



FACULTAD DE MATEMÁTICAS

DEPARTAMENTO DE ANÁLISIS MATEMÁTICO

Trabajo Fin de Grado

La curva de Hilbert y otras curvas que rellenan áreas o volúmenes

Miriam Gandul Jiménez

Dirigido por:

Juan Carlos García Vázquez

Junio 2019

Resumen

El presente trabajo de fin de grado, en su primera parte, estudia con detalle la construcción, propiedades geométricas y analíticas de la curva de Hilbert y de otras curvas similares, que son funciones continuas definidas en el intervalo unidad $[0, 1]$ y con valores en \mathbb{R}^2 o \mathbb{R}^3 tales que la imagen del intervalo unidad tiene contenido de Jordan positivo.

En la segunda parte del trabajo se explica la construcción de la curva de Lebesgue, la cual se basa en el teorema de Hausdorff que demuestra que todo compacto es imagen continua del conjunto de Cantor. Todo esto sirve de idea para estudiar que tipos de conjuntos de \mathbb{R}^n pueden ser la imagen de un intervalo mediante una aplicación continua, es decir, el teorema de Hahn- Mazurkiewick. También se demuestra previamente que una función continua definida del intervalo unidad al cuadrado unidad no puede ser biyectiva.

Por último se estudia la existencia de curvas definidas del intervalo unidad a \mathbb{R}^n continuas e inyectivas con medida de Lebesgue positiva, como la curva de Osgood.

Abstract

The present final graduate project, in the first part, studies with details the construction, geometric and analytical properties of the Hilbert's space-filling curve and others similar curves, which are continuous functions defined in the unit interval $[0, 1]$ and with values in \mathbb{R}^2 or \mathbb{R}^3 such that the unit interval image has positive Jordan measure.

In the second part of the work, we explain the construction of the Lebesgue curve, which is based on the Hausdorff theorem which shows that every compact is a continuous image of the Cantor set. All this serves as an idea to study what types of sets of \mathbb{R}^n can be the image of a range by a continuous application, that is, Hahn-Mazurkiewick's theorem. It is also shown previously that a continuous function defined from the unit interval to the unit square can not be bijective.

Finally, we study the existence of continuous and injective curves defined of the unit interval to \mathbb{R}^n with a positive Lebesgue measure, such as the Osgood curve.

Índice general

Resumen	III
Abstract	V
Introducción	1
1. La curva de Hilbert	5
1.1. Construcción de la curva de Hilbert	5
1.2. Representación compleja de la curva de Hilbert	8
1.3. Forma aritmética de la curva de Hilbert	16
1.4. La curva de Hilbert no es diferenciable. Prueba analítica.	19
1.5. Aproximación por polígonos de la curva de Hilbert	22
1.6. La curva de Hilbert tridimensional	24
1.7. La curva de Peano	29
1.8. La curva de Sierpinski	30
2. El teorema de Hahn-Mazurkiewicz	33
2.1. El conjunto de Cantor	33
2.2. Propiedades del conjunto de Cantor	36
2.3. Definición de la curva de Lebesgue	42
2.4. Aproximación por polígonos de la curva de Lebesgue	44
2.5. Preliminares de Topología	48
2.6. Teorema de Netto	51
2.7. Conjuntos localmente conexos	52
2.8. Teorema de Hausdorff	54

2.9. Conexo por caminos	57
2.10. El teorema de Hahn-Mazurkiewicz	61
3. Curva de Jordan con medida de Lebesgue positiva	65
3.1. Curva de Jordan de Osgood	65
3.2. Las curvas de Jordan de Sierpinski y Knopp	70
3.3. Otras curvas de Osgood	76
Complemento	79
Procesamiento de imágenes	79
Otras aplicaciones	82
Bibliografía	87

Introducción

El objetivo del TFG es, en primer lugar, estudiar con detalle tanto la construcción como las propiedades geométricas y analíticas de la curva de Hilbert y de otras curvas similares (funciones continuas definidas en el intervalo $[0, 1]$ y con valores en \mathbb{R}^2 o \mathbb{R}^3 tales que la imagen del intervalo unidad tiene contenido de Jordan > 0). Esto lo hacemos en las diferentes secciones del Capítulo 1 de la memoria.

Ponemos a continuación en su contexto histórico este tipo de curvas. La historia comienza con un artículo de Cantor publicado en el año 1878 en el que demuestra que los conjuntos $[0, 1]$ y $[0, 1]^2$ son biyectivos (de hecho prueba algo más general, prueba que cualesquiera dos variedades diferenciables de dimensión finita son biyectivas). La idea original de la prueba es relativamente sencilla: dado un punto $(x, y) \in [0, 1]^2$, se escriben los desarrollos decimales de cada una de las coordenadas $x = 0.x_1x_2x_3\dots$, $y = 0.y_1y_2y_3\dots$ y se asocia con el punto $0.x_1y_1x_2y_2x_3y_3\dots$ que obviamente pertenece al intervalo $[0, 1]$. Dirichlet, con quien Cantor mantuvo correspondencia, se dio cuenta que esta idea tiene algunos problemas, pero tales problemas se pueden arreglar. Cantor lo arregló usando fracciones continuas en vez del desarrollo decimal para representar las coordenadas, pero también se puede ir a la idea original de Cantor y usar el teorema de Schroder-Bernstein que dice lo siguiente: si X e Y son dos conjuntos cualesquiera, existe una inyección de X en Y y otra de Y en X , entonces X e Y son biyectivos.

No obstante, la biyección de Cantor no es continua y abre puerta a la pregunta natural: ¿puede darse una biyección continua entre $[0, 1]$ y $[0, 1]^2$? Esta pregunta la responde Netto en 1879: no es posible, cualquier biyección debe ser discontinua (Netto usa un argumento topológico de conexión que estudiamos en la sección 2.6).

El siguiente paso, después de abandonar la idea de la existencia de una biyección

continua, fue encontrar funciones continuas sobreyectivas de $[0, 1]$ en $[0, 1]^2$. En 1890 aparece una primera construcción debida a Peano que nosotros describiremos en la sección 1.7. Es la primera curva que fue bautizada con el término en inglés space-filling curve y que nosotros traduciremos por curvas que rellenan espacio.

Después de la curva de Peano aparecieron otras curvas, entre las cuales destacamos:

1. Hilbert (1891)
2. Moore (1900)
3. Lebesgue (1904)
4. Sierpinski (1912)

En general a todas estas curvas se les llama curvas de Peano en honor a la primera que apareció. Veamos la siguiente definición de curva que rellena espacio:

Definición 0.1. *Una curva de Peano o curva que rellena espacio, es una aplicación continua*

$$f : [0, 1] \rightarrow \mathbb{R}^n$$

de modo que $J_n(f([0, 1])) > 0$, donde J_n es el contenido de Jordan en \mathbb{R}^n .

Recordamos que el contenido de Jordan de un conjunto acotado n -dimensional A , viene dado por la fórmula

$$J_n(A) = \sup\{m(P) : P \subset A\}$$

donde P recorre la colección de conjuntos elementales contenidos en A . Un conjunto P es elemental si es una unión finita de cajas en \mathbb{R}^n y $m(P)$ representa la medida natural de tales conjuntos.

Casi todas estas curvas se construyen realizando un proceso geométrico recursivo (como veremos con detalle en el Capítulo 1) y en casi todos los casos se generan curvas no derivables en ningún punto del intervalo $[0, 1]$. Nosotros hemos centrado el TFG en la curva de Hilbert por los siguientes motivos: su construcción es elegante, versátil, generalizable y es muy utilizada en las aplicaciones.

Seguimos con el hilo histórico. Después de la aparición de tales ejemplos, algunos matemáticos siguieron trabajando en el caso inyectivo. En esta dirección, Osgood encontró un ejemplo de función $f : [0, 1] \rightarrow \mathbb{R}^2$ continua e inyectiva tal que $m(f([0, 1])) > 0$ siendo m la medida de Lebesgue. De hecho, construye una familia de tales curvas (f_α) , una para cada $0 < \alpha < 1$, de modo que el rango de cada una de ellas tiene medida de Lebesgue exactamente α . De hecho la colección converge uniformemente a la curva original de Peano cuando $\alpha \rightarrow 1$. Esto lo estudiamos en el Capítulo 3. También estudiaremos la curva de Knopp, la cual aparece en 1917 y tiene las mismas propiedades que la curva de Osgood, pero añade una cualidad más: en este caso, la imagen de cualquier subintervalo de $[0, 1]$ tiene medida de Lebesgue estrictamente positiva. Si tomamos límite en este caso, aparece la curva de Sierpinski. Estos resultados los veremos en el Capítulo 3.

Después de la colección de ejemplos que ponen de manifiesto lo poco intuitiva que es la definición de curva como la imagen de una función continua definida sobre un intervalo, posiblemente la pregunta más importante que cabe plantearse es la siguiente: ¿qué tipo de conjuntos de \mathbb{R}^n pueden ser la imagen de un intervalo mediante una aplicación continua? La respuesta a esta pregunta es el teorema de Hahn-Mazurkiewick de 1917: un conjunto A de \mathbb{R}^n es la imagen de un intervalo mediante una aplicación continua si y solo si A es compacto, conexo y localmente conexo. Este teorema y su conexión con la construcción de la curva de Lebesgue la estudiamos en el Capítulo 2 del TFG.

Por último damos una idea de algunas aplicaciones de estas curvas a algunos problemas concretos, aunque hemos elegido contar con más detalle la aplicación de la curva de Hilbert al procesamiento de imágenes. En el libro de Michel Bader *Space-filling curves, an introduction with applications in scientific computing*, como en otros artículos de la bibliografía, podemos estudiar muchas más.

Capítulo 1

La curva de Hilbert

1.1. Construcción de la curva de Hilbert

El matemático David Hilbert describe la curva de Hilbert en 1891, como una curva fractal continua que recubre el plano, en la que cada una de las curvas que aproximan a la curva final es simple, es decir, no se corta a sí misma.

Su construcción y definición como una curva que rellena espacio se basa en un procedimiento recursivo:

1. Primero dividimos el intervalo unidad \mathcal{I} en cuatro subintervalos iguales,

$$\mathcal{I}_1 = [0, 1/4], \mathcal{I}_2 = [1/4, 1/2], \mathcal{I}_3 = [1/2, 3/4], \mathcal{I}_4 = [3/4, 1]$$

y el cuadrado unidad \mathcal{Q} en cuatro subcuadrados $\mathcal{Q}_1, \mathcal{Q}_2, \mathcal{Q}_3, \mathcal{Q}_4$, cada uno con longitud lateral $1/2$ y de centro $(1/4, 1/4), (1/4, 3/4), (3/4, 3/4)$ y $(3/4, 1/4)$. Luego asignamos cada subintervalo a uno de los subcuadrados de la siguiente forma: el intervalo \mathcal{I}_1 lo asignamos al subcuadrado inferior izquierdo \mathcal{Q}_1 , el intervalo \mathcal{I}_2 al subcuadrado superior izquierdo \mathcal{Q}_2 , el \mathcal{I}_3 al superior derecho \mathcal{Q}_3 y el intervalo \mathcal{I}_4 al subcuadrado inferior derecho \mathcal{Q}_4 . (Ver Fig. 1.1(a))

2. Posteriormente, dividimos cada subintervalo de \mathcal{I} del paso anterior en cuatro subintervalos de igual longitud, $1/16$, que denominamos I_{kj} con $k, j = 1, 2, 3, 4$

$$I_{kj} = \left[\frac{k-1}{4} + \frac{j-1}{4^2}, \frac{k-1}{4} + \frac{j}{4^2} \right].$$

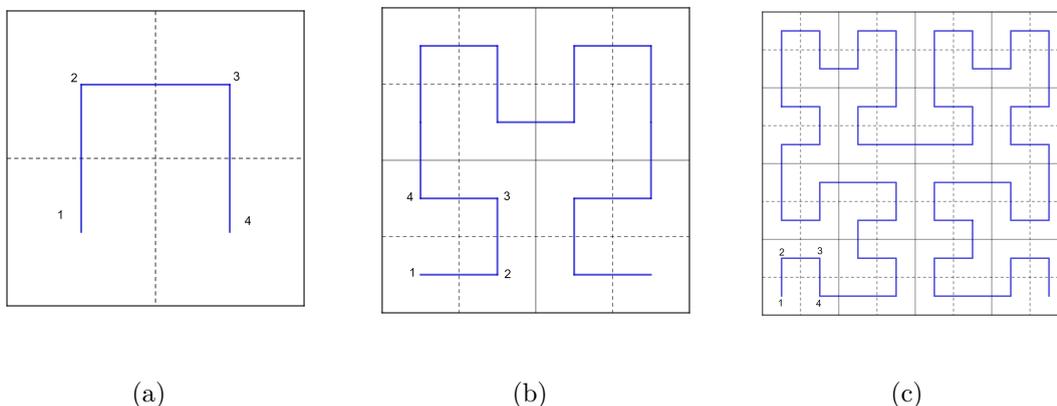


Figura 1.1. Construcción de la curva de Hilbert.

y dividimos cada subcuadrado Q_k en cuatro subcuadrados Q_{kj} , cada uno con una longitud lateral de $1/4$. Luego establecemos una correspondencia entre ellos $I_{kj} \leftrightarrow Q_{kj}$ de modo que subintervalos consecutivos se transforman en subcuadrados adyacentes (comparten uno de los lados). Ver Fig. 1.1.

Hilbert demostró que el proceso puede repetirse tantas veces como se quiera, consiguiendo en cada paso de la construcción que:

- subintervalos adyacentes se transformen en subcuadrados adyacentes.
- si un intervalo se corresponde a un cuadrado, cada uno de sus subintervalos se corresponde con un subcuadrado contenido en el cuadrado inicial (se preserva la relación de inclusión).

Se va realizando este proceso hasta que el área de cada subdivisión tienda a cero y la curva tienda a ocupar toda la superficie del cuadrado. El proceso permite la siguiente definición geométrica

Definición 1.1. Cada $t \in \mathcal{I}$ está determinada de forma única por una secuencia de intervalos cerrados anidados (que se generan mediante nuestra partición sucesiva), cuyas longitudes tienden a 0. A esta secuencia le corresponde una única secuencia de cuadrados cerrados anidados, cuyas diagonales se contraen a un punto, que definen un único punto en \mathcal{Q} , $f_h(t)$. Denominamos a $f_h(\mathcal{I})$ la Curva de Hilbert.

| Teorema 1.1. *La aplicación $f_h : \mathcal{I} \rightarrow \mathcal{Q}$ está bien definida, es continua, sobreyectiva y no es diferenciable en ningún punto.*

Demostración. Si t es el punto final de uno de los subintervalos (distintos de 0 y 1), entonces t pertenece a dos secuencias diferentes de intervalos anidados, pero como intervalos adyacentes se asignan a cuadrados adyacentes, esto conduce a la misma imagen, es decir la aplicación f_h está bien definida.

Es sobreyectiva porque dado cualquier punto $p \in \mathcal{Q}$ podemos dar una sucesión de subcuadrados encajados cuya intersección es p y esta subsucesión de subcuadrados tiene asociada una sucesión de subintervalos encajados en \mathcal{I} que define un cierto punto $t \in \mathcal{I}$ que obviamente cumple $f_h(t) = p$.

La aplicación f_h es continua pues si $|t - t'| < 1/4^n$ entonces o bien pertenecen al mismo intervalo \mathcal{I} de la etapa n o bien pertenecen a dos subintervalos consecutivos y por ello $f_h(t)$ y $f_h(t')$ o bien pertenecen al mismo subcuadrado o bien pertenecen a subcuadrados adyacentes de donde $\|f_h(t) - f_h(t')\| \leq \sqrt{5}/2^n$. Esto prueba la continuidad uniforme de f_h en \mathcal{I} .

