son en algún sentido «operaciones» naturalmente inversas, fué demostrado, en su forma más generalizada, por Barrow[7]. Los trabajos de Newton y Leibnitz proporcionan una sistematización del cálculo integral, pero carecen del rigor exigido a la formalización de cualquier concepto matemático. Podemos pensar que actualmente, tanto la derivada como la integral, se definen como un límite, que utilizando la notación que se ha empleado en la primera parte de este trabajo, pueden expresarse como,

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}, \tag{9}$$

$$\int_a^b f(x)dx = \lim_{n \to \infty} \sum_{k=0}^{n-1} \frac{b-a}{n} f(a + \frac{b-a}{n}k), \tag{10}$$

El concepto de límite y su formalización no se darán hasta la llegada de Augustin Cauchy (1789-1857), y comparando fechas podemos ver, que esto no sucederá hasta unos cuantos años después de la muerte de los protagonistas anteriormente nombrados. Esto puede darnos una idea de que las definiciones de derivada y de integral, que es la que particularmente nos interesa aquí, dadas por Newton y Leibniz, eran más una intuición que una formalizaicón. Eso sí, una vez formalizada la idea de límite, ya tenemos todo lo neceario para que trabaje Riemann.

[7]Una primera demostración de una versión más restringida del teorema se debe a James Gregory (1638-1675).

La idea detrás de la cuál descansa la formalización de integral definida, desarrollada por Riemann, es lógica e intuitiva. Se tiene una curva cualquiera[8], que podemos denotar por $f(x)$, cuyo dominio está limitado por dos extremos -extremos que definen lo que conocemos comunmente como intervalo de integración, y que denotaremos por $[a,b]$-, y se quiere determinar el área que delimita, que después de lo que se ha expuesto en el apartado anterior, resultara natural que denotemos por $\int_a^b f(x)dx$. Una forma de aproximar el valor de este área sería utilizar un rectángulo (ver Figura 3), ya que desde la geometría clásica es sencillo determinar el área de este tipo de figura plana. Se trataría, por tanto, de elegir el rectánculo que pueda suponer una buena aproximación. La base será, de forma eviente, la que coincide con el intervalo de integración, pues en cualquier otro caso no estaríamos cubriendo toda la figura. Para la altura, tendrémos que conformarnos con decir que ha de estar acotarda entre los valores máximo y mínimo de la curva, y que realmente, en un principio, no es posible saber qué altura daría una mejor aproximación al área verdadera. Es cierto que esta aproximación será, en general, bastante burda, y si se pretende encontrar una mejor aproximación se ha de mejorar en algo la forma en que «rellenamos» la zona de la cuál queremos conocer el área. Para ello se podría dividir el intervalo de integración en dos trozos y aplicar la misma idea a cada uno, es decir, utilizaría el área de dos rectángulos como una aproximación al área de la curva, en cada uno de los trozos que hemos definido. Si se procede de esta manera, haciendo cada vez una división más fina del intervalo de integración, el error que estoy cometiendo será menor, puesto que cada vez los rectángulos ajustan mejor a la superficie objetivo. De esta manera, para una división del intervalo de integración en n subintervalos, que por comodidad podemos suponer iguales y de longitud $(b-a)/n$, el área quedará aproximada por

$$\sum_{k=0}^{n-1} \frac{b-a}{n} f(a + \frac{b-a}{n} k)$$

El caso límite soluciona el problema y me conduce a la expresión (10).

La idea que utiliza Riemann aquí, es la misma que utiliza el griego Eudoxo (408-355 aC), y que más tarde adaptó Arquímedes (287-212 aC) para aproximar el área de figuras curvas[9]. Actualmente se conoce como método de exahución de Eudoxo, método de agotamiento de Arquímedes, o método de integración geométrica.

2.2. El método de exahución de Eudoxo.

El método que propone Eudoxo es, de nuevo, increiblemente intuitivo y tremendamente sencillo. Imaginemos que queremos calcular el área de la zona delimitada por una circunferencia. De alguna manera, ésta será más pequeña que el área del cuadrado tangente que la contiene, y más grande que el área del cuadrado tangente contenido en ella (ver figura 4). El área de un cuadrado se puede calcular fácilmente, y por lo tanto sabemos obtener una acotación para el área de la circunferencia. Ésta acotación puede resultarnos bastante mala, y, puesto que puede resultar intuitivo, que al añadir un lado más a nuestra figura podemos ajustar mejor el polígono a la circunferencia, se podría utilizar un péntagono, en sustitución del cuadrado, para mejorar esta acotación. En general, cada vez que añado un lado a la figura poligonal plana, ésta acotará mejor el área de la zona delimitada por la circunferencia. Esto

[8] «Cualquiera» en el sentido del lenguaje común. Como hemos visto en la primera sección, una función ha de cumplir ciertas condiciones para ser integrable.

[9] Como nos recuerda [12], no es realmente un método de exhaución, ya que no se agota un proceso, sino que lo que el método realmente pretende es evitar el infinito. Como propone, un nombre más adecuado sería quizá *Principio de Convergencia de Eudoxo*. Actualmente se reconoce el método de exahución griego como los origines del cálculo integral, y se le suele llamar informalmente integración geométrica.