Falta ver que la función f_h no es derivable en ningún punto del intervalo \mathcal{I} , para ello ninguna de sus dos funciones componentes lo son. Pongamos $f_h = (\varphi, \psi)$. Fijemos $t \in \mathcal{I}$. Sea $n \geq 3$. Fijemos t_n de modo que $|t - t_n| \leq 10/2^{2n}$ y que $f_h(t)$ esté separado de $f_h(t_n)$ horizontalmente por al menos un subcuadrado, esto es,

$$|\varphi(t) - \varphi(t_n)| \geq \frac{1}{2^n}$$

Esto siempre es posible gracias a como se ordenan los subcuadrados de un orden. Entonces tenemos

$$\frac{|\varphi(t) - \varphi(t_n)|}{|t - t_n|} \geq \frac{1/2^n}{10/2^{2n}} = \frac{2^n}{10}$$

con lo que se demuestra que φ no puede ser derivable en t . Lo mismo puede hacerse para ψ . Luego f_h no es derivable en ningún punto. **|**

1.2. Representación compleja de la curva de Hilbert

Mirando la Fig. 1.1 vemos que la curva de Hilbert se origina en el subcuadrado inferior izquierdo y termina en el subcuadrado inferior derecho. Como los subcuadrados se reducen a puntos, se deduce que la curva de Hilbert comienza en el origen (punto de entrada en \mathcal{Q}) y termina en el punto $(1,0)$ (punto de salida de \mathcal{Q}). Los subcuadrados tienen que estar orientados de tal manera que el punto de salida de cada subcuadrado coincide con el punto de entrada del siguiente subcuadrado. En la Fig. 1.2 hemos indicado la orientación correcta de los subcuadrados de los dos primeros pasos del proceso de iteración. Para obtener la configuración en la Fig. 1.2(b), tenemos que someter a \mathcal{Q} en la Fig. 1.2 (a) a las siguientes transformaciones: sea $z \in \mathbb{C}$. Primero reducimos \mathcal{Q} en la relación 2:1 uniformemente hacia el origen ($z' = \frac{1}{2}z$). Luego, rotamos el cuadrado resultante 90° ($z'' = z'i$) y, finalmente, hacemos una reflexión sobre el eje imaginario ($z''' = \bar{z}''$). Combinando estas tres transformaciones y nombrando a la transformación compuesta T_0 , obtenemos

$$T_0z = \frac{1}{2}\bar{z}i$$

Esto produce el cuadrado en la esquina inferior izquierda de la Fig. 1.2(b). Para obtener el cuadrado en la esquina superior izquierda de la Fig. 1.2(b), reducimos \mathcal{Q} como antes y luego lo desplazamos $1/2$ unidades hacia arriba. Esto produce la transformación

$$T_1z = \frac{1}{2}z + \frac{i}{2}$$

El cuadrado de la esquina superior derecha se obtiene reduciendo \mathcal{Q} y desplazándolo $1/2$ unidades hacia arriba y $1/2$ unidades hacia la derecha:

$$T_2z = \frac{1}{2}z + \frac{1}{2} + \frac{i}{2}$$

Finalmente, obtenemos el cuadrado en la esquina inferior derecha de la Fig. 1.2 (b) reduciendo \mathcal{Q} , girándolo -90° ($z' = -\frac{1}{2}zi$), haciendo una reflexión sobre el eje imaginario ($z'' = -\bar{z}'$), y desplazando el resultado 1 unidad a la derecha y $1/2$

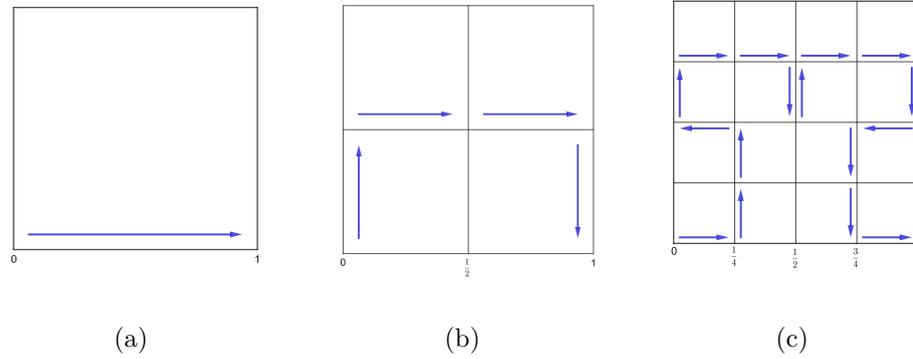


Figura 1.2. Aplicación de las transformaciones al cuadrado unidad.

unidades hacia arriba para obtener:

$$T_3 z = -\frac{1}{2}\bar{z}i + 1 + \frac{i}{2}$$

La aplicación de estas cuatro transformaciones a la Fig. 1.2 (a) produce la Fig. 1.2 (b). Si las aplicamos a la Fig. 1.2 (b) obtendríamos la Fig. 1.2 (c), etc.

Las cuatro aplicaciones complejas permiten escribir la asignación de subcuadrados a intervalos en la etapa n de la construcción de la curva de Hilbert. Pongamos que

$$t = 0_4 q_1 q_2 q_3 \dots = \frac{q_1}{4} + \frac{q_2}{4^2} + \frac{q_3}{4^3} + \dots, \quad \text{donde } q_j = 0, 1, 2, 3.$$

es un desarrollo en base 4 del número $0 \leq t \leq 1$. El primer dígito q_1 nos indica a que subintervalo de la etapa 1 pertenece: $[q_1/4, (q_1 + 1)/4]$ y ese subintervalo está relacionado con el subcuadrado \mathcal{Q}_{q_1+1} siendo:

- $\mathcal{Q}_1 = T_0(\mathcal{Q}) = [0, 1/2] \times [0, 1/2]$ (con centro el punto $(1/4, 1/4)$).
- $\mathcal{Q}_2 = T_1(\mathcal{Q}) = [0, 1/2] \times [1/2, 1]$ (con centro el punto $(1/4, 3/4)$).
- $\mathcal{Q}_3 = T_2(\mathcal{Q}) = [1/2, 1] \times [1/2, 1]$ (con centro el punto $(3/4, 3/4)$).
- $\mathcal{Q}_4 = T_3(\mathcal{Q}) = [1/2, 1] \times [0, 1/2]$ (con centro el punto $(3/4, 1/4)$).

El primer y segundo dígito $q_1 q_2$ nos indica el subintervalo de la etapa 2 al que pertenece t (dentro del intervalo de la etapa 1 al que pertenece, está ocupando la posición que indica q_2). Pertenece al subcubo de la etapa 2 dado por:

$$\mathcal{Q}_{q_1, q_2} = T_{q_1} T_{q_2}(\mathcal{Q})$$

Siguiendo este razonamiento, y teniendo en cuenta que los subcuadrados se reducen a puntos, llegamos a:

| Teorema 1.2. Si $t = 0_3q_1q_2q_3 \dots q_n$, entonces

$$f_h(t) = \lim_{n \rightarrow \infty} T_{q_1} T_{q_2} T_{q_3} \dots T_{q_n}(\mathcal{Q})$$

Este teorema nos sirve para calcular $f_h(t)$ cuando t es un cuaternario finito, esto es cuando existe n tal que $q_{n+k} = 0$ para todo $k \in \mathbb{N}$ y t se escribe

$$t = 0_4q_1 \dots q_n.$$

| Teorema 1.3. Si $t = 0_4q_1 \dots q_n$ entonces $f_h(t) = T_{q_1} \dots T_{q_n}(0)$.

Demostración. Observemos primero que

$$0_4q_1q_2q_3 \dots q_n = \frac{q_1}{4} + \frac{q_2}{4^2} + \frac{q_3}{4^3} + \dots + \frac{q_n}{4^n} = 0_4q_1q_2q_3 \dots q_n 0000 \dots$$

Entonces por el teorema 1.2 tenemos

$$f_h(0_4q_1q_2q_3 \dots q_n) = \lim_{n \rightarrow \infty} T_{q_1} T_{q_2} T_{q_3} \dots T_{q_n} T_0 T_0 T_0 \dots(\mathcal{Q})$$

Recordando que $T_0 = \frac{1}{2}\bar{z}i$ es fácil ver que

$$T_0 T_0 T_0 \dots z = \lim_{n \rightarrow \infty} T_0^n z = \lim_{n \rightarrow \infty} \frac{1}{2^n} \begin{cases} i\bar{z} & \text{si } n \text{ es impar} \\ z & \text{si } n \text{ es par} \end{cases}$$

entonces $T_0 T_0 T_0 \dots z = 0$ y por tanto,

$$f_h(0_4q_1q_2q_3 \dots q_n) = T_{q_1} T_{q_2} T_{q_3} \dots T_{q_n}(0). \quad (1.1)$$

|

Escribimos ahora las transformaciones T_0, T_1, T_2 y T_3 de otro modo:

$$T_q(z) = H_q(z) + h_q$$

donde:

$$H_0(z) = \bar{z}i, \quad H_1(z) = z, \quad H_2(z) = z, \quad H_3(z) = -\bar{z}i$$

$$h_0 = 0, \quad h_1 = i, \quad h_2 = 1 + i, \quad h_3 = 2 + i$$

Teorema 1.4. Si $t = 0_4 q_1 \dots q_n$ entonces

$$f_h(t) = \sum_{j=2}^n \frac{1}{2^j} H_{q_1} \dots H_{q_{j-1}}(h_{q_j}) + \frac{1}{2} h_{q_1}.$$

Demostración. Sustituyendo sucesivamente observemos que

$$\begin{aligned} T_{q_1} T_{q_2}(0) &= \frac{1}{2} H_{q_1} \{T_{q_2}(0)\} + \frac{1}{2} h_{q_1} \\ &= \frac{1}{2} H_{q_1} \left\{ \frac{1}{2} H_{q_2}(0) + \frac{1}{2} h_{q_2} \right\} + \frac{1}{2} h_{q_1} \\ &= \frac{1}{2^2} H_{q_1} H_{q_2}(0) + \frac{1}{2^2} H_{q_1}(h_{q_2}) + \frac{1}{2} h_{q_1} \end{aligned}$$

Para tres se obtiene que

$$\begin{aligned} T_{q_1} T_{q_2} T_{q_3}(0) &= \frac{1}{2} H_{q_1} \{T_{q_2} T_{q_3}(0)\} + \frac{1}{2} h_{q_1} \\ &= \frac{1}{2} H_{q_1} \left\{ \frac{1}{2} H_{q_2} [T_{q_3}(0)] + \frac{1}{2} h_{q_2} \right\} + \frac{1}{2} h_{q_1} \\ &= \frac{1}{2} H_{q_1} \left\{ \frac{1}{2} H_{q_2} \left[\frac{1}{2} H_{q_3}(0) + \frac{1}{2} h_{q_3} \right] + \frac{1}{2} h_{q_2} \right\} + \frac{1}{2} h_{q_1} \\ &= \frac{1}{2^3} H_{q_1} H_{q_2} H_{q_3}(0) + \frac{1}{2^3} H_{q_1} H_{q_2}(h_{q_3}) + \frac{1}{2^2} H_{q_1}(h_{q_2}) + \frac{1}{2} h_{q_1} \end{aligned}$$

Entonces, de forma general tendríamos

$$\begin{aligned} T_{q_1} T_{q_2} T_{q_3} \dots T_{q_n}(0) &= \frac{1}{2} H_{q_1} \{T_{q_2} T_{q_3} \dots T_{q_n}(0)\} + \frac{1}{2} h_{q_1} \\ &= \frac{1}{2} H_{q_1} \left\{ \frac{1}{2} H_{q_2} [T_{q_3} \dots T_{q_n}(0)] + \frac{1}{2} h_{q_2} \right\} + \frac{1}{2} h_{q_1} \\ &= \frac{1}{2} H_{q_1} \left\{ \frac{1}{2} H_{q_2} \left[\frac{1}{2} H_{q_3} (T_{q_4} \dots T_{q_n}(0)) + \frac{1}{2} h_{q_3} \right] + \frac{1}{2} h_{q_2} \right\} + \frac{1}{2} h_{q_1} \\ &= \dots \\ &= \left(\frac{1}{2^n} \right) H_{q_1} H_{q_2} H_{q_3} \dots H_{q_n}(0) + \left(\frac{1}{2^n} \right) H_{q_1} H_{q_2} H_{q_3} \dots H_{q_{n-1}}(h_{q_n}) \\ &\quad + \left(\frac{1}{2^{n-1}} \right) H_{q_1} H_{q_2} H_{q_3} \dots H_{q_{n-2}}(h_{q_{n-1}}) + \dots + \frac{1}{2} h_{q_1} \end{aligned}$$

Si suponemos que $H_{q_0}(z) = z$ y como $H_{q_1} H_{q_2} \dots H_{q_n}(0) = 0$, pues $H_q(0) = 0$ para todo q , nos queda

$$T_{q_1} T_{q_2} T_{q_3} \dots T_{q_n}(0) = \sum_{j=1}^n \left(\frac{1}{2^j} \right) H_{q_0} H_{q_1} H_{q_2} H_{q_3} \dots H_{q_{j-1}}(h_{q_j}).$$

Por tanto, sustituyendo en (1.1) tenemos

$$f_h(0_4q_1q_2q_3 \dots q_n) = \sum_{j=1}^n \left(\frac{1}{2^j}\right) H_{q_0} H_{q_1} H_{q_2} \dots H_{q_{j-1}}(h_{q_j})$$

Debido a la continuidad de f_h podemos deducir el siguiente teorema.

Teorema 1.5. *Si $t = 0_4q_1 \dots q_n \dots$ entonces*

$$f_h(t) = \sum_{j=2}^{\infty} \frac{1}{2^j} H_{q_1} \dots H_{q_{j-1}}(h_{q_j}) + \frac{1}{2} h_{q_1}.$$

Ahora teniendo en cuenta las relaciones $H_1(z) = H_2(z) = z$ y $H_0^2(z) = H_3^2(z) = z$ y la relación de conmutación $H_0H_3(z) = H_3H_0(z)$ podemos reescribir la fórmula del teorema anterior así

$$f_h(t) = \sum_{j=2}^{\infty} \frac{1}{2^j} H_0^{e(0,j)} H_3^{e(3,j)}(h_{q_j}) + \frac{1}{2} h_{q_1}$$

donde $e(k, j)$ es el número de veces en módulo 2 que aparece el dígito k antes de la posición j en el desarrollo cuaternario de t .

Aunque se trata de una expansión en términos de potencias de $1/2$, no es la representación binaria del punto de imagen, ya que algunos coeficientes pueden ser negativos.

Comprobemos que la fórmula dada no depende del desarrollo cuaternario en los casos en los que un número t admite dos desarrollos diferentes, sería el caso de los cuaternarios finitos.

Lema 1.1. Un cuaternario finito también puede escribirse como un cuaternario infinito con la repetición de 3, es decir

$$0_4q_1q_2 \dots q_nq = 0_4q_1q_2 \dots q_n(q-1)\bar{3}$$

Demostración. Tenemos

$$\begin{aligned} 0_4q_1q_2 \dots q_nq &= \frac{q_1}{4} + \frac{q_2}{4^2} + \dots + \frac{q_n}{4^n} + \frac{q}{4^{n+1}} \\ 0_4q_1q_2 \dots q_n(q-1)\bar{3} &= \frac{q_1}{4} + \frac{q_2}{4^2} + \dots + \frac{q_n}{4^n} + \frac{q-1}{4^{n+1}} + \frac{3}{4^{n+2}} + \frac{3}{4^{n+3}} + \dots \end{aligned}$$

Como podemos ver, para que ambos desarrollos cuaternarios coincidan se debe de verificar

$$\frac{q}{4^{n+1}} = \frac{q-1}{4^{n+1}} + \frac{3}{4^{n+2}} + \frac{3}{4^{n+3}} + \dots = \frac{q}{4^{n+1}} - \frac{1}{4^{n+1}} + \frac{3}{4^{n+2}} + \frac{3}{4^{n+3}} + \dots$$

es decir, tenemos que ver que

$$-\frac{1}{4^{n+1}} + \frac{3}{4^{n+2}} + \frac{3}{4^{n+3}} + \dots = 0$$

pero eso lo tenemos pues

$$\begin{aligned} \frac{3}{4^{n+2}} + \frac{3}{4^{n+3}} + \frac{3}{4^{n+4}} + \dots &= \frac{3}{4^{n+2}} \left(1 + \frac{1}{4} + \frac{1}{4^2} + \dots \right) \\ &= \frac{3}{4^{n+2}} \sum_{m=0}^{\infty} \left(\frac{1}{4} \right)^m = \frac{3}{4^{n+2}} \left(\frac{1}{1 - 1/4} \right) \\ &= \frac{3}{4^{n+2}} \frac{4}{3} = \frac{1}{4^{n+1}} \end{aligned}$$

Obteniendo así el resultado que queríamos probar. |

Lema 1.2.

$$f_h(0_4 q_1 q_2 q_3 \dots q_n q) = f_h(0_4 q_1 q_2 q_3 \dots q_n (q-1)\bar{3})$$

Demostración. Por el lema anterior tenemos que $0_4 q_1 q_2 q_3 \dots q_n q = 0_4 q_1 q_2 q_3 \dots q_n (q-1)\bar{3}$, pero veamos que sus imágenes también coinciden. Recordando que

$$f_h(t) = \sum_{j=2}^{\infty} \frac{1}{2^j} H_0^{e(0,j)} H_3^{e(3,j)}(h_{q_j}) + \frac{1}{2} h_{q_1},$$

tenemos

$$f_h(0_4 q_1 q_2 q_3 \dots q_n q) = \sum_{j=2}^n \frac{1}{2^j} H_0^{e(0,j)} H_3^{e(3,j)}(h_{q_j}) + \frac{1}{2^{n+1}} H_0^{e(0,n+1)} H_3^{e(3,n+1)}(h_q) + \frac{1}{2} h_{q_1}$$

y volviendo a aplicar la fórmula anterior teniendo en cuenta el teorema 1.5

$$\begin{aligned} f_h(0_4 q_1 q_2 q_3 \dots q_n (q-1)\bar{3}) &= \sum_{j=2}^n \frac{1}{2^j} H_0^{e(0,j)} H_3^{e(3,j)}(h_{q_j}) + \frac{1}{2^{n+1}} H_0^{e(0,n+1)} H_3^{e(3,n+1)}(h_{q-1}) \\ &\quad + \sum_{j=n+2}^{\infty} \frac{1}{2^j} H_0^{e(0,n+1)} H_3^{e(3,n+1)} H_{q-1} H_{q_j} \dots H_{q_{j-1}}(h_3) + \frac{1}{2} h_{q_1} \end{aligned}$$

Recordando que $H_3^2(z) = z$ y simplificando, nos quedaría

$$\sum_{j=n+2}^{\infty} \frac{1}{2^j} H_0^{e(0,n+1)} H_3^{e(3,n+1)} H_{q-1} H_{q_j} \dots H_{q_{j-1}}(h_3) =$$

$$= \frac{1}{2^{n+2}} H_0^{e(0,n+1)} H_3^{e(3,n+1)} H_{q-1} \left(h_3 + \frac{1}{2} H_3(h_3) + \frac{1}{2^2} h_3 + \frac{1}{2^3} H_3(h_3) + \cdots \right)$$

Por otro lado, como

$$1 + \frac{1}{4} + \frac{1}{2^4} + \frac{1}{2^6} + \cdots = \sum_{n=0}^{\infty} \frac{1}{2^{2n}} = \sum_{n=0}^{\infty} \left(\frac{1}{4}\right)^n = \frac{1}{1-1/4} = \frac{4}{3}$$

y

$$\frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^5} + \cdots = \frac{1}{2} \sum_{n=0}^{\infty} \frac{1}{2^{2n}} = \frac{2}{3}$$

tenemos que

$$\begin{aligned} \left(I + \frac{1}{2} H_3 + \frac{1}{4} I + \frac{1}{8} H_3 + \cdots \right) (h_3) &= \left(\frac{4}{3} + \frac{2}{3} H_3 \right) (h_3) \\ &= \frac{4}{3} (2+i) + \frac{2}{3} H_3(2+i) \\ &= \frac{4}{3} (2+i) - \frac{2}{3} \overline{(2+i)} i \\ &= \frac{8}{3} + \frac{4}{3} i - \frac{4}{3} i - \frac{2}{3} = 2 \end{aligned}$$

Entonces nos quedaría

$$\begin{aligned} f_h(0_4 q_1 q_2 q_3 \cdots q_n (q-1)\bar{3}) &= \sum_{j=2}^n \frac{1}{2^j} H_0^{e(0,j)} H_3^{e(3,j)} (h_{q_j}) + \frac{1}{2^{n+1}} H_0^{e(0,n+1)} H_3^{e(3,n+1)} (h_{q-1}) \\ &\quad + \frac{1}{2^{n+2}} H_0^{e(0,n+1)} H_3^{e(3,n+1)} H_{q-1}(2) + \frac{1}{2} h_{q_1} \end{aligned}$$

o lo que es lo mismo

$$= \sum_{j=2}^n \frac{1}{2^j} H_0^{e(0,j)} H_3^{e(3,j)} (h_{q_j}) + \frac{1}{2^{n+1}} H_0^{e(0,n+1)} H_3^{e(3,n+1)} (h_{q-1} + \frac{1}{2} H_{q-1}(2)) + \frac{1}{2} h_{q_1}$$

Asimismo, también se verifica que

$$h_q = h_{q-1} + \frac{1}{2} H_{q-1}(2) \quad \text{para } q = 1, 2, 3.$$

veámoslo:

$$\begin{aligned} h_1 &= h_0 + \frac{1}{2} H_0(2) = 0 + \frac{1}{2} 2i = i \\ h_2 &= h_1 + \frac{1}{2} H_1(2) = i + \frac{1}{2} 2 = 1 + i \\ h_3 &= h_2 + \frac{1}{2} H_2(2) = 1 + i + \frac{1}{2} 2 = 2 + i \end{aligned}$$

Luego

$$f_h(0_4 q_1 q_2 q_3 \cdots q_n q) = f_h(0_4 q_1 q_2 q_3 \cdots q_n (q-1)\bar{3})$$

También se puede escribir la curva de Hilbert del siguiente modo.

Teorema 1.6. Si $t = 0_4 q_1 q_2 \dots$ entonces

$$f_h(t) = \sum_{j=1}^{\infty} \text{signo}(q_j) \sqrt{(q_j - 1)^2 + 1} \exp \left(i\pi \frac{d_j}{2} + i(-1)^{d_j} \arcsin \left(\frac{1}{\sqrt{(q_j - 1)^2 + 1}} \right) \right)$$

$$\text{donde } d_j = e(0, j) + e(3, j) \text{ mod } 2 \text{ y } \text{signo}(x) = \begin{cases} 0 & \text{si } x = 0 \\ 1 & \text{si } x > 0 \end{cases}.$$

Demostración. Por un lado tenemos

$$H_0(z) = i\bar{z} = \exp(i\pi/2)\bar{z}$$

$$H_1(z) = z$$

$$H_2(z) = z$$

$$H_3(z) = -i\bar{z} = \exp(-i\pi/2)\bar{z}$$

y como teníamos que $h_0 = 0$, $h_1 = i$, $h_2 = 1 + i$ y $h_3 = 2 + i$, entonces

$$h_n = \text{signo}(n)[(n - 1) + i]$$

utilizando coordenadas polares y que $\exp(ix) = \cos(x) + i\sin(x)$, nos quedaría

$$\begin{aligned} h_n &= \text{signo}(n) \sqrt{(n - 1)^2 + 1} \exp \left(i \arctg \left(\frac{1}{n - 1} \right) \right) \\ &= \text{signo}(n) \sqrt{(n - 1)^2 + 1} \exp \left[i \arcsen \left(\frac{1}{\sqrt{(n - 1)^2 + 1}} \right) \right] \end{aligned}$$

De todo este llegamos a que

$$H_3^{e(3,j)}(h_{q_j}) = \exp(-i\pi e(3, j)/2) \text{signo}(q_j)[(q_j - 1) + (-1)^{e(3,j)}i]$$

si además definimos

$$d_j = e(0, j) + e(3, j) \text{ mod } 2$$

nos quedaría

$$\begin{aligned} H_0^{e(0,j)} H_3^{e(3,j)}(h_{q_j}) &= \exp \left(\frac{i\pi e(0, j)}{2} \right) \exp \left(\frac{i\pi e(3, j)}{2} \right) \text{signo}(q_j)[(q_j - 1) + (-1)^{d_j}i] \\ &= (-1)^{e(3,j)} \exp \left(\frac{i\pi d_j}{2} \right) \text{signo}(q_j)[(q_j - 1) + (-1)^{d_j}i] \\ &= (-1)^{e(3,j)} \text{signo}(q_j) \sqrt{(q_j - 1)^2 + 1} \\ &\quad \times \exp \left[\frac{i\pi d_j}{2} + i(-1)^{d_j} \arcsen \left(\frac{1}{\sqrt{(q_j - 1)^2 + 1}} \right) \right] \end{aligned}$$

y sustituyendo este resultado en nuestra fórmula, nos produce la representación compleja de la curva de Hilbert que buscábamos. |

1.3. Forma aritmética de la curva de Hilbert

La descomposición de las transformaciones T_0, T_1, T_2, T_3 en parte real y parte imaginaria produce las siguientes transformaciones de similitud:

$$\begin{aligned}
 T_0 \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \frac{1}{2} H_0 \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} h_0 \\
 T_1 \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2} H_1 \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} h_1 \\
 T_2 \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2} H_2 \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} h_2 \\
 T_3 \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{1}{2} H_3 \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} h_3
 \end{aligned} \tag{1.2}$$

Observamos que en la nueva interpretación se sigue verificando

$$\lim_{n \rightarrow \infty} T_0^n(\mathcal{Q}) = \lim_{n \rightarrow \infty} \frac{1}{2^n} H_0^n(\mathcal{Q}) = \lim_{n \rightarrow \infty} \frac{1}{2^n} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^n (\mathcal{Q}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

y

$$f_h(0_4 q_1 q_2 q_3 \dots) = \sum_{j=1}^{\infty} \left(\frac{1}{2^j} \right) H_0^{e_{0j}} H_3^{e_{3j}}(h_{q_j}) \tag{1.3}$$

Es fácil ver que si $q_j = 0$, obtenemos el término $H_0^{e_{0j}} H_3^{e_{3j}}(h_0) = H_0^{e_{0j}} H_3^{e_{3j}} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Veamos ahora que para todos los posibles $(e(0, j), e(3, j))$ podemos escribir (1.3) del siguiente modo

| Teorema 1.7. *Si $t = 0_4 q_1 q_2 q_3 \dots$ entonces*

$$f_h(t) = \sum_{j=1}^{\infty} \left(\frac{1}{2^j} \right) (-1)^{e_{0j}} \operatorname{sgn}(q_j) \begin{pmatrix} (1 - d_j)q_j - 1 \\ 1 - d_j q_j \end{pmatrix}$$

Demostración. Tenemos que ver que para cada $(e_{0j}, e_{3j}) = (0, 0), (1, 0), (0, 1), (1, 1)$ y para cada $q_j = 1, 2, 3$ obtenemos el mismo resultado que en (1.3), para ello

$$H_0^{e_{0j}} H_3^{e_{3j}}(h_{q_j}) = (-1)^{e_{0j}} \begin{pmatrix} (1 - d_j)q_j - 1 \\ 1 - d_j q_j \end{pmatrix} \tag{1.4}$$

- Veámoslo primero para $(e_{0j}, e_{3j}) = (0, 0)$:

En este caso $d_j = 0$, por tanto el segundo término de (1.8) nos queda $\begin{pmatrix} q_j - 1 \\ 1 \end{pmatrix}$.

Para $q_j = 1$

$$H_0^0 H_3^0(h_1) = h_1 = i = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Para $q_j = 2$

$$H_0^0 H_3^0(h_2) = h_2 = 1 + i = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Para $q_j = 3$

$$H_0^0 H_3^0(h_3) = h_3 = 2 + i = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

- Si $(e_{0j}, e_{3j}) = (0, 1)$:

En este caso $d_j = 1$, por tanto el segundo término de (1.8) nos queda $\begin{pmatrix} -1 \\ 1 - q_j \end{pmatrix}$.

Para $q_j = 1$

$$H_0^0 H_3^1(h_1) = H_3(h_1) = H_3(i) = -(\bar{i})i = -1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Para $q_j = 2$

$$H_0^0 H_3^1(h_2) = H_3(h_2) = H_3(1 + i) = -(\overline{1 + i})i = -1 - i = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

Para $q_j = 3$

$$H_0^0 H_3^1(h_3) = H_3(h_3) = H_3(2 + i) = -(\overline{2 + i})i = -1 - 2i = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

- Si $(e_{0j}, e_{3j}) = (1, 0)$:

En este caso $d_j = 1$, por tanto el segundo término de (1.8) nos queda $-\begin{pmatrix} -1 \\ 1 - q_j \end{pmatrix}$.

Para $q_j = 1$

$$H_0^1 H_3^0(h_1) = H_0(h_1) = H_0(i) = \bar{i}i = -ii = 1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Para $q_j = 2$

$$H_0^1 H_3^0(h_2) = H_0(h_2) = H_0(1 + i) = (\overline{1 + i})i = 1 + i = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = -\begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

Para $q_j = 3$

$$H_0^1 H_3^0(h_3) = H_0(h_3) = H_0(2 + i) = (\overline{2 + 1})i = 1 + 2i = \begin{pmatrix} 1 \\ 2 \end{pmatrix} = -\begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

- Si $(e_{0j}, e_{3j}) = (1, 1)$:

En este caso $d_j = 0$, por tanto el segundo término de (1.8) nos queda $-\binom{q_j-1}{1}$.

Para $q_j = 1$

$$H_0^1 H_3^1(h_1) = H_0(-1) = -(\overline{-1})i = -i = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Para $q_j = 2$

$$H_0^1 H_3^1(h_2) = H_0(-1 - i) = (\overline{-1 - i})i = -1 - i = \begin{pmatrix} -1 \\ -1 \end{pmatrix} = -\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Para $q_j = 3$

$$H_0^1 H_3^1(h_3) = H_0(-1 - 2i) = (\overline{-1 - 2i})i = -2 - i = \begin{pmatrix} -2 \\ -1 \end{pmatrix} = -\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Probando así que se verifica el teorema. |

Dado que el cálculo de los numeradores de los términos en la suma que aparece en el teorema 1.7 solo implica la multiplicación y la resta de enteros, podemos programarlo para obtener las imágenes exactas de todos los datos finitos $t = 0_4 q_1 q_2 q_3 \dots q_n$ en forma de fracciones con 2^n en el denominador. De la misma manera, podemos calcular aproximaciones a las coordenadas de la imagen para cualquier $t \in \mathcal{I}$ con cualquier grado de precisión.

Veamos un ejemplo aplicando este último teorema: Encontremos $f_h(0_4 203)$, tenemos

$$e(0, 1) = 0, \quad e(0, 2) = 0, \quad e(0, 3) = 1, \quad e(3, 1) = 0, \quad e(3, 2) = 0, \quad e(3, 3) = 0$$

$$d_1 = 0, \quad d_2 = 0, \quad d_3 = 1$$

Por tanto

$$\begin{aligned} f_h(0_4 203) &= \frac{1}{2}(-1)^0 \binom{1, 2-1}{1} + 0 + \frac{1}{8}(-1)^1 \binom{0-1}{1-1, 3} = \\ &= \binom{1/2}{1/2} + \binom{1/8}{1/4} = \binom{5/8}{3/4} \end{aligned}$$

Esto representa el punto de entrada en el cuarto subcuadrado de la tercera partición dentro del primer subcuadrado de la segunda partición dentro del tercer subcuadrado de la primera partición. (Ver Fig. 1.3).

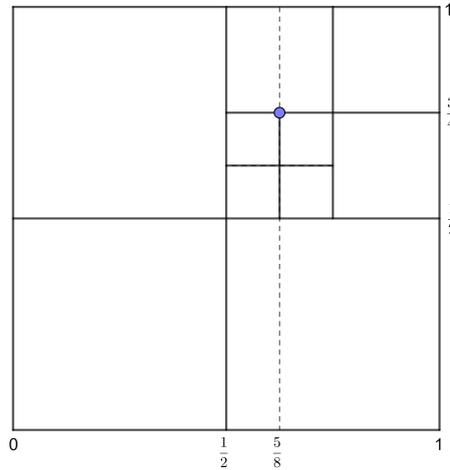


Figura 1.3. $f_h(0_4203)$

1.4. La curva de Hilbert no es diferenciable. Prueba analítica.

En la sección 1.2 hemos demostrado que la curva de Hilbert no es diferenciable en ningún punto de \mathcal{I} . Esta prueba se basó, en cierta medida, en la visión geométrica. Veamos en esta sección una prueba analítica:

Teorema 1.8. *La función $f_h : \mathcal{I} \rightarrow \mathcal{Q}$ no es diferenciable en ningún punto.*

Demostración. Denotemos $f_h = (\varphi_h, \psi_h)$ a las funciones coordenadas de f_h . Realizamos la prueba para la primera componente φ_h , ya que la prueba para ψ_h es análoga. Para abreviar, definimos $d(d_j, q_j)$ a la primera componente del vector bajo la suma de la función del teorema 1.7, es decir

$$d(d_j, q_j) = (1 - d_j)q_j - 1$$

entonces,

$$d(1, q_j) = -1 \quad \text{para } q_j = 1, 2, 3 \quad \text{y} \quad d(0, 1) = 0, \quad d(0, 2) = 1, \quad d(0, 3) = 2. \quad (1.5)$$

Nos queda que

$$\varphi_h(t) = \sum_{j=1}^{\infty} \frac{1}{2^j} (-1)^{e(0,j)} \text{sgn}(q_j) d(d_j, q_j) \quad (1.6)$$

Sea $t = 0_4q_1q_2q_3\dots \in \mathcal{I}$. Tenemos que distinguir tres casos:

1. Si hay infinitos dígitos 1 o 2 en la representación cuaternaria de t .

Si q_{k_1} representa el primer 1 o 2 después del punto cuaternario, q_{k_2} el segundo 1 o 2 ... y definimos

$$\underline{t}_n = 0_4q_1q_2q_3 \cdots q_{k_{n-1}}0, \quad \bar{t}_n = 0_4q_1q_2q_3 \cdots q_{k_{n-1}}3.$$

Entonces

$$\bar{t}_n - \underline{t}_n = \frac{q_1}{4} + \frac{q_2}{4^2} + \cdots + \frac{q_{k_{n-1}}}{4^{k_{n-1}}} + \frac{3}{4^{k_n}} - \left(\frac{q_1}{4} + \frac{q_2}{4^2} + \cdots + \frac{q_{k_{n-1}}}{4^{k_{n-1}}} \right) = \frac{3}{4^{k_n}}$$

Además como los k_{n-1} primeros términos de \underline{t} y \bar{t} son iguales, de (1.5) y (1.6) obtenemos,

$$|\varphi_h(\bar{t}_n) - \varphi_h(\underline{t}_n)| = \left| \frac{(-1)^{e(0,k_n)} d(d_{k_n}, 3)}{2^{k_n}} \right| \geq 1/2^{k_n}$$

De donde tenemos,

$$\left| \frac{\varphi_h(\bar{t}_n) - \varphi_h(\underline{t}_n)}{\bar{t}_n - \underline{t}_n} \right| = \frac{1/2^{k_n}}{3/4^{k_n}} = 2^{k_n}/3 \longrightarrow \infty \quad \text{si } n \longrightarrow \infty$$

es decir, φ_h no es diferenciable en t .