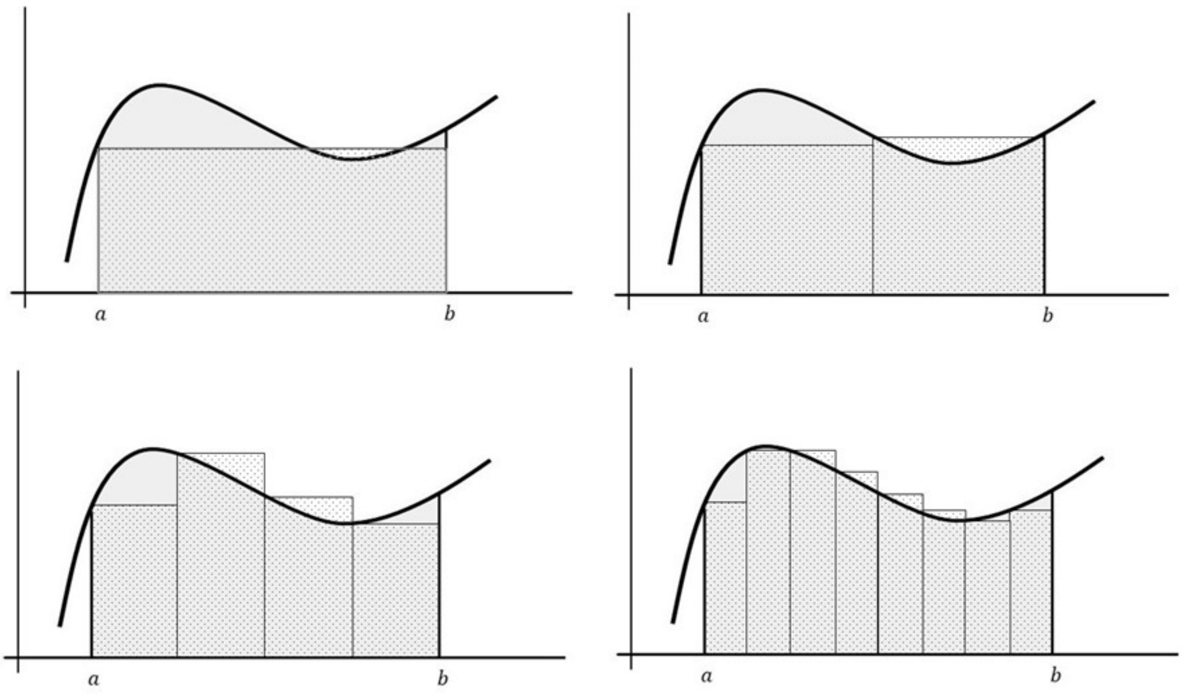

Figura 3: Idea intuitiva de suma de Riemann para el cálculo del área bajo una curva.

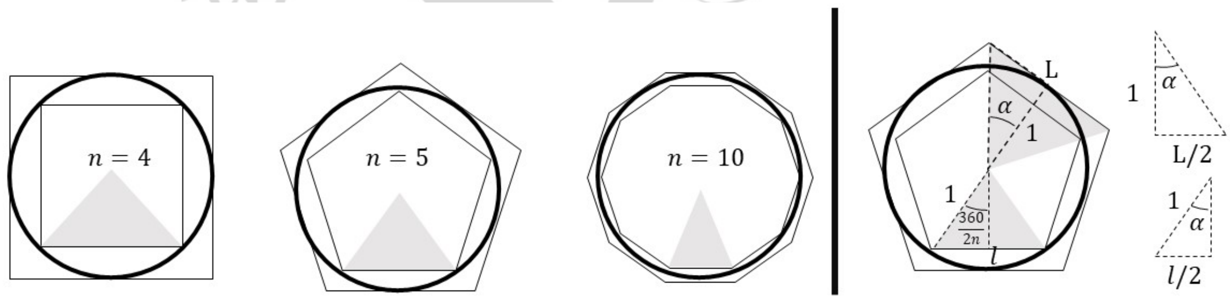

Figura 4: Izquierda: Aproximación de la circunferencia por medio de polígonos regulares. Derecha: Determinación de los valores notables de cada triángulo que determina el polígono regular, a partir del número de lados y del radio de la circunferencia.

permite generalizar el proceso, pero requiere de la capacidad de calcular el área, de forma genérica, de cualquier polígono regular. La forma más sencilla sea quizá, la de imaginar un polígono regular compuesto por triángulos isósceles, cuya base y altura coinciden con el lado y con el apotema del polígono. Puesto que desde la geometría plana clásica, sabemos que el área de un triángulo es la mitad de la base, b, por la altura, h, y que un polígono regular de n lados estará formado por exáctamente n triángulos iguales, el área de dicho polígono se podrá calcular como

$$\frac{b \cdot h}{2} n$$

Para un valor de n suficientemente grande, e independientemente de si se aproxima por exceso (utilizando el polígono exterior) o por defecto (a través del polígono interior), dicho área puede ser considerada una buena aproximación del área del círculo. Además, el caso límite daría el valor exacto del área de la circunferencia.

Notese que en este caso estamos hablando de acotación y no de determinación de un

n	3	4	5	10	20	50	100
a(n)	1,299	2	2,378	2,939	3,090	3,133	3,140
A(n)	5,196	4,000	3,633	3,249	3,168	3,146	3,143

Tabla 1: Valores del área (redondeada a tres cifras decimales) del polígono regular inscrito, $a(n)$, y del polígono regultar cicunscrito, $A(n)$, en función del número de lados, n.

valor. Esto se debe a que en la antigua Grecia el concepto de infinito era deliberadamente evitado[10], y se conformarían con una aproximación que fuera «suficientemente» buena.

Calcular el área de un polígono regular de forma sistemática a partir de la suma de las áreas de n triángulos es relativamente sencillo. Utilizando trigonometría básica, podemos determinar los valores del lado del triángulo para el caso del triángulo interior y del triángulo exterior, respectivamente. Si observamos la Figura 4 (derecha), en cada caso podemos determinar un triángulo rectángulo para el cuál el radio de la circunferencia, en este caso 1, es uno de los valores de sus lados. El ángulo interno de cada uno de los triángulos corresponde a la mitad de los 360° de la circunferencia divididos entre el número de lados del polígono regular, es decir, $180°/n$. De esta manera, tenemos que el cateto correspondiente al lado del triángulo inerior (que determina el lado del polígono regular inscrito en la circunferencia), tiene un valor de $l = 2\sin(180°/n)$, y el correspondiente al triángulo exterior (que determina el lado del polígono regular circunscrito en la circunferencia), $L = 2\tan(180°/n)$.