2. Si la representación en base cuatro de t contiene infinitos dígitos 3. Este caso incluye las representaciones en base cuatro finitas ($\neq 0$) ya que las podemos convertir en representaciones infinitas añadiendo infinitos 3 al final, como vimos en el lema 1.1.

En particular,

$$0_4\bar{3} = \frac{3}{4} \left(1 + \frac{1}{4} + \frac{1}{4^2} + \cdots \right) = \frac{3}{4} \sum_{n=0}^{\infty} \left(\frac{1}{4} \right)^n = \frac{3}{4} \left(\frac{1}{1 - 1/4} \right) = 1.$$

Sea q_{k_1} el primer 3 que aparece después del punto cuaternario y está precedido por un número par de 0 o 3, q_{k_2} el segundo 3 que está precedido por un número par de 0 o 3, ... Si

$$t = 0_4q_1q_2q_3 \cdots q_{k_{n-1}}3q_{k_{n+1}} \cdots$$

y definimos

$$t_{k_n} = 0_4 q_1 q_2 q_3 \cdots q_{k_n-1} 0.$$

Entonces,

$$t - t_{k_n} = 0_4 000 \cdots 03 q_{k_n+1} \cdots \leq 1/4^{k_n-1}$$

nuevamente por (1.6), teniendo en cuenta que los k_{n-1} primeros términos se repiten y que $d(d_{k_n}, 3) = 0$ pues q_{k_n} es el enésimo dígito 3 precedido por un número par de dígitos 0 o 3, tenemos

$$\begin{aligned} \varphi_h(t) - \varphi_h(t_{k_n}) &= (-1)^{e_0 k_n} / 2^{k_n} d(d_{k_n}, 3) \\ &+ \sum_{m=1}^{\infty} [(-1)^{e(0, k_n+m)} / 2^{k_n+m}] \text{signo}(q_{k_n+m}) d(d_{k_n+m}, q_{k_n+m}) \\ &= (-1)^{e_0 k_n} / 2^{k_n-1} + \sum_{m=1}^{\infty} [(-1)^{e(0, k_n+m)} / 2^{k_n+m}] \text{signo}(q_{k_n+m}) d(d_{k_n+m}, q_{k_n+m}) \\ &= (-1)^{e(0, k_n)} / 2^{k_n-1} \\ &+ [e(0, m) (-1)^{e(0, k_n)} / 2^{k_n}] \sum_{m=1}^{\infty} [(-1)^{e(0, m)} / 2^m] \text{signo}(q_{k_n+m}) d(d_m, q_{k_n+m}) \\ &= (-1)^{e(0, k_n)} / 2^{k_n-1} + [(-1)^{e(0, k_n)} / 2^{k_n}] \varphi_h(0_4 q_{k_n+1} q_{k_n+2} q_{k_n+3} \cdots). \end{aligned}$$

Como $|\varphi_h(t)| \leq 1$, obtenemos

$$\begin{aligned} |\varphi_h(t) - \varphi_h(t_{k_n})| &\geq |(-1)^{e(0, k_n)} / 2^{k_n-1}| + |(-1)^{e(0, k_n)} / 2^{k_n} \varphi_h(0_4 q_{k_n+1} q_{k_n+2} q_{k_n+3} \cdots)| \\ &\geq 1/2^{k_n-1} - 1/2^{k_n} = 1/2^{k_n} \end{aligned}$$

Por tanto

$$\left| \frac{\varphi(t) - \varphi(t_{k_n})}{t - t_{k_n}} \right| \geq \frac{1/2^{k_n}}{1/4^{k_n-1}} \geq 2^{k_n-2} \longrightarrow \infty \quad \text{si } n \longrightarrow \infty$$

Es decir, φ no es diferenciable en t .

3. Por último, el caso $t=0$.

Sea $t_n = 3/4^n$. Entonces $|0 - t_n| = 3/4^n$ y

$$|\varphi(0) - \varphi_h(t_n)| = | [(-1)^{e(0, n)} / 2^n] d(d_n, 3) | \geq 1/2^n$$

Por lo tanto,

$$\left| \frac{\varphi_h(0) - \varphi_h(t_n)}{0 - t_n} \right| \geq \frac{1/2^n}{3/4^n} = 2^n/3 \longrightarrow \infty \quad \text{si } n \longrightarrow \infty$$

entonces tenemos que φ_h no es diferenciable en 0.

Esto agota todas las posibilidades sobre t , por lo que se prueba el teorema. |

1.5. Aproximación por polígonos de la curva de Hilbert

Dentro de cada subcuadrado en la figura 1.1, la distancia entre los puntos de la curva de Hilbert y la línea poligonal está delimitada por la longitud de su diagonal, ya que en un cuadrado la máxima distancia entre dos puntos viene determinada por su diagonal. Si uno extiende estos polígonos continuamente por segmentos de línea desde su punto de inicio hasta el origen, y desde su punto final hasta el punto $(0, 1)$, entonces aproximan la curva de Hilbert uniformemente en \mathcal{I} .

| Definición 1.2. *La línea poligonal que une los puntos*

$$\{f(k/4^n) : k = 0, 1/2^{2n}, 2/2^{2n}, \dots, 1\}$$

se denomina poligonal aproximante de la curva de Hilbert de orden n .

En la Fig. 1.4 hemos dibujado los tres primeros polígonos aproximados, indicando la orientación de los segmentos individuales de la poligonal aproximante.

Podemos pensar que $p_n : \mathcal{I} \rightarrow \mathcal{Q}$ es una parametrización de dicha curva definida de manera natural, es decir,

$$p_n(t) = 2^{2n} \left(t - \frac{k}{2^{2n}} \right) f_h \left(\frac{k+1}{2^{2n}} \right) - 2^{2n} \left(t - \frac{k+1}{2^{2n}} \right) f_h \left(\frac{k}{2^{2n}} \right), \quad (1.7)$$

para $k/2^{2n} \leq t \leq (k+1)/2^{2n}$, $k = 0, 1, 2, 3, \dots, 2^{2n} - 1$, y podemos decir que

$$\|f_h - p_n\| \leq \sqrt{2}/2^n \quad \forall t \in \mathcal{I}$$

es decir, p_n converge uniformemente a la curva de Hilbert. Dado que los p_n son continuos, esto proporciona una prueba alternativa de la continuidad de la curva de Hilbert.

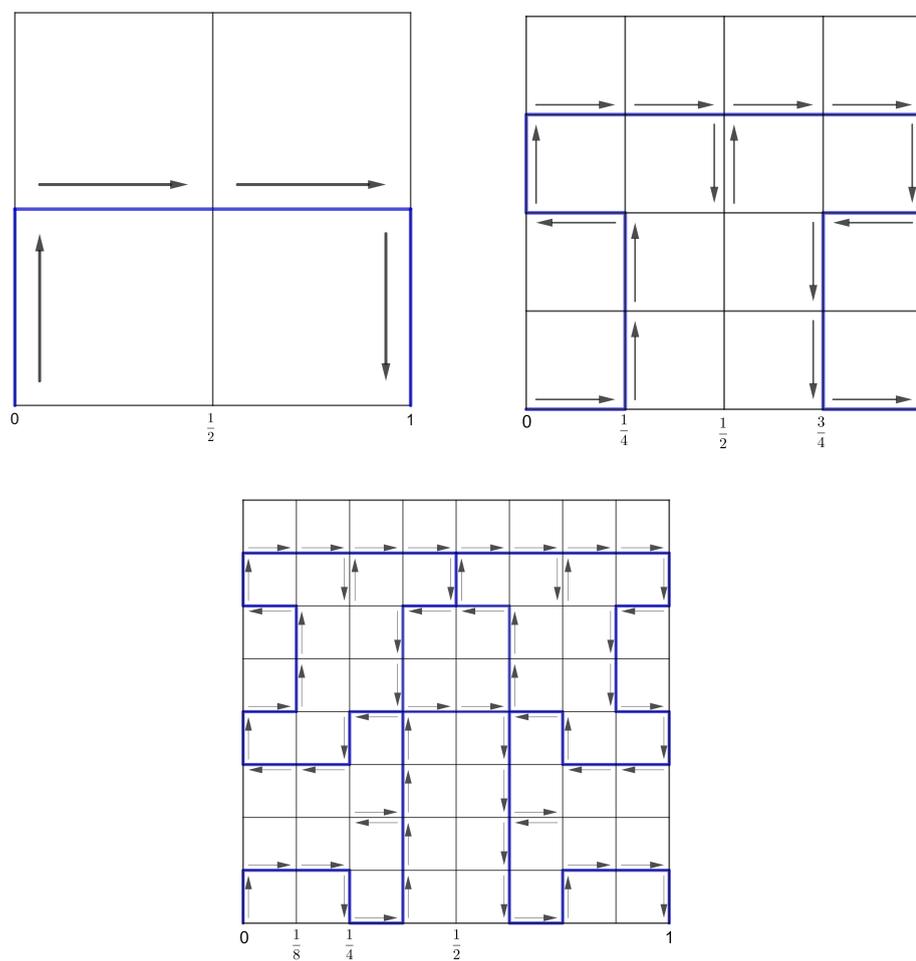


Figura 1.4. Poligonal aproximante de la curva de Hilbert

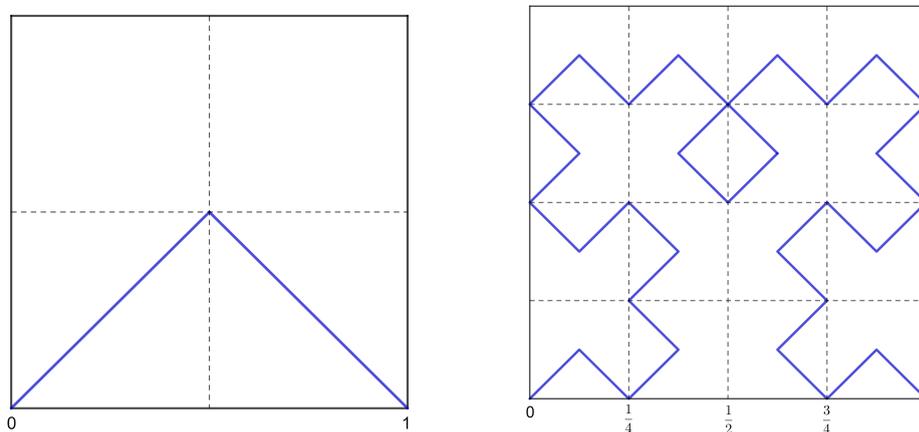


Figura 1.5. Triángulo isósceles como conjunto inicial

Las curvas aproximantes definidas en (1.7) y representados en la Fig. 1.4 para $n = 1, 2, 3$, pueden obtenerse recursivamente del segmento de línea (dirigido) de $(0,0)$ a $(1,0)$ por aplicación repetida de las transformaciones T_0, T_1, T_2, T_3 definidas en (1.2). Tomando límite cuando $n \rightarrow \infty$, se obtiene la curva de Hilbert. Los conjuntos de puntos que se obtienen de forma recursiva de esta manera se denominan de forma general como fractales. Por lo tanto, la curva de Hilbert se puede ver como un fractal con el segmento de línea de $(0,0)$ a $(1,0)$ como conjunto inicial.

Podemos observar que cualquier curva que entre en el cuadrado \mathcal{Q} en $(0,0)$ y salga en $(1,0)$ puede servir como conjunto inicial para la construcción de curvas aproximantes. Por ejemplo en la Fig. 1.5 hemos dibujado la tercera curva aproximante con un triángulo isósceles como conjunto inicial.

1.6. La curva de Hilbert tridimensional

Para definir una curva de Hilbert tridimensional, procedemos como en la Sección 1.1, excepto que ahora dividimos \mathcal{I} en cada paso en 2^{3n} subintervalos congruentes y, en consecuencia, el cubo unidad \mathcal{W} en 2^{3n} subcubos congruentes. Veamos su construcción:

1. Primero dividimos el intervalo unidad \mathcal{I} en $2^3 = 8$ subintervalos de igual

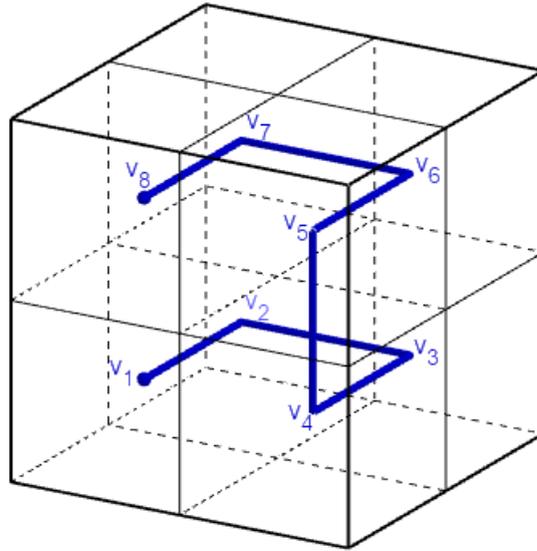


Figura 1.6. Primer paso de la curva tridimensional

longitud

$$\mathcal{I}_k = \left[\frac{k-1}{8}, \frac{k}{8} \right] \quad \text{donde } k = 1, 2, \dots, 8.$$

y el cubo unidad \mathcal{W} en 8 subcubos, \mathcal{W}_k con $k = 1, \dots, 8$, de longitud lateral $1/2$ y con centros respectivamente $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, $(\frac{1}{4}, \frac{3}{4}, \frac{1}{4})$, $(\frac{3}{4}, \frac{3}{4}, \frac{1}{4})$, $(\frac{3}{4}, \frac{1}{4}, \frac{1}{4})$, $(\frac{3}{4}, \frac{1}{4}, \frac{3}{4})$, $(\frac{3}{4}, \frac{3}{4}, \frac{3}{4})$, $(\frac{1}{4}, \frac{3}{4}, \frac{3}{4})$ y $(\frac{1}{4}, \frac{1}{4}, \frac{3}{4})$. Luego asignamos de forma continua cada subintervalo \mathcal{I}_k al subcubo \mathcal{W}_k .

2. Posteriormente repetimos este mismo proceso para cada subintervalo \mathcal{I}_k de manera que si un intervalo se corresponde con un cubo, cada uno de sus subintervalos se corresponde con un subcubo contenido en el cubo al que pertenecía el intervalo inicial, es decir, se sigue preservando la relación de inclusión.

Este proceso satisface los requisitos necesarios para la definición de una curva de relleno de espacio: subintervalos adyacentes asignados a subcubos adyacentes con una cara común y cada asignación conserva la anterior. El punto final de la curva obtenida de esta manera se encuentra una unidad por encima del punto inicial en dirección vertical.

Para obtener una representación aritmética-analítica de esta curva, procederemos

como en las secciones 1.2 y 1.3. Colocamos el sistema de coordenadas de modo que el origen coincida con la esquina frontal inferior izquierda de \mathcal{W} , el eje x apuntando hacia la derecha, el eje y hacia atrás y el eje z hacia arriba. Supondremos que el punto $t \in \mathcal{I}$ se representa en base 8, es decir,

$$t = 0_{\text{8}}w_1w_2w_3\dots, \quad \text{donde } w_j = 0, 1, \dots, 7.$$

El primer dígito w_1 nos indica que pertenece al intervalo \mathcal{I}_{w_1+1} de la primera etapa y este subintervalo hemos visto que estaba relacionado con el subcubo \mathcal{W}_{w_1+1} , los dígitos w_1w_2 nos indican que pertenecen al intervalo $\mathcal{I}_{w_1+1, w_2+1}$, es decir

$$t \in \left[\frac{(w_1 + 1) - 1}{8} + \frac{(w_2 + 1) - 1}{8^2}, \frac{(w_1 + 1) - 1}{8} + \frac{(w_2 + 1)}{8^2} \right]$$

y así sucesivamente.

Veamos un ejemplo, si $t = 0_{\text{8}}02$, entonces el dígito 0 nos está indicando que pertenece al subintervalo $\mathcal{I}_1 = [0, 1/8]$, entonces ya sabemos que nuestra t está en dicho intervalo, que además sabemos que está relacionado con el subcubo \mathcal{W}_1 que es el cubo de lado $1/2$ y centro $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. El siguiente dígito nos indica el subintervalo de \mathcal{I}_1 al que pertenece t , en este caso al subintervalo $\mathcal{I}_{13} = \left[\frac{(0+1)-1}{8} + \frac{(2+1)-1}{8^2}, \frac{(0+1)-1}{8} + \frac{(2+1)}{8^2} \right] = \left[\frac{2}{8^2}, \frac{3}{8^2} \right]$ que está relacionado con el subcubo \mathcal{W}_{13} , de radio $1/4$ y centro $(\frac{3}{8}, \frac{3}{8}, \frac{1}{8})$. Hemos llegado a que $t \in \left[\frac{2}{8^2}, \frac{3}{8^2} \right]$ y su imagen $f_h(t) \in \mathcal{W}_{13}$.

Llamamos T_0 a la aplicación que transforma \mathcal{W} en el primer subcubo de la primera partición, T_1 a la aplicación que transforma \mathcal{W} en el segundo subcubo, etc. Entonces

$$f_h(0_{\text{8}}w_1w_2w_3\dots w_n) = T_{w_1}T_{w_2}T_{w_3}\dots T_{w_n}(\mathcal{W})$$

Teniendo en cuenta que

$$0_{\text{8}}w_1w_2w_3\dots w_n = 0_{\text{8}}w_1w_2w_3\dots w_n000\dots,$$

obtenemos

$$f_h(0_{\text{8}}w_1w_2w_3\dots w_n) = T_{w_1}T_{w_2}T_{w_3}\dots T_{w_n}T_0T_0T_0\dots(\mathcal{W}). \quad (1.8)$$

Veamos quienes son dichas transformaciones:

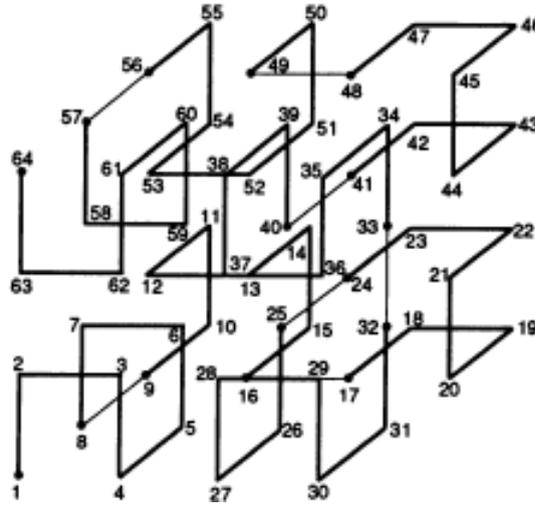


Figura 1.7. Segundo paso de la curva tridimensional.

Sagan, H., *Space-Filling Curves*. Springer-Verlag(1994).

Para obtener el primer subcubo orientado (Fig. 1.7) de \mathcal{W} , tenemos que transformar el polígono de la Fig. 1.6 en el que se encuentra en la esquina frontal inferior izquierda de la Fig. 1.7. Esto se logra mediante una reducción de las dimensiones de \mathcal{W} en la relación 2 : 1, una rotación de 90° sobre el eje x , y una reflexión con respecto al plano (x, z) :

$$\frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Para obtener el segundo subcubo, reducimos \mathcal{W} como antes, giramos 90° alrededor del eje y , reflexionamos con respecto al plano (y, z) , y trasladamos $1/2$ unidades en la dirección del eje y :

$$\frac{1}{2} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

De la misma forma, llegamos finalmente a

$$\begin{aligned}
 T_0 \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2} H_0 \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} h_0 \\
 T_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2} H_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} h_1 \\
 T_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2} H_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} h_2 \\
 T_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{2} H_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} h_3 \\
 T_4 \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{2} H_4 \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} h_4 \\
 T_5 \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{2} H_5 \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} h_5 \\
 T_6 \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \frac{1}{2} H_6 \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} h_6 \\
 T_7 \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \frac{1}{2} H_7 \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} h_7
 \end{aligned}$$

Ya que

$$T_0 T_0 T_0 \dots (\mathcal{W}) = \lim_{n \rightarrow \infty} T_0^n(\mathcal{W}) = \lim_{n \rightarrow \infty} \frac{1}{2^n} H_0^n(\mathcal{W}) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y suponiendo que H_{w_0} es la identidad, de (1.8) podemos obtener

$$f_h^{(3)}(0_8 w_1 w_2 w_3 \dots w_n) = \sum_{j=1}^n \frac{1}{2^j} H_{w_0} H_{w_1} H_{w_2} \dots H_{w_{j-1}}(h_{w_j})$$

y por consiguiente, tenemos

$$f_h^{(3)}(0_8 w_1 w_2 w_3 \dots) = \sum_{j=1}^{\infty} \frac{1}{2^j} H_{w_0} H_{w_1} H_{w_2} \dots H_{w_{j-1}}(h_{w_j})$$

Como los productos de matrices en general no son conmutativos no podemos simplificarlo como hicimos en la sección 1.3.

1.7. La curva de Peano

Siguiendo el mismo tipo de construcción que dimos para la curva de Hilbert, también podemos definir la curva de Peano, encontrada por G. Peano en 1890, como el límite de una construcción recursiva de iteraciones. Veamos su construcción:

1. Primero dividimos el intervalo unidad \mathcal{I} en nueve subintervalos de igual longitud $\mathcal{I}_1, \dots, \mathcal{I}_9$ y dividimos el cuadrado unidad \mathcal{Q} en nueve subcuadrados iguales cada uno con una longitud lateral de $1/3$, $\mathcal{Q}_1 \dots \mathcal{Q}_9$. Y asignamos cada \mathcal{I}_k con \mathcal{Q}_k con $k = 1 \dots 9$, de forma que subintervalos adyacentes se transformen en subcuadrados adyacentes. Ver Fig. 1.8.
2. Luego aplicamos sucesivamente el mismo procedimiento a cada subintervalo, de modo que se obtienen otras nueve subdivisiones por cada subdivisión anterior. Además se debe de preservar la relación de inclusión, es decir, si un intervalo se asigna a un cuadrado, cada uno de sus subintervalos se asigna a un subcuadrado del cuadrado anterior.

Para ver que la curva de Peano es una curva de relleno de espacio nos faltaría ver que la aplicación que lleva el intervalo unidad \mathcal{I} al cuadrado unidad \mathcal{Q} es continua y sobreyectiva, aunque dicha prueba es análoga a la prueba que dimos para la curva de Hilbert.

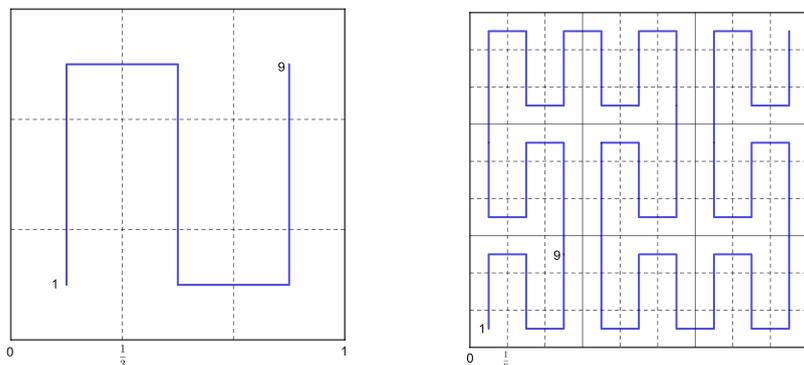


Figura 1.8. Construcción de la curva de Peano

1.8. La curva de Sierpinski

Hasta ahora las curvas que rellenan espacio que hemos visto se basan en una estructura recursiva de cuadrados, pero en este apartado estudiaremos una curva que rellena espacio que puede construirse geoméricamente utilizando una estructura recursiva basada en triángulos, la curva de Sierpinski. Esta curva lleva el nombre de Waclaw Sierpinski, quien en 1912 la presentó como la solución de ciertas ecuaciones funcionales.

Veamos su construcción:

1. Primero se divide el intervalo unidad \mathcal{I} en 2 intervalos iguales $\mathcal{I}_1 = [0, 1/2]$ y $\mathcal{I}_2 = [1/2, 1]$, y también dividimos el triángulo isósceles de vértices $(0, 0)$, $(1, 1)$, $(2, 0)$, al que denotaremos por \mathcal{T} , en 2 subtriángulos congruentes, de forma que la hipotenusa quede dividida en dos partes iguales, por lo que nos quedarían los triángulos \mathcal{T}_1 de vértices $(0, 0)$, $(1, 1)$, $(1, 0)$ y el triángulo \mathcal{T}_2 de vértices $(1, 0)$, $(1, 1)$, $(2, 0)$. Y asignamos cada \mathcal{I}_k con \mathcal{T}_k , para $k = 1, 2$.
2. Luego repetimos este procedimiento, asignando siempre cada subintervalo con un subtriángulo de manera que subintervalos adyacentes se asignen a subtriángulos adyacentes con un borde en común y preservando la relación de inclusión, es decir, si un intervalo se asigna a un triángulo, cada uno de sus subintervalos se asigna a un subtriángulo del triángulo anterior. Ver Fig. 1.9.

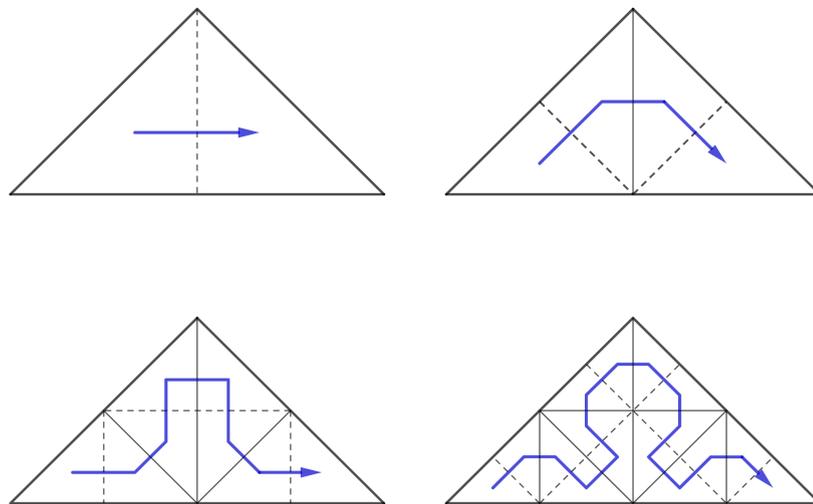


Figura 1.9. Construcción de la curva de Sierpinski

Como los subtriángulos se contraen a puntos, la curva empieza en $(0, 0)$ y termina en $(2, 0)$. Por tanto, para que la curva de Sierpinski sea una curva de relleno de espacio solo nos faltaría ver que es continua y sobreyectiva, pero dicha prueba es análoga a la prueba que dimos para la curva de Hilbert.

Capítulo 2

El teorema de Hahn-Mazurkiewicz

Hasta ahora en el capítulo 1 hemos visto la construcción de varias curvas que rellenan espacio y todas ellas se basaban en la división y asignación del intervalo $[0, 1]$ a un subconjunto de \mathbb{R}^2 o \mathbb{R}^3 (generalmente el cuadrado unidad). Aunque la curva de Lebesgue también es una curva que rellena el espacio, su construcción es totalmente diferente al resto de curvas, por ello la estudiaremos en este capítulo, además esta curva nos da pie para estudiar que tipos de conjuntos del plano pueden ser la imagen continua del intervalo unidad, es decir, nos da pie para llegar a nuestro principal objetivo, probar el teorema de Hahn-Mazurkiewicz, con el cual concluiremos este capítulo.

2.1. El conjunto de Cantor

El *Conjunto de Cantor*, definido por Georg Cantor en 1883, es un subconjunto fractal del intervalo unidad \mathcal{I} . El conjunto de Cantor, \mathcal{C} , se construye mediante el límite de una sucesión de conjuntos \mathcal{C}_n , con $n \in \mathbb{N}$.

Primero dividimos el intervalo unidad en tres intervalos iguales y eliminamos el intervalo central (quedándonos siempre con los extremos). De esta forma nos queda

$$\mathcal{C}_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

Seguidamente, repetimos el proceso anterior a cada uno de los intervalos que forman \mathcal{C}_1 , es decir, dividimos cada intervalo en tres partes iguales y volvemos a eliminar

las partes centrales. Así llegamos a

$$\mathcal{C}_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

Repetimos este proceso indefinidamente. Es claro que cada término de la sucesión está contenido en el término que le precede, es decir $\mathcal{C}_{n+1} \subset \mathcal{C}_n$ para $n \in \mathbb{N}$. El conjunto de Cantor se define como la intersección de todos ellos:

$$\mathcal{C} = \bigcap_{n=0}^{\infty} \mathcal{C}_n$$

Como ya sabemos, todos los números de \mathcal{I} se pueden representar en forma ternaria, $t = 0_3 t_1 t_2 t_3 \dots$ donde $t_j = 0, 1$ ó 2 . Luego, en el primer paso, al dividir \mathcal{I} en tres intervalos y eliminar el intervalo central, estamos eliminando todos los números de la forma $0_3 1 t_2 t_3 t_4 \dots$ excepto el $0_3 1$, que lo podemos reescribir por $0_3 0\bar{2}$, dado que en la primera iteración estamos eliminando el intervalo $(\frac{1}{3}, \frac{2}{3})$ y

$$0_3 0\bar{2} = 2 \left(\frac{1}{3^2} + \frac{1}{3^3} + \frac{1}{3^4} + \dots \right) = \frac{2}{3} \sum_{n=1}^{\infty} (1/3)^n = \frac{2}{3} \frac{1/3}{1 - 1/3} = \frac{1}{3} = 0_3 1.$$

En el siguiente paso, eliminamos todos los puntos de la forma $0_3 01 t_3 t_4 \dots$ y $0_3 21 t_3 t_4 \dots$, excepto $0_3 01 = \frac{1}{9} = 0_3 00\bar{2}$ y $0_3 21 = \frac{7}{9} = 0_3 20\bar{2}$, ya que estamos eliminando todos los puntos de los intervalos $(\frac{1}{9}, \frac{2}{9})$ y $(\frac{7}{9}, \frac{8}{9})$.

Veamos que de forma general se verifica que $0_3 t_1 \dots t_n 1 = 0_3 t_1 \dots t_n 0\bar{2}$

$$\begin{aligned} 0_3 t_1 \dots t_n 0\bar{2} &= \frac{t_1}{3} + \dots + \frac{t_n}{3^n} + 0 + \frac{2}{3^{n+2}} + \frac{2}{3^{n+3}} + \dots \\ &= \frac{t_1}{3} + \dots + \frac{t_n}{3^n} + \frac{2}{3^{n+2}} \sum_{j=0}^{\infty} \left(\frac{1}{3}\right)^j \\ &= \frac{t_1}{3} + \dots + \frac{t_n}{3^n} + \frac{2}{3^{n+2}} \left(\frac{1}{1 - 1/3} \right) \\ &= \frac{t_1}{3} + \dots + \frac{t_n}{3^n} + \frac{2}{3^{n+2}} \frac{3}{2} \\ &= \frac{t_1}{3} + \dots + \frac{t_n}{3^n} + \frac{1}{3^{n+1}} = 0_3 t_1 \dots t_n 1 \end{aligned}$$

| Definición 2.1. Sea \mathcal{C}^* el conjunto de puntos que tienen una representación decimal en base 3 sin contener el dígito 1.

En el siguiente lema veremos que en realidad $\mathcal{C} = \mathcal{C}^*$.

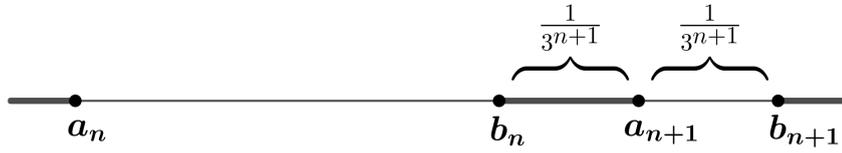


Figura 2.1. Construcción del conjunto de Cantor

Lema 2.1. $t \in \mathcal{C}$ si y solo si al menos una representación en base 3 de t no contiene ningún dígito 1, es decir, $\mathcal{C} = \mathcal{C}^*$.