El área de cada uno de estos triángulos queda entonces determinada a partir del número de lados del polígono regular que determinará de la forma:

- Polígono inscrito: $a(n) = \frac{nla_p}{2} = n\sin(\frac{180}{n})\cos(\frac{180}{n})$
- Polígono circunscrito: $A(n) = \frac{nLa_p}{2} = n\tan(\frac{180}{n})$

donde a_p denota la altura de cada uno del triángulos[11] (equivalente al apotema del polígono regultar que determina). En la tabla 1 se muestran los valores de estas áreas para diferentes valores de n. Puede observarse cómo, cuanto mayor es el valor de n (más lados tiene el polígono regular utilizado para aproximar la circunferencia), mejor es la aproximación del valor de π (que evidentemente es el área de una circunferencia de radio 1).

3. Aplicación de la integral definida: longitud de curva.

De igual manera que la geometría clásica nos ofrece multitud de «fórmulas» que nos permiten calcular el área de figuras planas, también nos ofrece otras tantas que nos permiten calcular longitudes de curva[12]. Sin embargo, existen limitaciones a la hora de calcular

[10]Podemos recordar a Zenón y el miedo que infundaba en los sabios antiguos con sus paradojas sobre el infinito.

[11]El mismo razonamiento que nos permite determinar el valor de los lados, l y L, de los triángulos a partir de las razones trigonométricas correspondientes, nos va a determinar que $\cos(\frac{180°}{n})$ corresponde a la apotema del triángulo inscrito.

[12]El cálculo de la lóngitud de un camino recto no supone un gran problema: $d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$. Si se sustituyera el segmento lineal por un arco de circunferencia, su longitud también podría ser determinada a partir de la parte proporcional del perímetro de una circunferencia, $L = 2\pi r\alpha/360$, en función del radio, r, y el ángulo que determina el arco, α. Podemos intuir que existan otras expresiones más complejas para determinar longitudes de algunos otros tipos de curva. Sin embargo, cuando queremos generalizar el cálculo de la longitud de una curva, no disponemos de otra herramienta que la propia integral.

la longitud de un segmento cualquiera[13]. Utilizando un razonamiento análogo al que nos ha permitido definir la integral como el área de una zona delimitada por una curva «cualquiera», podemos ahora formalizar el concepto de *longitud de curva* desde el cálculo integral.

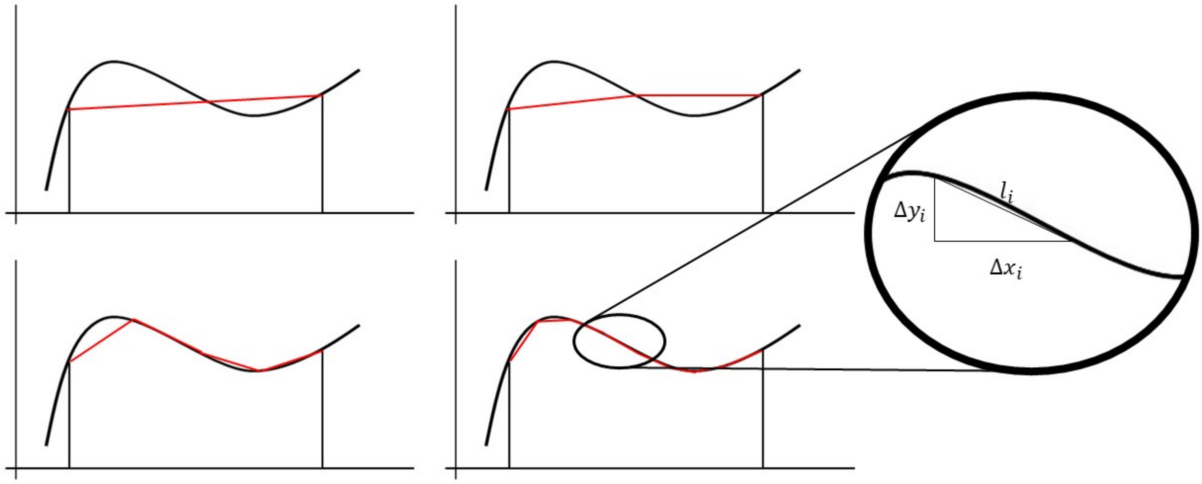

Figura 5: Aproximación discreta a la longitud de una curva por medio de un polinomio interpolador.

Supongamos que se quiere determinar la longitud de una curva, $f(x)$, en un intervalo dado, $[a, b]$. Una forma de estimar dicha longitud sería a través del cálculo de la longitud de n rectas interpoladoras que pudieran aproximar la trayectoria de la curva dada. Para ello, pueden definirse $n+1$ puntos dentro del intervalo de la forma $x_i = a + i \cdot (b-a)/n$ que determinan las imágenes $y_i = f(x_i)$, con $i = 0, 1, 2, ..., n$. Toda pareja de puntos consecutivos (x_i, y_i) y (x_{i+1}, y_{i+1}) determina un segmento de recta de longitud

$$l_i = \sqrt{(x_{i+1}-x_i)^2 + (y_{i+1}-y_i)^2} = \sqrt{\Delta x_i^2 + \Delta y_i^2} = \sqrt{1 + (\frac{\Delta y_i}{\Delta x_i})^2}\Delta x_i.$$

La suma de las longitudes de todos estos segmentos de recta, coincidirá con la longitud de la curva f,

$$\tilde{L} = \sum_{i=0}^{n-1} \sqrt{1 + (\frac{\Delta y_i}{\Delta x_i})^2}\Delta x_i$$

cuyo paso al límite, haciendo $\Delta x_i \to 0$, nos lleva a que

$$L = \int_a^b \sqrt{1 + f'(x)^2}dx. \tag{11}$$

3.1. Motivación 1: la linea curva como elemento fundamental en la construcción.

La linea recta es el camino más corto, pero no es el camino más fuerte. Esto implica, que en un diseño arquitectónico destinado a soportar cierta cantidad de carga, la forma recta no es la más adecuada. Por esta razón, es muy común que cuando queremos salvar el vano de una puerta o una ventana, cuando construimos un muro para almacenar agua en un embalse o cuando perforamos un túnel, encontremos que el diseño en arco sustituye al diseño en linea.