Demostración. Hemos visto que para $n = 1$ se elimina el intervalo (a_1, b_1) si $a_1 = 0_31 = \frac{1}{3}$ y $b_1 = 0_32 = \frac{2}{3}$. Para $n = 2$ se elimina el intervalo (a_2, b_2) si $a_2 = 0_301 = \frac{1}{3^2}$ y $b_2 = 0_302 = \frac{2}{3^2}$ o $a_2 = 0_321 = \frac{2}{3} + \frac{1}{3^2}$ y $b_2 = 0_322 = \frac{2}{3} + \frac{2}{3^2}$. En general, se elimina el intervalo (a_n, b_n) si y solo si a_n y b_n pueden escribirse de la forma

$$\begin{aligned} a_n &= 0_3(2t_1)(2t_2)(2t_3) \dots (2t_{n-1})1 \\ b_n &= 0_3(2t_1)(2t_2)(2t_3) \dots (2t_{n-1})2 \end{aligned} \tag{2.1}$$

Supongamos que esto es cierto. Vamos a demostrar que también se verifica para $n + 1$. En el siguiente paso podríamos tener

$$a_{n+1} = b_n + \frac{1}{3^{n+1}} \quad y \quad b_{n+1} = a_{n+1} + \frac{1}{3^{n+1}}$$

(ver Fig. 2.1) y por tanto,

$$\begin{aligned} a_{n+1} &= 0_3(2t_1)(2t_2)(2t_3) \dots (2t_{n-1})2 + \frac{1}{3^{n+1}} \\ &= 0_3(2t_1)(2t_2)(2t_3) \dots (2t_{n-1})21 \\ b_{n+1} &= a_{n+1} + \frac{1}{3^{n+1}} \\ &= 0_3(2t_1)(2t_2)(2t_3) \dots (2t_{n-1})21 + \frac{1}{3^{n+1}} \\ &= 0_3(2t_1)(2t_2)(2t_3) \dots (2t_{n-1})22 \end{aligned}$$

o podríamos tener

$$\begin{aligned}
 a_{n+1} &= a_n - \frac{2}{3^{n+1}} \\
 &= 0_3(2t_1)(2t_2)(2t_3) \dots (2t_{n-1})1 - \frac{2}{3^{n+1}} \\
 &= 0_3(2t_1)(2t_2)(2t_3) \dots (2t_n - 1)01 \\
 b_{n+1} &= a_n - \frac{1}{3^{n+1}} \\
 &= 0_3(2t_1)(2t_2)(2t_3) \dots (2t_{n-1})1 - \frac{1}{3^{n+1}} \\
 &= 0_3(2t_1)(2t_2)(2t_3) \dots (2t_{n-1})02
 \end{aligned}$$

Volviéndose a verificar (2.1) en este caso para $n+1$. Por inducción tenemos que (2.1) se verifica para todo n .

Si $t \in (a_n, b_n)$, entonces

$$t = 0_3(2t_1)(2t_2)(2t_3) \dots (2t_{n-1})1\tau_{n+1}\tau_{n+2} \dots$$

Si todos los $\tau_{n+j} = 0$ tendríamos que $t = a_n \notin (a_n, b_n)$, o si todos los $\tau_{n+j} = 2$ tendríamos que $t = b_n \notin (a_n, b_n)$. Por lo tanto, al menos uno de los $\tau_{n+j} \neq 0$ y al menos uno de los $\tau_{n+j} \neq 2$.

De modo que un punto t se elimina en la construcción de \mathcal{C} si y solo si sus representaciones ternarias contienen al menos un dígito 1, demostrando así el teorema. |

Así, podemos decir que el Conjunto de Cantor es el conjunto de los puntos de \mathcal{I} que pueden escribirse en base 3 sin contener al 1, es decir

$$\mathcal{C} = \left\{ t \in \mathcal{I} : t = \sum_{n=1}^{\infty} \frac{t_n}{3^n}, \text{ donde } t_n = 0, 2 \forall n \in \mathbb{N} \right\}.$$

Este conjunto también se puede representar por

$$\mathcal{C} = \{0_3(2t_1)(2t_2)(2t_3) \dots, \text{ donde } t_j = 0, 1\}.$$

2.2. Propiedades del conjunto de Cantor

El conjunto de Cantor es compacto, pues es cerrado (por ser intersección numerable de cerrados) y acotado (pues $\mathcal{C} \in \mathcal{I}$).

Lema 2.2. El conjunto de Cantor, \mathcal{C} , es perfecto, es decir, es un conjunto cerrado tal que todos sus puntos son puntos de acumulación. Además todos los puntos de la forma $0_{\dot{3}}(2t_1)(2t_2)\dots(2t_n)\bar{2} \in \mathcal{C}$ son puntos de acumulación por la derecha, todos los puntos de la forma $0_{\dot{3}}(2t_1)(2t_2)\dots(2t_n) \in \mathcal{C}$ son puntos de acumulación por la izquierda y el resto de puntos son puntos de acumulación por la derecha y por la izquierda.

Demostración. Sabemos que el conjunto \mathcal{C} es cerrado. Falta ver que todos sus puntos son puntos de acumulación.

- Sea

$$t = 0_{\dot{3}}(2t_1)(2t_2)\dots(2t_n)\bar{2}$$

y sea

$$t^{(k)} = 0_{\dot{3}}(2t_1)(2t_2)\dots(2t_n)22\dots20\bar{2}$$

donde el dígito 0 ocupa la posición $(n+k)$. Luego la sucesión $t^{(k)} \in \mathcal{C}$ y $\{t^{(k)}\}$ tiende a t por la izquierda. Por tanto t es un punto de acumulación por la izquierda.

- Sea

$$t = 0_{\dot{3}}(2t_1)(2t_2)\dots(2t_n)$$

y sea

$$t^{(k)} = 0_{\dot{3}}(2t_1)(2t_2)\dots(2t_n)00\dots02$$

donde el dígito 2 ocupa la posición $(n+k)$. Luego la sucesión $t^{(k)} \in \mathcal{C}$ y $\{t^{(k)}\}$ tiende a t por la derecha. Por tanto t es un punto de acumulación por la derecha.

- Por último, sea

$$t = 0_{\dot{3}}(2t_1)(2t_2)\dots(2t_n)\dots$$

con infinitos dígitos 0 y 2, vamos a dar dos sucesiones, tal que una tienda a t por la izquierda y otra por la derecha. Sea

$$t^{(k)} = 0_{\dot{3}}(2t_1)(2t_2)\dots(2t_{i-1})0(2t_{i+1})\dots(2t_n)\dots$$

el mismo número que t , salvo que el k -ésimo dígito 2 lo sustituimos por 0. Es decir, si

$$t = 0_3 022002 \dots$$

entonces

$$t^{(1)} = 0_3 002002 \dots$$

$$t^{(2)} = 0_3 020002 \dots$$

$$t^{(3)} = 0_3 022000 \dots$$

y así sucesivamente. De esta forma tenemos que la sucesión $\{t^{(k)}\}$ pertenece a \mathcal{C} y además tiende a t por la izquierda.

Siguiendo un razonamiento similar definimos la sucesión $\{\tau^{(k)}\}$ donde

$$\tau^{(k)} = 0_3 (2t_1)(2t_2) \dots (2t_{j-1})2(2t_{j+1}) \dots (2t_n) \dots$$

en este caso sustituimos el k -ésimo dígito 0 por un 2. Es decir, si

$$\tau = 0_3 022002 \dots$$

entonces

$$\tau^{(1)} = 0_3 222002 \dots$$

$$\tau^{(2)} = 0_3 022202 \dots$$

$$\tau^{(3)} = 0_3 022022 \dots$$

y así sucesivamente. Luego la sucesión $\{\tau^{(k)}\}$ pertenece a \mathcal{C} y además tiende a t por la derecha.

Por tanto, en este caso t es un punto de acumulación tanto por la derecha como por la izquierda.

■

Lema 2.3. \mathcal{C} tiene medida nula, $m(\mathcal{C}) = 0$.

Demostración. Si nos fijamos en la primera construcción que dimos del conjunto de Cantor, partimos de un intervalo de medida 1. En la primera iteración, nos quedamos

con 2 intervalos de medida $\frac{1}{3}$, en la segunda teníamos 2^2 intervalos de medida $\frac{1}{3^2}$, ... Es decir, cada conjunto \mathcal{C}_n está formado por 2^n intervalos de medida $\frac{1}{3^n}$. Luego,

$$m(\mathcal{C}_n) = 2^n \frac{1}{3^n} = \left(\frac{2}{3}\right)^n$$

Por tanto, la medida del conjunto de Cantor es

$$m(\mathcal{C}) = \lim_{n \rightarrow \infty} m(\mathcal{C}_n) = \lim_{n \rightarrow \infty} \left(\frac{2}{3}\right)^n = 0$$

pues sabemos que si $A_n \subset A_{n+1} \forall n > 0$, entonces

$$m\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} m(A_n)$$

.

Definimos las aplicaciones continuas, $f : \mathcal{C} \rightarrow \mathcal{Q}$ y $F : \mathcal{C} \rightarrow \mathcal{W}$, como

$$f(0_3(2t_1)(2t_2)(2t_3)\dots) = \begin{pmatrix} 0_2 t_1 t_3 t_5 \dots \\ 0_2 t_2 t_4 t_6 \dots \end{pmatrix} \quad y \quad F(0_3(2t_1)(2t_2)(2t_3)\dots) = \begin{pmatrix} 0_2 t_1 t_4 t_7 \dots \\ 0_2 t_2 t_5 t_8 \dots \\ 0_2 t_3 t_6 t_9 \dots \end{pmatrix} \quad (2.2)$$

donde $t_j = 0, 1$. Es claro que estas aplicaciones no son inyectivas, veamos un ejemplo:

Si $t = 0_3 2$ entonces

$$f(0_3 2) = \begin{pmatrix} 0_2 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}$$

pero también tenemos que para $t = 0_3 0\bar{0}2$ se tiene

$$f(0_3 0\bar{0}2) = \begin{pmatrix} 0_2 0\bar{1} \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}$$

esto es debido a que

$$0_2 0\bar{1} = \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \dots = \frac{1}{2} \frac{1/2}{1 - 1/2} = \frac{1}{2}.$$

Luego f no puede ser inyectiva. Un ejemplo parecido nos prueba que F no puede ser inyectiva:

Sea de nuevo $t = 0_3 2$ entonces

$$F(0_3 2) = \begin{pmatrix} 0_2 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 0 \\ 0 \end{pmatrix}$$

y sea $t = 0_3\overline{0002}$, entonces

$$F(0_3\overline{0002002\dots}) = \begin{pmatrix} 0_2\overline{01} \\ 0_2\overline{0} \\ 0_2\overline{0} \end{pmatrix} = \begin{pmatrix} 1/2 \\ 0 \\ 0 \end{pmatrix}$$

por el mismo razonamiento que en el ejemplo anterior. Por tanto F tampoco puede ser inyectiva, pero veamos que si son sobreyectivas, continuas y que no son diferenciables en ningún punto.

Teorema 2.1. *Las aplicaciones definidas en (2.1) son sobreyectivas, continuas y no son diferenciables en ningún punto de \mathcal{I} .*

Demostración. Lo probaremos para la aplicación $f : \mathcal{C} \rightarrow \mathcal{Q}$, ya que para el otro caso la demostración es análoga.

- Veamos primero que la aplicación es sobreyectiva:

Sea $p \in \mathcal{Q}$. Podemos representar sus coordenadas en base 2 de la siguiente forma

$$p = \begin{pmatrix} 0_2 a_1 a_2 a_3 \dots \\ 0_2 b_1 b_2 b_3 \dots \end{pmatrix}.$$

Si $t = 0_3(2a_1)(2b_1)(2a_2)(2b_2)(2a_3)(2b_3)\dots$, entonces por (2.1) tenemos que $f(t) = p$. Luego la aplicación $f : \mathcal{C} \rightarrow \mathcal{Q}$ es sobreyectiva.

- Probemos ahora que es continua:

Suponemos que $|t - t_0| < \frac{1}{3^{2n}}$. Vamos a demostrar que t y t_0 coinciden en los primeros $2n$ términos, es decir

$$t_0 = 0_3(2t_1)(2t_2)\dots(2t_{2n})(2t_{2n+1})\dots \quad y$$

$$t = 0_3(2t_1)(2t_2)\dots(2t_{2n})(2\tau_{2n+1})\dots$$

Supongamos que $t_{2k} \neq \tau_{2k} \forall k = 1, \dots, n$, entonces $|t_{2k} - \tau_{2k}| = 1$ y

$$t - t_0 = \frac{2\tau_{2k} - 2t_{2k}}{3^{2k}} + \frac{2\tau_{2n+1} - 2t_{2n+1}}{3^{2n+1}} + \dots$$

como $|t_{2n+i} - \tau_{2n+i}| \leq 1 \forall i$, entonces

$$|t - t_0| \geq \frac{2}{3^{2k}} - \frac{2}{3^{2n+1}}(1 + \frac{1}{3} + \frac{1}{9} + \dots) = \frac{2}{3^{2k}} - \frac{1}{3^{2n}} \geq \frac{1}{3^{2n}} \quad \forall k = 1, \dots, n$$

Hemos llegado a una contradicción, pues por hipótesis $|t - t_0| < 1/3^{2n}$, luego $t_{2k} = \tau_{2k} \forall k = 1, \dots, n$. Tenemos que

$$f(t) - f(t_0) = \begin{pmatrix} (\tau_{2n+1} - t_{2n+1})/2^{n+1} + (\tau_{2n+3} - t_{2n+3})/2^{n+2} + \dots \\ (\tau_{2n+2} - t_{2n+2})/2^{n+1} + (\tau_{2n+4} - t_{2n+4})/2^{n+2} + \dots \end{pmatrix}$$

de donde obtenemos

$$\|f(t) - f(t_0)\| < \frac{\sqrt{2}}{2^n}$$

lo que prueba la continuidad. Además como \mathcal{I} es un compacto y f hemos probado que es continua, tenemos que f es uniformemente continua.

- Por último, vamos a demostrar que la aplicación no es diferenciable en ningún punto:

Sea t un punto cualquiera de \mathcal{I}

$$t = 0_3(2t_1)(2t_2) \dots (2t_{2n})(2t_{2n+1})(2t_{2n+2}) \dots,$$

y sean

$$t_n = 0_3(2t_1)(2t_2) \dots (2t_{2n})(2\tau_{2n+1})(2t_{2n+2}) \dots,$$

donde $\tau_{2n+1} = t_{2n+1} + 1 \pmod{2}$. Entonces

$$|t - t_n| = \frac{2}{3^{2n+1}}.$$

Si denotamos por φ a la primera componente de f , nos queda

$$\varphi(t) - \varphi(t_n) = \frac{t_{2n+1} - \tau_{2n+1}}{2^{n+1}}$$

y, por lo tanto

$$\frac{|\varphi(t) - \varphi(t_n)|}{|t - t_n|} = \frac{3^{2n+1}}{2^{n+2}} = \frac{3}{4} \left(\frac{9}{2}\right)^n \rightarrow \infty,$$

Hemos visto entonces que φ no es diferenciable en \mathcal{C} . De forma análoga podemos probarlo para la segunda componente de f . Lo que nos lleva a que la aplicación $f : \mathcal{C} \rightarrow \mathcal{Q}$ no es diferenciable en ningún punto.

■

2.3. Definición de la curva de Lebesgue

En 1904, Lebesgue extendió por linealidad la aplicación $f : \mathcal{C} \rightarrow \mathcal{Q}$ a todo el intervalo \mathcal{I} de la siguiente forma:

Si (a, b) es uno de los intervalos abiertos que hemos quitado en la construcción de \mathcal{C} , entonces f queda definida en ese intervalo por

$$f(t) = \frac{1}{b-a}(f(a)(b-t) + f(b)(t-a))$$

para todo $t \in (a, b)$. Esto se puede hacer debido a que \mathcal{Q} es un convexo y por ello, si dos puntos pertenecen a \mathcal{Q} entonces también pertenece todo el segmento que los une.

| Teorema 2.2. *La curva construida por Lebesgue $f : \mathcal{I} \rightarrow \mathcal{Q}$ es continua, sobreyectiva y diferenciable en casi todo punto de \mathcal{I} .*

Demostración. ■ Veamos primero que es continua. Para ello tenemos que demostrar que f es continua en todo punto $t_0 \in \mathcal{I}$. Distingamos casos:

- Si $t_0 \in \mathcal{C}^c = [0, 1] \setminus \mathcal{C}$, f es continua en t_0 por construcción.
- Si $t_0 \in \mathcal{C}$, por el lema 2.2 sabemos que t_0 es un punto de acumulación por la izquierda, por la derecha o por ambos lados, por tanto distingamos los siguientes casos:
 - t_0 punto de acumulación de \mathcal{C} por la derecha.
 - t_0 punto de acumulación de \mathcal{C} por la izquierda.
 - t_0 punto de acumulación de \mathcal{C} por ambos lados.

Vamos a demostrar solamente el primer caso ya que la prueba para los casos restantes es semejante. Si t_0 es un punto de acumulación por la derecha, entonces es el extremo derecho de uno de los intervalos que quitamos en la construcción de \mathcal{C} .

La continuidad por la izquierda es inmediata, ya que estaríamos dentro del intervalo quitado y f restringida a \mathcal{C}^c estaba definida por linealidad, luego se ve claro que es continua y como t_0 es el extremo derecho

$$\lim_{t \rightarrow t_0^-} f(t) = f(t_0).$$

Vamos a demostrar que f en t_0 es continua por la derecha. Sea $\varepsilon > 0$ y sea $\delta_1 > 0$ el número que existe por la continuidad uniforme sobre \mathcal{C} de f , de modo que $\|f(t) - f(t')\| < \varepsilon$ siempre que $\forall t, t' \in \mathcal{C}$ y $|t - t'| < \delta_1$. Es seguro que existe $t_1 \in \mathcal{C}$ tal que $t_0 < t_1 < t_0 + \delta_1$ ya que t_0 es punto de acumulación por la derecha. Tomamos $\delta = t_1 - t_0 > 0$. Vamos a probar que para todo $t \in (t_0, t_0 + \delta)$ ocurre que $\|f(t) - f(t_0)\| < \varepsilon$. Sea pues t un punto de este intervalo, tenemos dos posibilidades $t \in \mathcal{C}$ o $t \in \mathcal{C}^c$.

Si $t \in \mathcal{C}$, como $t_0 \in \mathcal{C}$, por lo que acabamos de ver tenemos

$$\|f(t) - f(t_0)\| < \varepsilon \quad \text{para } |t - t_0| < \delta. \quad (2.3)$$

Si $t \in \mathcal{C}^c$, entonces t pertenece a uno de los intervalos que quitamos en la construcción de \mathcal{C} , digamos $t \in (a, b)$, donde $(a, b) \subset \mathcal{C}^c$ y $a, b \in \mathcal{C}$, además (a, b) debe cumplir $t_0 < a < t < b < t_0 + \delta = t_1$, de esta forma $|a - t_0| < \delta$ y $|b - t_0| < \delta$ entonces como $a, b, t_0 \in \mathcal{C}$ podemos aplicar (2.3), es decir, $\|f(b) - f(t_0)\| < \varepsilon$ y $\|f(a) - f(t_0)\| < \varepsilon$.

Por la definición de la curva de Lebesgue y teniendo en cuenta que $t \in \mathcal{C}^c$ y $t_0 \in \mathcal{C}$

$$\begin{aligned} f(t) - f(t_0) &= \frac{1}{b-a}(f(a)(b-t) + f(b)(t-a)) - f(t_0) \\ &= \frac{1}{b-a}[f(a)(b-t) + f(b)(t-a) - f(t_0)(b-a)] \\ &= \frac{1}{b-a}[f(a)(b-t) + f(b)(t-a) - f(t_0)(b-a) + f(t_0)t - f(t_0)t] \\ &= \frac{1}{b-a}[f(a)(b-t) + f(b)(t-a) - f(t_0)(b-t) - f(t_0)(t-a)] \\ &= \frac{1}{b-a}[(f(a) - f(t_0))(b-t) + (f(b) - f(t_0))(t-a)] \end{aligned}$$

para todo $t \in (a, b)$. Entonces nos queda

$$\begin{aligned} \|f(t) - f(t_0)\| &\leq \frac{1}{b-a}(\|f(a) - f(t_0)\|(b-t) + \|f(b) - f(t_0)\|(t-a)) \\ &< \frac{1}{b-a}(b-t+t-a)\varepsilon = \varepsilon \end{aligned}$$

para todo $t \in (a, b) \subset \mathcal{C}^c$.

Luego hemos probado la continuidad de f en cualquier punto de acumulación por la derecha.

- Veamos ahora que f es sobreyectiva.

Tenemos que f restringida a \mathcal{C}^c es sobreyectiva por construcción. Por otro lado, en el teorema 2.1 probamos que \mathcal{C} era sobreyectiva, por consiguiente, f restringida a \mathcal{C} es sobreyectiva. Con lo cual probamos la sobreyectividad de f en \mathcal{I} .

- Por último probemos que f es diferenciable en casi todo punto de \mathcal{I} :

Por el teorema 2.1 sabemos que f restringida a \mathcal{C} no es diferenciable en ningún punto, pero como \mathcal{C} es un conjunto de medida nula, nos basta probar que f es diferenciable en \mathcal{C}^c .

Por definición de f , vemos claramente que es derivable en dicho conjunto. Sea $t_0 \in (a, b)$,

$$\begin{aligned} f'(t_0) &= \lim_{t \rightarrow t_0} \frac{f(t) - f(t_0)}{t - t_0} \\ &= \lim_{t \rightarrow t_0} \frac{f(a)(b-t) + f(b)(t-a) - f(a)(b-t_0) - f(b)(t_0-a)}{(b-a)(t-t_0)} \\ &= \lim_{t \rightarrow t_0} \frac{f(a)(b-t-b+t_0) + f(b)(t-a-t_0+a)}{(b-a)(t-t_0)} \\ &= \lim_{t \rightarrow t_0} \frac{(f(b) - f(a))(t-t_0)}{(b-a)(t-t_0)} = \lim_{t \rightarrow t_0} \frac{f(b) - f(a)}{b-a} \\ &= \frac{f(b) - f(a)}{b-a} \end{aligned}$$

Esto se verifica para todo $t_0 \in (a, b)$ donde (a, b) es uno de los subconjuntos que quitamos en la construcción de \mathcal{C} , luego f es diferenciable en \mathcal{C}^c , lo que implica que f es diferenciable en casi todo punto de \mathcal{I} .

■

2.4. Aproximación por polígonos de la curva de Lebesgue

Para llegar al concepto de poligonal aproximante de la curva de Lebesgue, primero vamos a ver su construcción de forma geométrica.

esto es debido a que

$$\frac{2}{3^{2n+1}} + \frac{2}{3^{2n+2}} + \cdots = \sum_{k=1}^{\infty} \frac{2}{3^{2n+k}} = \frac{2}{3^{2n}} \sum_{k=1}^{\infty} \left(\frac{1}{3}\right)^k.$$

Por tanto la imagen de t_b viene dada por

$$f(t_b) = \begin{pmatrix} 0_2 t_1 t_3 \dots t_{2n-1} \bar{1} \\ 0_2 t_2 t_4 \dots t_{2n} \bar{1} \end{pmatrix}.$$

Tenemos que

$$\begin{aligned} 0_2 t_1 t_3 \dots t_{2n-1} \bar{1} &= \frac{t_1}{2} + \frac{t_3}{2^2} + \cdots + \frac{t_{2n-1}}{2^n} + \sum_{k=1}^{\infty} \frac{1}{2^{n+k}} \\ &= 0_2 t_1 t_3 \dots t_{2n-1} + \frac{1}{2^n} \left(\frac{1/2}{1 - 1/2} \right) = 0_2 t_1 t_3 \dots t_{2n-1} + \frac{1}{2^n} \end{aligned}$$

y

$$\begin{aligned} 0_2 t_2 t_4 \dots t_{2n} \bar{1} &= \frac{t_2}{2} + \frac{t_4}{2^2} + \cdots + \frac{t_{2n}}{2^n} + \sum_{k=1}^{\infty} \frac{1}{2^{n+k}} \\ &= 0_2 t_2 t_4 \dots t_{2n} + \frac{1}{2^n} \left(\frac{1/2}{1 - 1/2} \right) = 0_2 t_2 t_4 \dots t_{2n} + \frac{1}{2^n} \end{aligned}$$

es decir, podemos ver que las coordenadas de $f(a)$ y $f(b)$ distan $1/2^n$, y por consiguiente como $f(a)$ y $f(b)$ pertenecen al subcuadrado $\mathcal{Q}_{t_1 t_2 t_3 \dots t_{2n}}$ cuya longitud lateral es de $1/2^n$, tenemos que $f(b)$ está en la diagonal con respecto $f(a)$ de dicho cuadrado. Con todo esto llegamos a que la curva de Lebesgue entra en el subcuadrado $\mathcal{Q}_{t_1 t_2 t_3 \dots t_{2n}}$ por $f(a)$ que es el vértice inferior izquierdo y sale de dicho cuadrado por $f(b)$ que sería el vértice superior derecho. Ahora unimos el punto de salida con el punto de entrada del siguiente subcuadrado mediante una línea recta, que es precisamente la interpolación lineal

$$f(t) = \frac{1}{b-a} [f(a)(b-t) + f(b)(t-a)]$$

Hemos representado los dos primeros polígonos de este tipo en la Fig. 2.4, donde hemos aproximado algunos vértices de dichos polígonos para poderlo entender mejor.

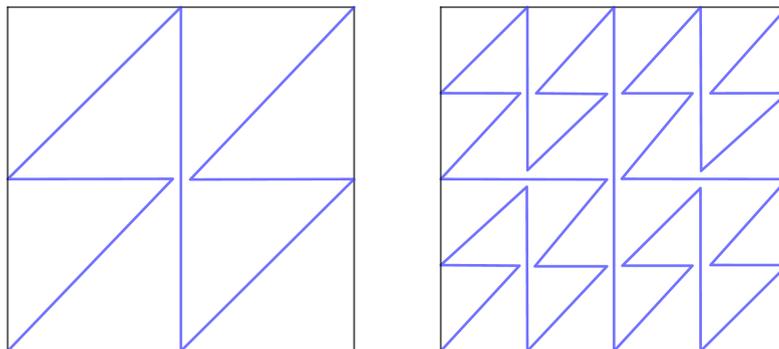


Figura 2.4. Primera y segunda poligonal aproximante de la curva de Lebesgue

Con esto damos por terminado el estudio de la curva de Lebesgue que, como ya hemos mencionado, nos sirve de guía para probar el teorema de Hahn-Mazurkiewicz, el cual nos decía que todo conjunto es la imagen de un intervalo mediante una aplicación continua si y solo si dicho conjunto es compacto, conexo y localmente conexo. Para llegar a dicho teorema necesitamos ver primero unas nociones de Topología.

2.5. Preliminares de Topología

Definición 2.2. Sea $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. Para cualquier $A \subset \mathbb{R}^n$ el conjunto imagen por f de A es el conjunto

$$f(A) = \{f(x) : x \in A \cap D\} \subset \mathbb{R}^m.$$

Para cada $B \subset \mathbb{R}^m$, la imagen inversa de B mediante f es el conjunto

$$f^{-1}(B) = \{x \in D : f(x) \in B\} \subset D \cap \mathbb{R}^n$$

Definición 2.3. f es continua en $a \in D$ si para todo $\varepsilon > 0$ existe $\delta > 0$ de modo que $f(B(a, \delta) \cap D) \subset B(f(a), \varepsilon)$.

Teorema 2.3. Sea $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. Son equivalentes:

1. f es continua en D .
2. Para todo abierto $G \subset \mathbb{R}^m$ existe un abierto $G_1 \subset \mathbb{R}^n$ de modo que $f^{-1}(G) = G_1 \cap D$.

| Teorema 2.4. Sea $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. Son equivalentes:

1. f es continua en D .
2. Para todo cerrado $F \subset \mathbb{R}^m$ existe un cerrado $F_1 \subset \mathbb{R}^n$ de modo que $f^{-1}(F) = F_1 \cap D$.

| Definición 2.4. Diremos que $K \subset \mathbb{R}^n$ es compacto si de todo recubrimiento por abiertos de K se puede extraer un subrecubrimiento finito.

| Teorema 2.5. Si $f : K \rightarrow \mathbb{R}^m$ es continua y K es compacto entonces $f(K)$ es un compacto.

Demostración. Sea $\{G_\gamma\}_{\gamma \in \Gamma}$ un recubrimiento por abiertos de $f(K)$. Entonces existen abiertos $U_\gamma \subset \mathbb{R}^n$ tales que $f^{-1}(G_\gamma) = U_\gamma \cap K$. La familia de abiertos $\{U_\gamma\}_{\gamma \in \Gamma}$ es un recubrimiento por abiertos de K claramente y por ser K compacto existe un subrecubrimiento finito $\{U_{\gamma_j}\}_{j=1}^p$ de K . Entonces $\{G_{\gamma_j}\}_{j=1}^p$ es un subrecubrimiento finito de $f(K)$. |

| Definición 2.5. Sean $A, B \subset \mathbb{R}^n$. Diremos que A es denso en B si $\bar{A} \supset B$.

| Definición 2.6. Diremos que $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ es uniformemente continua en D si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\|f(x) - f(y)\| < \varepsilon$$

siempre que $x, y \in D$ y $\|x - y\| < \delta$.

| Teorema 2.6. Supongamos que D es un conjunto denso en I , M es compacto de \mathbb{R}^n y $g : D \rightarrow M$ es uniformemente continua. Entonces existe una única $f : I \rightarrow M$ continua de modo que $f(t) = g(t)$ para todo $t \in D$ (existe una única extensión continua de f de D a todo I).

Demostración. Sea $t \in I \setminus D$. Como t es de acumulación de D existe una sucesión $\{t_n\} \subset D$ de modo que $t_n \rightarrow t$. Como $\{t_n\}$ es una sucesión convergente es de Cauchy. Entonces la sucesión $\{g(t_n)\}$ también es de Cauchy gracias a la continuidad uniforme de g . Por lo tanto la sucesión $\{g(t_n)\}$ es convergente. Definimos entonces

$$f(t) := \lim_{n \rightarrow +\infty} g(t_n)$$

f está bien definida en el conjunto $\mathcal{I} \setminus D$ porque el límite anterior no depende de la sucesión $\{t_n\}$ elegida, cualquier sucesión contenida en D que tienda al punto t genera el mismo valor $f(t)$ gracias a la continuidad uniforme de g . Por supuesto ponemos $f(t) = g(t)$ para $t \in D$. Ya tenemos f bien definida para todo $t \in \mathcal{I}$. Veamos ahora que f es uniformemente continua sobre \mathcal{I} . Sean $t', t'' \in \mathcal{I}$, asumimos sin pérdida de generalidad que $t' < t''$ de modo que $|t' - t''| < \delta$. Elegimos las sucesiones $\{t'_n\}$ y $\{t''_n\}$ con $t'_n, t''_n \in D$ tal que la primera converge de forma decreciente a t' y la segunda converge de forma creciente a t'' . Entonces

$$\|f(t') - f(t'')\| \leq \|f(t') - f(t'_n)\| + \|f(t'_n - f(t''_n))\| + \|f(t''_n) - f(t'')\|.$$

Por construcción $|t'_n - t''_n| < \delta$, luego por la continuidad uniforme de f en D existe un $\varepsilon > 0$ tal que $\|f(t'_n) - f(t''_n)\| < \varepsilon/3$. Como $\{t'_n\} \rightarrow f(t')$ y $\{t''_n\} \rightarrow f(t'')$ podemos elegir n, m tal que $\|f(t') - f(t'_n)\| < \varepsilon/3$ y $\|f(t''_m) - f(t'')\| < \varepsilon/3$. Por tanto

$$\|f(t') - f(t'')\| < \varepsilon \quad \forall t', t'' \in \mathcal{I} \text{ tal que } |t' - t''| < \delta$$

es decir, f es uniformemente continua en \mathcal{I} .

Además, como M es compacto, $f(t) \in M$ para todo $t \in \mathcal{I}$. Falta ver que la función f es única. Supongamos que no es única, es decir que existe una función $f' : \mathcal{I} \rightarrow M$ tal que $f' = g(t) \forall t \in D$. Si definimos $h(t) = f(t) - f'(t) \forall t \in \mathcal{I}$, tenemos que $h(t) = 0 \forall t \in D$. Sea $t_0 \in \mathcal{I}$ y sea $\{t_n\}$ una sucesión que converge a t_0 , con $t_n \in D$. Entonces

$$h(t_0) = \lim_{n \rightarrow \infty} h(t_n) = 0$$

Por tanto $h(t) = 0$ para todo $t \in \mathcal{I}$, es decir, $f = f'$. Luego f es única. |

| Definición 2.7. $A \subset \mathbb{R}^n$ es conexo si no existen dos abiertos G_1, G_2 de \mathbb{R}^n tales que $A \cap G_1 \neq \emptyset, A \cap G_2 \neq \emptyset, (A \cap G_1) \cap (A \cap G_2) = \emptyset$ y $(A \cap G_1) \cup (A \cap G_2) = A$

| Definición 2.8. Si existen los abiertos G_1, G_2 diremos que A no es conexo y que G_1 y G_2 son una desconexión de A .

| Teorema 2.7. Si $D \subset \mathbb{R}^n$ es conexo y $f : D \rightarrow \mathbb{R}^m$ es continuo entonces $f(D)$ es conexo.

| Teorema 2.8. *Supongamos que $C \subset \mathbb{R}^n$ es un conjunto conexo y cerrado. Entonces, dados dos puntos cualesquiera $x, x' \in C$ y dado cualquier $\varepsilon > 0$, existe una colección finita de puntos $x = x_1, x_2, x_3, \dots, x_p = x'$ pertenecientes a C de modo que $\|x_j - x_{j-1}\| < \varepsilon$ para $j = 1, \dots, p$ (existe una ε -cadena finita de x a x').*

Demostración. Sea $x' \in A$, arbitrario pero fijo. Sea $\varepsilon > 0$. Sea

$$A = \{x \in C : \text{existe una } \varepsilon\text{-cadena finita de } x' \text{ a } x\}$$

Queremos demostrar que $A = C$.

A no es el conjunto vacío porque $x' \in A$ de forma trivial.

Sea $\Omega = \bigcup_{a \in A} B(a, \varepsilon)$. Ω es un abierto no vacío. Si $x \in C \cap \Omega$ entonces $x \in C \cap B(a, \varepsilon)$ para algún $a \in A$ y por ello hay una ε -cadena finita de x' a x a través de a luego $x \in A$. Hemos demostrado que $C \cap \Omega \subset A$. Ahora al revés. Sea $x \in A$. Entonces $x \in B(x, \varepsilon) \subset \Omega$ y como $A \subset C$ tenemos que $A \subset C \cap \Omega$. Hemos demostrado que $A = C \cap \Omega$.