[13]De nuevo aquí, «cualquiera» ha de tomarse en el sentido del lenguaje común de la palabra, no tanto en un sentido matemático, pues esta curva «cualquiera» ha de poder ser descrita por medio de una función que ha de cumplir una serie de condiciones (por ejemplo, ha de tener una cantidad finita de discontinuidades)

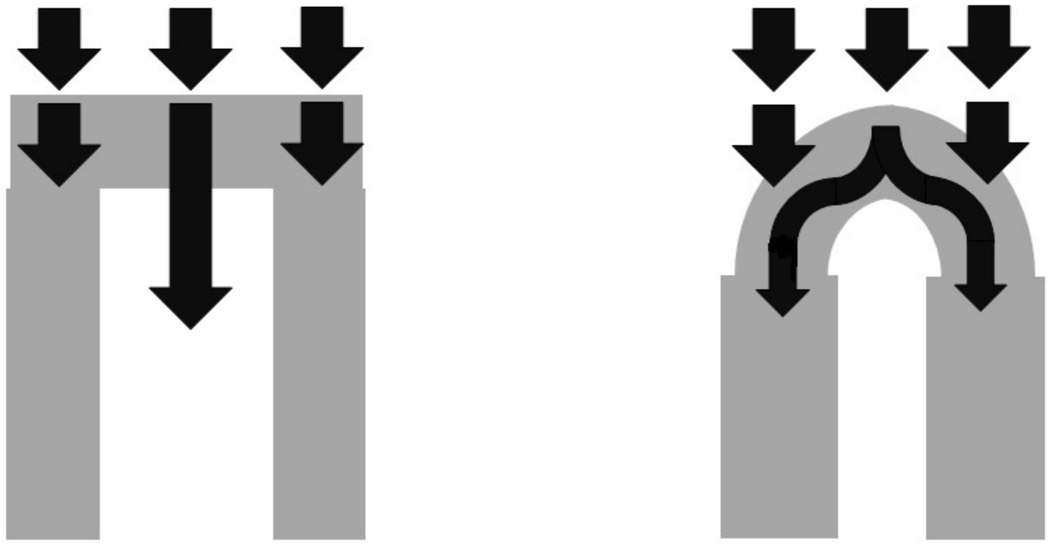

Figura 6: Distribución de cargas en un diseño tipo dintel (izquierda) y en un diseño en arco (derecha).

Esto no es un avance de nuestro tiempo. El uso del arco fue introducido en Europa gracias al imperio romano, que lo utilizó de forma sistemática en sus construcciones[14]. El diseño en arco ofrece una ventaja definitiva sobre el diseño tipo dintel. Mientras que éste último ha de soportar el peso de forma homogénea (Figura 6), lo que en muchas ocasiones conduce a que la estructura ceda en las zonas más alejadas de los puntos de apoyo; el arco trasmite toda, o gran parte de la carga, a las columnas sobre la cuál está apoyado, de forma que se convierte en un diseño mucho más robusto a la hora de soportar los pesos que gravitan el vano[15].

Será Robert Hook [16] quien justifique que la curva catenaria es la curva óptima para que un arco se sostenga. Técnicamente diríamos que una curva catenaria es el lugar geométrico de todos los centros de gravedad de un elemento lineal sometido únicamente a cargas verticales. Esto hace que de entre todos los posibles diseños en curva para la construcción de un arco, aquel que siga una curva catenaria será el que optimice el aguante de cargas verticales. La curva catenaria es un concepto sencillo de entender, pues en palabras comunes podríamos decir que es la curva que describe una cadena -sostenida por sus extremos- por el sólo efecto de una gravedad homogénea[17], pero no tan fácil de modelar. La expresión matemática que la describe es, $y = a\cosh(x/a)$, donde a es un factor de escala que ha de ser calculado en función del peso del cable y de la tensión horizontal que soporta. En la imagen de la izquiera de la Figura 7 se han representado distintas curvas catenarias, correspondientes a distintos valores del parámetro a (un menor valor de a da lugar a una curva más pronunciada).

[14]Veasé la forma aparente de cualquier construcción que lleve el adjetivo de «romano»: el Coliseo de Roma, el puente de Córdoba, el acueducto de Segovia, cualquier «arco del triunfo», y tantos otros

[15]En un principio, el diseño en arco se realiza de forma intuitiva, a partir de observar su presencia en la naturaleza. Podemos imaginar cómo es la forma en que el viento o el agua erosionan la roca. Ejemplos muy ilustrativos de esta erosión en forma de arco se pueden encontrar en «La ciudad encantada» de Cuenca, o visualizando imagenes de los desiertos americanos (Por ejemplo, en el Parque Natural de Utah, se encuentra un puente natural donde se puede observar esta forma de arco que ahora reproduce la arquitectura humana). Como brevísimo apunte, dirémos que el primero que se acerca a esta forma en arco, desde un punto de vista más riguroso o formal, es Leonardo da Vinci (sí, el que comienza su *Cuaderno de Nota*s con la frase «abstengase de leer este libro todo aquel que no sea matemático»).

[16]Sí, el de la Ley de Hooke/Ley de elasticidad.

[17]Podemos imaginar la linea curva que describe el cableado telefónico entre dos postes consecutivos.

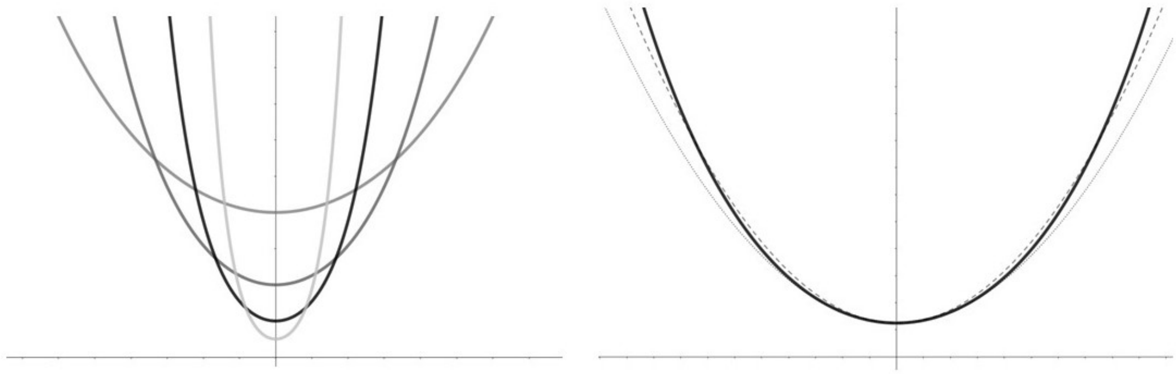

Figura 7: Izquierda: Curva catenaria para distintos valores del parámetro a. Derecha: Aproximación por medio de una curva parabólica (grises-punteados) en un entorno del vértice de una curva catenaria (negro sólido).

La parábola es una curva que aproxima bastante bien a una curva catenaria en un entorno de sus vertices[18], pero además, tiene la ventaja de que puede describirse de forma algebraica por medio de una expresión realmente sencilla, $y = a_0 + a_1x + a_2x^2$, donde a_0, a_1 y a_2, son parámetros que determinarán la forma (concava-convexa, ó más abierta - más profunda) y la posición de la parábola. Además, su expresión explícita (valores concretos de los parámetros) puede ser determinada simplemente conociendo tres puntos por los que pase.

Es casi obvia la necesidad de conocer la longitud de nuestro segmento curvo: desde la utilidad de conocer la distancia efectiva entre dos puntos, hasta la necesidad de presupuestar el material para la construcción - reconstrucción - reacondicionamiento de cierta estructura. Para ilustrar esta situación, y el cómo aplicariamos el concepto de integral definida para este fin, en la siguiente sección supondremos que tenemos que asfaltar una carretera que une dos puntos sobre una presa construida en curva.

3.2. Ejemplo: asfaltado de la carretera sobre una presa.

Supongamos que queremos construir una presa entre dos montañas. Para soportar las cargas producidas por el peso del agua, el diseño más adecuado es utilizar lineas curvadas en lugar de lineas rectas, pues, como hemos argumentado, estas lineas curvas trasladan el peso a los extremos del muro y haran que éste recaiga sobre la superficie natural (la montaña), en lugar de directamente sobre las paredes. Supongamos que sobre esta presa irá una carretera que unirá ambos margenes del río, y que queremos determinar cuantos metros lineales de asfalto necesitaremos para pavimentarla. Suponiendo un ancho de carretera constante, lo que necesitaríamos conocer es la longitud de esta presa.

En la Figura 8 se muestra una presa que puede servir para ilustrar nuestro ejemplo. Sobre ella se han trazado (en perspectiva) unos ejes cartesianos y se ha dibujado una parábola que aproxime la trayectoria curva que describe. Supongamos que la distancia en linea recta de los dos extremos de la presa es 2α. Por simplicidad en los cálculos, y sin perdida de generalidad, podemos utilizar esta línea recta imaginaria para situar el eje de abcisas de nuestro sistema de referencia, y suponer que el origen se encuentra en el punto medio (de esta ma-

[18]En la imagen de la derecha de la Figura 7 se ha representado una curva catenaria cualquiera (en color negro grueso), y se ha aproximado, en un entorno de su vertice, por medio de una curva parabólica (en gris punteado).

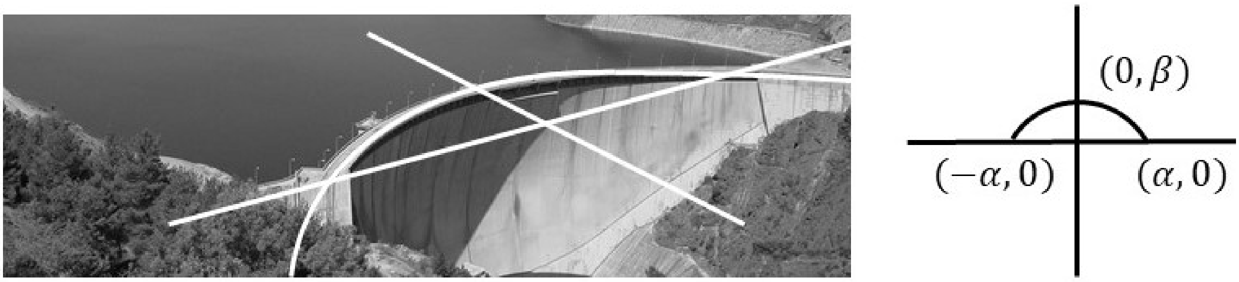

Figura 8: Carretera curva aproximada por una parábola referenciada en un eje cartesiano que la situa sobre los puntos $(\alpha,0)$, $(-\alpha,0)$ y $(0,\beta)$.

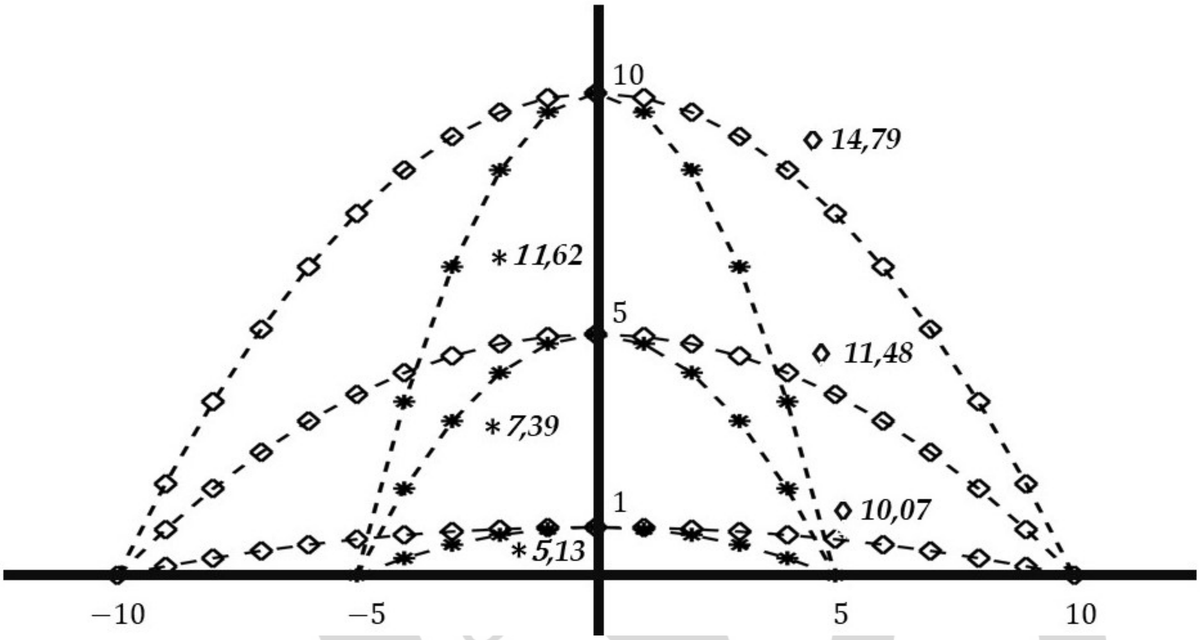

Figura 9: Longitudes de la parábola que une dos puntos separados a 10 m. (diamantes) y 5 m. (asteriscos) para distintos grados de convexidad.

nera el vertice de la parábola caerá sobre el eje de ordenadas). De esta manera la parábola pasará por los puntos $(\alpha,0)$, $(-\alpha,0)$ y $(0,\beta)$, a partir de los cuales es sencillo determinar su ecuación. Resolviendo el sistema resultante de sustituir estos puntos en la expresión genérica $y = a_0 + a_1x + a_2x^2$, obtenemos que $a_0 = \beta$, $a_1 = 0$, y $a_2 = -\beta/\alpha^2$. Por simplicidad en la escritura, en lo sucesivo trabajaremos con la expresión genérica simplicada $y = a_0 + a_2x^2$.

Para determinar la longitud de la carretera, tenemos que calcular la integral obtenida en (11) en el intervalo $[-\alpha,\alpha]$, para $f(x) = a_0 + a_2x^2$, lo que podríamos escribir

$$L = \int_{-\alpha}^{\alpha} \sqrt{1 + (2a_2x)^2}dx.$$

Esta función no tiene una primitiva inmediata, y las técnicas de integración necesarias para resolverla exceden lo exigible en un curso de Bachillerato[19]. Sin embargo, con un poco de

[19]Un primer cambio de variable $y = 2a_2x$ simplifica la expresión en la integral transformandola a $(1/2a)\int\sqrt{1+y^2}dy$. Esta integral se resolverá de forma, algo tediosa pero bastante mecánica, por sustitución trigonométrica $y = tan(u)$.

paciencia, el estudiante sí será capaz de demostrar que la derivada $F'(x) = f(x)$, siendo

$$F(x) = \frac{1}{2}x\sqrt{4a_2^2x^2+1} + \frac{1}{4a_2}\ln\left|2a_2x + \sqrt{1+4a_2^2x^2}\right|.$$

Aplicando el resultado del Teorema 1.19, tendremos que $L = F(\alpha) - F(-\alpha)$. De esta manera podremos computar la longitud de la carretera en función de los valores de los parámetros α (mitad de la distancia entre los margenes del río) y β (distancia entre el vértice de la parábola y la hipotética carretera en linea recta). En la Figura 9 se han representado las longitudes de la carretera para distintos valores de estos parámetros[20]. Para cada curva, el valor de α utilizado puede ser deducido a partir de los puntos de corte con el eje de abcisas, y el valor de β, desde el punto de corte con el eje de ordenadas. Para cada curva se adjunta el valor numérico de la longitud del arco que une los dos extremos representados.

3.3. Motivación 2: Curvas complejas.

Pensemos en una partícula orbitando alrededor de un punto dado, y en un primer escenario, supongamos, a una distancia constante. Cuando la partícula ha dado una vuelta completa alrededor de su punto de referencia, habrá descrito una circunferencia, y, haciendo uso de la geometría plana clásica se podrá calcular la longitud del camino recorrido. Supongamos ahora, en un segundo escenario, que la partícula no orbita a una distancia constante, sino que está siendo repelida, de forma que en cada vuelta se encuentra más lejos de su centro de referencia que en la vuelta anterior. La trayectoria que describe, en este caso, la partícula, es la de una espiral. Hay quizá dos espirales que, debido a su aplicabilidad y su presencia en la naturaleza, podríamos decir que tienen una fama especial. Éstas son, la espiral de arquímedes[21], donde la variación del radio en cada vuelta sigue una progresión aritmética, y la espiral logarítmica[22], donde la variación del radio se encuentra en progresión geométrica.

La espiral de arquímedes puede describirse de forma muy sencilla, en coordenadas polares, a partir de la expresión $r = a + b\theta$, donde r representa el radio (distancia de la partícula al punto de referencia) cuando ésta se ha movido un ángulo θ. Los valores a y b determinan la posición inicial y el paso de la particula, respectivamente. Utilizando la misma notación, la espiral logarítmica puede ser expresada, en coordenadas polares, de la forma $r = ab^\theta$.

Sabemos que la longitud de la circunferencia puede ser determinada, haciendo uso de los conocimientos de geometría plana, a partir de su radio r, como $L = 2\pi r$. En este sentido, no hay razón aparente para utilizar la integral con este proposito. Sin embargo, sí lo necesitaremos para determinar la longitud de una espiral. Puesto que la espiral logarítmica es una espiral mucho más atractiva que la arquimediana, nos centraremos en ella.

[20] Quizá sea de especial interes remarcar, para el caso de una separación de 10 $m.$, cómo una desviación de 1 $m.$ en la vertical únicamente nos requiere la construcción de 7 $cm.$ más de carretera. ¿No nos resulta extraño esto? ¿Y si lo reformulamos? Imaginemos que no tenemos una presa, que simplemente queremos construir un puente para poder salvar este río. Lo construiremos, por tanto, en linea recta y en llano. Una desviación de 7 $cm.$ en nuestra medida, es decir, una viga que mida 10,07 $m.$ en lugar de exactamente 10 $m.$ producirá un abombamiento en la carretera de 1$m.$, es decir, en sólo 5 metros de carretera se produce un desnivel de 1 metro, un 20 % de desnivel. Ahora, si 7 centimetros en 10 metros te parece un error despreciable, coge tu bicicleta y ve a subir esa cuesta.

[21] La espiral de arquímedes es la que aparece al enrollar una cuerda sobre sí misma. Tiene un basto abanico de aplicaciones ([13]): se utiliza en la creación de bombas de compresión ([7]), en la grabación de los discos de vinilo, en medicina como test neurológico.

[22] Esta espiral ha alcanzado su fama y atraido a personalidades matemáticas tan destacadas como Descartes o Bernouilli, debido a algunas propiedades que la hacen interesante y a su extraordinaria presencia en fenómenos y estructuras naturales ([6]).

3.4. Ejemplo: La espiral logarítmica.

Como ya hemos dicho, una espiral logarítmica queda representada, en coordenadas polares, por la expresión $r = ab^\theta$, donde a define el punto de partida, b define el tamaño del paso de una vuelta a otra, y θ denota el ángulo con respecto al eje de referencia. La espiral logarítmica se puede genrar para $\theta \in (-\infty, +\infty)$. Para simplificar los cálculos es conveniente representar la curva en coordenadas paraméticas. Esto es muy sencillo desde el momento en que sabemos que una circunferencia de radio r se expresa en estas coordenadas de la forma

$$(x(\theta), y(\theta)) = (r\sin(\theta), r\cos(\theta)).$$

Para representar la espiral, sólo será necesario hacer que el valor del radio, que es constante en el caso de la circunferencia, varíe dependiendo del valor del ángulo de la forma $r = ab^\theta$. De esta forma, la espiral logarítmica expresada en coordenadas paramétricas quedaría

$$(x(\theta), y(\theta)) = (ab^\theta sin(\theta), ab^\theta cos(\theta)).$$

Una pequeña modificación en esta expresión, va a hacer mucho más intuitiva la interpretación del movimiento de esta espiral. Modificamos el argumento θ a través del cambio de variable $\theta = 2\pi t$, estaremos obligando a que la espiral de una vuelta completa en cada instante de tiempo, por lo que el contador $t \in (-\infty, +\infty)$ puede interpretarse como el número de fueltas que ha dado sobre sí misma la espiral. De esta manera tenemos,

$$(x(t), y(t)) = (ab^{2\pi t} sin(2\pi t), ab^{2\pi t} cos(2\pi t)).$$

El cálculo de la longitud de una curva expresada en coordenadas paramétricas, entre los instantes de tiempo t_0 y t_1[23], se realiza a través de la expresión

$$L(t_0, t_1) = \int_{t_0}^{t_1} \sqrt{[x'(t)]^2 + [y'(t)]^2} dt. \tag{12}$$

que es análoga a la dada en (11).

Tras unos calculos relativamente sencillos, aunque quizá algo tediosos, obtenemos que,

$$(x'(t), y'(t)) = 2\pi^a b^{2\pi t}(Ln(b)\cos(2\pi t) - \sin(2\pi t), Ln(b)\sin(2\pi t) + \cos(2\pi t)).$$

El integrando en (12) será por lo tanto la raíz cuadrada de,

$$[x'(t)]^2 + [y'(t)]^2 = (2\pi ab^{2\pi t})^2[Ln^2(b) + 1].$$

Tenemos por lo tanto que

$$L(t_0, t_1) = \int_{t_0}^{t_1} 2\pi ab^{2\pi t}\sqrt{Ln^2(b) + 1} dt.$$

Resolviendo convenientemente tendremos que,

$$L(t_0, t_1) = \frac{a\sqrt{Ln^2(b) + 1}}{Ln(b)}[b^{2\pi t_1} - b^{2\pi t_0}]. \tag{13}$$

Algunas longitudes notables.

Supongamos que tenemos la espiral logarítmica $r = ab^{2\pi t}$. Según hemos dicho en el apartado anterior, y entendiento que para $t = 0$ tenemos una posición inicial, el valor de t, cuyo soporte en un principio es $(-\infty, +\infty)$, puede ser interpretado como la cantidad de vueltas

[23]También se puede interpretar como «entre la vuelta t_0 y la vuelta t_1».

que da la espiral sobre si misma. Esto es, si representamos la espiral para $t|in[0,1]$ tendremos la primera vuelta de la integral, si la representamos para $t \in [1,2]$ tendrémos la segunda vuelta, y si la prepresentamos para $t \in [0,2]$ tendremos las dos primeras vueltas. En general, $t \in [t_0, t_1]$ nos dará lugar a una espiral que da $t_1 - t_0$ vueltas sobre sí misma, desde la vuelta t_0, hasta la vuelta t_1. Estos números pueden ser en general positivos o negativos, interpretando que los positivos son vueltas que hacen que la espiral crezca, se aleje del origen, y los negativos como vueltas que hacen que la espiral mengue o se hacerque hacia el origen.

Ya hemos encontrado la expresión (13) que nos permite determinar la longitud de una espiral de este tipo en el intervalo $[t_0, t_1]$. Cuando t_1 crece indefinidamente ($t_0 \to \infty$), la espiral va alejandose del origen, expandiendose, por lo que resulta intuitivo que la longitud de la espiral también crece indefinidamente y que, por lo tanto, ha de terminar teniendo longitud infinita. Esto se puede confirmar determinando $L(t_0, \infty)$ en 13)[24],

$$L(t_0, \infty) = \frac{a\sqrt{Ln^2(b)+1}}{Ln(b)}[b^{2\pi\infty} - b^{2\pi t_0}] = \infty.$$

Sin embargo, en el sentido contrario, cuando t_0 se hace pequeño indefinidamente ($t_0 \to -\infty$), la espiral va acercandose cada vez más al origen y para ello va comprimiendose. Por un lado, podemos pensar que la espiral se está representando en el intervalo $[-\infty, t_1]$, por lo que gira infinitas veces sobre si misma. Pero, ¿tiene también una longitud infinita?. En este caso debemos calcular el valor de $L(-\infty, t_1)$,

$$L(-\infty, t_1) = \frac{a\sqrt{Ln^2(b)+1}}{Ln(b)}[b^{2\pi t_1} - b^{-2\pi\infty}] = \frac{a\sqrt{Ln^2(b)+1}}{Ln(b)}b^{2\pi t_1}$$

que es un número finito. Por lo tanto, aunque es verdad que la espiral da infintas vueltas, la longitud del camino que recorre es finita[25], algo que en un principio puede resultar completamente antiintuitivo.

Para terminar, haremos notar un caso particular a la par que interesante. Sea $r = e^{2\pi t}$ la espiral logarítmica de paso e con punto inicial en el eje cartesiano $(0,1)$. Como hemos visto, resulta contraintuitivo, que aunque la espiral gira infinitas veces sobres si misma al ir acercandose a su centro, la longitud de la curva desde cualquier punto ha dicho centro es finita. ¿Cuál es la longitud de esta «espiral neperiana» desde el punto de partida hasta el origen?. Se trata de calcular $L(-\infty, 0)$ en (13), quedando,

$$L(-\infty, 0) = \frac{\sqrt{Ln^2(e)+1}}{Ln(e)} = \sqrt{2}$$

es decir, la espiral está girando infinitas veces sobre si misma para terminar recorriendo un camino de longitud igual a la diagonal de un cuadrado de lado 1.

Referencias

[1] Apostol, T. M., *Análisis matemático.*, Editorial Reverté, S.A., (2010)[26]

[24] Notese que $L(t_0, +\infty)$ da lugar a una integral impropia que requiere el paso al límite $lim_{t_1 \to +\infty} L(t_0, t_1)$ para una resolución rigurosa del caso. La forma en que expresamos la solución aquí puede ser tomada como la parte final de la resolución de dicho límite, o simplemente, como una «relajación» en la notación, interpretando ∞ como un número concreto (y atendiendo después a sus particularidades).

[25] ¿Qué tiene que decir Zenon ahora?

[26] La mayoría de los textos utilizados como referencia en este trabajo son textos técnicos, manuales reconocidos, y algunos apuntes cualquiera de alguna universidad cualquiera de un curso de análisis matemático de primeros niveles universitarios. Por ello, sólo se comentarán las referencias que no se puedan encuadrar en alguno de esos bloques.

[2] Apostol, T. M., *Calculus. vol. I.*, George Springer, S.A., (1967)

[3] Bell, E. T., *Los grandes matemáticos. Su vida y sus obras.*, Editorial Losada, S.A., (2010)
Un paseo por los grandes descubrimientos de las matemáticas desde la vida de los grandes matemáticos. En el presente trabajo han tenido gran influencia los capítulos segundo (Zenón, Eudoxio, Arguímedes), Cuarto (Fermat), Sexto (Newton) y Séptimo (Leibniz), pero el libro es muy recomendable, no sólo en su contenido, si no que también en su forma y en la interesante distinción que hace entre las matemáticas y los matemáticos intuicionistas y las matemáticas y los matemáticos formalistas.

[4] Bichteler, K., *Stochastic Integration With Jumps.*, Cambridge University Press, (2010)
Manual técnico cuyo contenido especializado no es objetivo de este trabajo, pero que ilustra sobradamente la existencia de otros horizontes en el campo de la investigación en el concepto de integral y su aplicación.

[5] Boyer, C.B., *A History of Mathematics.*, Wiley International Edition, (1968)
¿El historiador más famoso en matemáticas? De especial intereses para este trabajo son los capítulos XVII, XVIII y XIX, aunque todo ésto se entenderá mejor si se ha recorrido antes el capítulo dedicado a Arquímedes.

[6] Cook, T. A., *The Curves of Life.*, Dover Pub. Inc., (1979)
Una recopilación masiva de ejemplos de cómo la curva que tanto inspiro a Jakob Bernoulli (eadem mutata resurgo) está presente en la naturaleza y es la forma en que ésta ha decidido desarrollarse y crecer. El discurso de este libro debería acabar con el misticismo relativo a número de oro y toda su hegemonía en la imaginaria popular.

[7] Mitsunaga, T. y Sato, K., *Compresor espiral.*, Oficina española de Patentes y Marcas, (1996)
En esta patente se muestra cómo la espiral arquimediana es una curva aplicada al diseño práctico de compresores digitales. En el interior del documento se puede ver un diseño específico muy interesante que ilustra cómo funciona la espiral en este objetivo.

[8] Ortega, J. M., *Introducción al análsis matemático.*, UAB, (1993)

[9] Ortega, J. M., *Análisis IV.*, (2001) disponible en http://www.maia.ub.es/cag/docencia/anali-IV.pdf,

[10] Perez-Gonzalez, F. J., *Cálculo Diferencial e Integral de funciones de una variable.*, Dto. de Análisis Matemático. Universidad de Granada, (2008)

[11] Spivak, M., *Calculus. Cálculo infinitesimal.*, Editorial Reverté, S.A., (1998)

[12] Martín-Suárez, M., *Orígenes del Cálculo Diferencial e Integral.*, Universidad de Granada, (2008)
Exposición histórica de los orígenes del cálculo, algo esquemático, pero bien ilustrativo. Se puede utilizar como línea del tiempo y profundizar en cada apartado utilizando otras referencias.

[13] Yates, J.R., *A Handbook on Curves and their Properties.*, Edwards Brothers, Inc. (1949)