Claramente $C \cap \Omega \neq \emptyset$ pues $x' \in A$. Sea x un punto de acumulación de A . Dado que C es cerrado, $x \in C$. Como $B(x, \varepsilon)^*$ debe contener un punto de A , digamos a , y como a está a distancia de $x \in C$ menor que ε es claro que existe una ε -cadena finita de x' a x vía el punto a . Así $x \in A$, por ello A es un cerrado y A^c es un abierto. Como $A = C \cap \Omega$ tenemos que $C \cap \Omega \cap A^c = \emptyset$.

Si $A^c \cap C$ fuese no vacío entonces Ω y A^c sería una desconexión de C ya que

$$A^c \cap C \neq \emptyset, \Omega \cap C \neq \emptyset$$

$$(A^c \cap C) \cup (\Omega \cap C) = C, (A^c \cap C) \cap (\Omega \cap C) = \emptyset.$$

Luego debe ocurrir que $A^c \cap C = \emptyset$. C no tiene puntos que no estén en A . Como $A \subset C$ concluimos que $A = C$ y esto termina el teorema.

2.6. Teorema de Netto

En este apartado vamos a demostrar que una función continua definida entre \mathcal{I} y \mathcal{Q} no puede ser biyectiva, lo que implicaría que no hay curvas inyectivas que rellenan \mathcal{Q} .

| Teorema 2.9. *Si $f : \mathcal{I} \rightarrow \mathcal{Q}$ es biyectiva, entonces f es necesariamente discontinua en \mathcal{I} .*

Demostración. Veámoslo por reducción a lo absurdo, es decir, supongamos que f es continua en \mathcal{I} .

Sea $g = f^{-1} : \mathcal{Q} \rightarrow \mathcal{I}$ la función inversa y sea $A \subset \mathbb{R}$ un conjunto cerrado. Entonces $A \cap \mathcal{I}$ es un compacto, por ser cerrado y acotado. Luego por el teorema 2.5 tenemos que $f(A \cap \mathcal{I})$ es un compacto, por tanto es un cerrado. Definimos $A_1 = f(A \cap \mathcal{I})$. Podemos observar que $g^{-1}(A) = A_1$ y $A_1 = A_1 \cap \mathcal{Q}$ luego por el teorema 2.6 tenemos que $g = f^{-1}$ es continua.

Si quitamos un punto $0 < t_0 < 1$ de \mathcal{I} y su imagen $f(t_0) \in \mathcal{Q}$, resulta que $\mathcal{Q} \setminus \{f(t_0)\}$ es conexo pero $\mathcal{I} \setminus \{t_0\}$ no lo es. Luego hemos llegado a una contradicción pues f^{-1} es continua y por ello debería transformar conexos en conexos. Por lo tanto f no puede ser continua en \mathcal{I} . |

2.7. Conjuntos localmente conexos

| Definición 2.9. *Un conjunto $A \subset \mathbb{R}^n$ es localmente conexo si para cada punto $a \in A$ y para cada $\varepsilon > 0$ existe $\delta = \delta(a, \varepsilon)$ de modo que para todo $a' \in B(a, \delta) \cap A$ existe un conjunto compacto y conexo $A' \subset B(a, \varepsilon) \cap A$ tal que $a, a' \in A'$.*

Observación 2.1. 1. Localmente conexo no implica conexo (ver Fig. 2.5).

2. Conexo no implica localmente conexo. Veamos un ejemplo: Consideremos los subconjuntos de \mathbb{R}^2 de la forma

$$X_n = \left\{ \left(x, \frac{x}{2^n} \right) : x \in [0, 2] \right\}$$

entonces, el conjunto $X = \bigcup_{n=0}^{\infty} X_n$ es conexo, pero como podemos ver en la Fig. 2.6 no es localmente conexo.

3. Localmente conexo no es invariante por transformaciones continuas, pero si además de localmente conexo se añade que es compacto entonces si lo es.

| Teorema 2.10. *Sea $f : \mathcal{I} \rightarrow \mathbb{R}^n$ una aplicación continua. Entonces $f(\mathcal{I})$ es localmente conexo.*

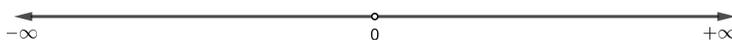


Figura 2.5. $\mathbb{R} \setminus \{0\}$ es localmente conexo pero no es conexo

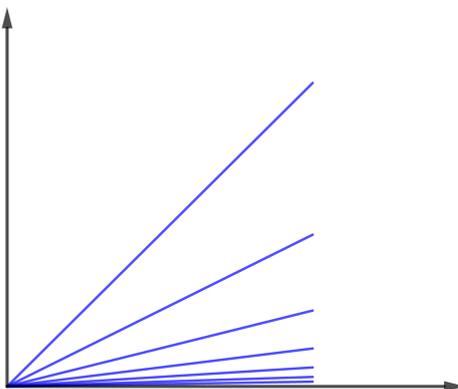


Figura 2.6. El conjunto X es conexo pero no es localmente conexo

Demostración. Probemoslo por reducción a lo absurdo, es decir, supongamos que $f(\mathcal{I}) = M$ no es localmente conexo. Entonces existe un punto $p \in M$ y un $\varepsilon > 0$ tal que para todo $\delta > 0$ existe un punto $p' \in B(p, \delta) \cap M$ de modo que todo compacto y conexo de M conteniendo p y p' tiene puntos fuera de $B(p, \varepsilon)$.

Consideremos (δ_j) una sucesión de números positivos decreciente tales que $\delta_j \rightarrow 0$. Para cada uno de los δ_j existe un punto $p'_j \in B(p, \delta_j) \cap M$ que no tiene la propiedad requerida. Por la construcción es claro que $p'_j \rightarrow p$.

Como $M = f(\mathcal{I})$, para cada j existe $t'_j \in \mathcal{I}$ de modo que $f(t'_j) = p'_j$. Como \mathcal{I} es compacto, existe una subsucesión convergente (t_{n_j}) digamos a $t' \in \mathcal{I}$. Por continuidad $f(t') = p$. El segmento que forman los puntos t_{n_j} y t' , que es o bien el segmento $[t'_{n_j}, t']$ o el segmento $[t', t_{n_j}]$, se transforma en un conjunto compacto y conexo $M' \subset M$ que contiene a p y a p'_{n_j} . Dado que f es continua, existe un $\delta > 0$ tal que

$$\|f(t) - f(t')\| < \varepsilon$$

para todo $t \in B(t', \delta) \cap \mathcal{I}$. Dado que (t_{n_j}) converge a t' , debe existir un j_0 de modo que $t'_{n_j} \in B(t', \delta) \cap \mathcal{I}$ para todo $j > j_0$. Por lo tanto la imagen del segmento que une

t'_{n_j} con t' está contenida en $B(p, \varepsilon)$ siempre que $j > j_0$. Esto es, existe un compacto y conexo $M' \subset M$ que pertenece a $B(p, \varepsilon)$ que contiene a p y a p'_{n_j} . Hemos llegado a una contradicción con la elección de los p_{n_j} , ya que de la hipótesis teníamos que todo compacto y conexo de M que contiene a p , p' tiene puntos fuera de $B(p, \varepsilon)$. Por lo que $f(\mathcal{I})$ es localmente conexo. |

2.8. Teorema de Hausdorff

| Teorema 2.11. *Todo conjunto compacto es la imagen continua del conjunto de Cantor.*

Demostración. Sea K un compacto. Es claro que

$$K \subset \bigcup_{x \in K} B(x, 1)$$

es un recubrimiento por abiertos de K , por lo que debe existir un subrecubrimiento finito, digamos

$$K \subset \bigcup_{i_1} B(x_{i_1}, 1)$$

donde $i_1 \in \{0, 1, 2, 3, 2^{n_1} - 1\}$ y n_1 es un número natural. (Siempre podemos conseguir que el número de elementos del recubrimiento sea una potencia de 2 repitiendo si es necesario uno de los entornos las veces que hagan falta).

Pongamos entonces para cada índice i_1 perteneciente al conjunto $\{0, 1, 2, 3, 2^{n_1} - 1\}$:

$$K_{i_1} = \overline{B}(x_{i_1}, 1) \cap K.$$

Es inmediato comprobar que $K = \bigcup_{i_1} K_{i_1}$. Cada K_{i_1} es un compacto contenido en K . Para cada uno de ellos podemos hacer una construcción similar al anterior pero esta vez con bolas de radio $1/2$. Podemos suponer que el número de elementos de los subrecubrimientos sea el mismo para cada i_1 añadiendo si hace falta entornos repetidos. Así existe un número natural n_2 y unos puntos $x_{i_1 i_2} \in K_{i_1}$ de modo que

$$K_{i_1} \subset \bigcup_{i_2} B(x_{i_1 i_2}, 1/2)$$

donde $i_2 \in \{0, 1, 2, \dots, 2^{n_2} - 1\}$.

Ponemos entonces para cada índice i_2 perteneciente al conjunto $\{0, 1, 2, 3, 2^{n_2} - 1\}$:

$$K_{i_1 i_2} = \overline{B}(x_{i_1 i_2}, 1/2) \cap K$$

Obviamente $K_{i_1} = \bigcup_{i_2} K_{i_1 i_2}$.

Continuamos el procedimiento con bolas de radio $1/2^2, 1/2^3, \dots$ obteniendo así una sucesión de números reales $n_1, n_2, \dots, n_k, \dots$ y una colección de compactos $K_{i_1 i_2 \dots i_k}$ donde $i_1 \in \{0, 1, 2, 3, \dots, 2^{n_1} - 1\}, i_2 \in \{0, 1, 2, 3, \dots, 2^{n_2} - 1\}, \dots, i_k \in \{0, 1, 2, 3, \dots, 2^{n_k} - 1\}, \dots$ y

$$K \supset K_{i_1} \supset K_{i_1 i_2} \supset K_{i_1 i_2 i_3} \supset \dots$$

y cada compacto $K_{i_1 i_2 \dots i_k}$ está contenido en una bola cerrada de radio $1/2^{k-1}$, luego cada compacto $K_{i_1 i_2 \dots i_k}$ tiene diámetro inferior a $1/2^{k-2}$.

Relacionamos ahora esta colección de compactos con el conjunto de cantor \mathcal{C} . Para ello cada elemento $t \in \mathcal{C}$ lo escribimos usando el desarrollo en base 3 del siguiente modo

$$t = 0_3(2b_1)(2b_2) \dots (2b_{n_1})(2b_{n_1+1}) \dots (2b_{n_1+n_2})(2b_{n_1+n_2+1}) \dots$$

donde cada b_j es o bien 0 o bien 1 y donde pueden existir infinitos $b_j = 0$.

Definimos la aplicación $f : \mathcal{C} \rightarrow K$ del siguiente modo

$$f(t) = K_{i_1} \cap K_{i_1 i_2} \cap K_{i_1 i_2 i_3} \cap \dots$$

donde los índices vienen dados por las siguientes relaciones

$$i_1 = (b_1 \dots b_{n_1})_2$$

$$i_2 = (b_{n_1+1} \dots b_{n_1+n_2})_2$$

$$i_3 = (b_{n_1+n_2+1} \dots b_{n_1+n_2+n_3})_2$$

y en general

$$i_k = (b_{n_1+\dots+n_{k-1}+1} \dots b_{n_1+n_2+\dots+n_k})_2$$

donde las expresiones de la derecha son representación binaria de los números de la izquierda.

Por el teorema de intersección de Cantor $f(t)$ es un único punto perteneciente al conjunto K . Tenemos una función bien definida.

En primer lugar, la aplicación f es sobreyectiva puesto que a cada punto de K le corresponde por lo menos una secuencia de compactos encajados $K_{i_1} \supset K_{i_1 i_2} \supset K_{i_1 i_2 i_3} \supset \dots$. Asociado a los índices obtenidos i_1, i_2, i_3, \dots tenemos una sucesión de desarrollos binarios que finalmente se corresponde con el desarrollo en base 3 de un cierto punto de \mathcal{C} .

Además vamos a demostrar que la función f es continua. Supongamos que t', t'' son dos puntos de \mathcal{C} de modo que $|t' - t''| < 1/3^{n+1}$. Entonces, como vimos en la demostración del teorema 2.1, en sus desarrollos en base 3 deben coincidir todos sus dígitos al menos hasta la posición n , de modo que

$$t' = 0_3(2b_1)(2b_2) \dots (2b_n)(2b_{n+1})(2b_{n+2}) \dots$$

$$t'' = 0_3(2b_1)(2b_2) \dots (2b_n)(2\beta_{n+1})(2\beta_{n+2}) \dots$$

Supongamos que

$$n_1 + n_2 + \dots + n_j \leq n < n_1 + n_2 + \dots + n_j + n_{j+1}$$

Entonces por la construcción de f se tiene que $f(t'), f(t'') \in K_{i_1 i_2 \dots i_j}$ y por lo tanto $\|f(t') - f(t'')\| \leq 1/2^{j-2}$ pues pertenecen a una bola cerrada de radio $1/2^{j-1}$. Con esto está claro que para todo $\varepsilon > 0$ podemos encontrar $\delta > 0$ tal que $\|f(t') - f(t'')\| < \varepsilon$ siempre que $|t' - t''| < \delta$ y t', t'' sean dos puntos de \mathcal{C} . Así f es continua (de hecho, es uniformemente continua en \mathcal{C} por ser el conjunto de Cantor un compacto.).

■

Hemos visto que todo compacto es la imagen continua del conjunto de Cantor. Si quisiéramos ahora seguir la idea de Lebesgue para construir una curva continua definida de \mathcal{I} en K nos haría falta asegurar que la extensión de \mathcal{C} a todo \mathcal{I} existe. Si K fuese convexo estaría hecho con la misma idea de Lebesgue, pero no todo compacto es convexo. Vamos a estudiar en las secciones restantes las condiciones que debe cumplir K para que dicha extensión exista.

2.9. Conexo por caminos

| Definición 2.10. *Un conjunto A es conexo por caminos si dados dos puntos cualesquiera $p', p'' \in A$ existe una función continua $f : I \rightarrow A$ tal que $f(0) = p'$ y $f(1) = p''$.*

| Teorema 2.12. *Supongamos que M es un conjunto compacto, conexo y localmente conexo. Entonces para todo $\varepsilon > 0$ existe $\delta = \delta(\varepsilon) > 0$ de modo que siempre que p', p'' sean dos puntos de M tales que $\|p' - p''\| < \delta$, existe un conjunto compacto y conexo $M' \subset M$ tal que contiene a los puntos p' y p'' y además $M' \subset B(p', \varepsilon) \cap B(p'', \varepsilon)$, esto es, p', p'' están conectados por un conjunto compacto y conexo muy próximo a p' y p'' .*

Demostración. Vamos a demostrarlo por reducción a lo absurdo, es decir, supongamos que Existe un $\varepsilon > 0$ tal que $\forall \delta_i > 0, i = 1, 2, 3, \dots$, con $\lim_{i \rightarrow \infty} \delta_i = 0$, existe un par de puntos p'_i, p''_i que verifican $\|p'_i - p''_i\| < \delta_i$ y además cualquier compacto conexo M'_i que contenga a p'_i, p''_i tiene puntos que están fuera de $B(p', \varepsilon) \cap B(p'', \varepsilon)$.

Como M es compacto, existe una subsucesión de $\{p'_i, p''_i\}, \{p'_{i_j}, p''_{i_j}\}$ de modo que los $\{p'_{i_j}\}$ convergen a algún $p' \in M$. También existe una subsucesión $\{p'_k, p''_k\}$ de la subsucesión anterior $\{p'_{i_j}, p''_{i_j}\}$, de manera que los $\{p'_k\}$ siguen convergiendo a p' y los $\{p''_k\}$ convergen a un $p'' \in M$. Tenemos por construcción que $\lim_{k \rightarrow \infty} p'_k = p'$, y $\lim_{k \rightarrow \infty} p''_k = p''$. Como $\|p'_k - p''_k\| < \delta_k$ y $\delta_k \rightarrow 0$, entonces $p' = p''$. Por simplificar, denominamos p a p' y p'' . Como M es cerrado, entonces $p \in M$. Como M es localmente conexo, para cualquier $\varepsilon/2 > 0$ existe un $\delta > 0$ tal que, para todo $\bar{p} \in B(p, \delta)$, p y \bar{p} están conectados por un subconjunto compacto $K \subseteq M$ que además $K \subset B(p, \varepsilon/2)$. Elegimos un k_0 tal que $p' \in B(p, \delta)$ y $p'' \in B(p, \delta) \forall k > k_0$. Luego p'_k y p están conectados por un subconjunto compacto $M' \subseteq M$ y p''_k y p están conectados por un subconjunto compacto $M'' \subseteq M$. Ambos subconjuntos pertenecen a $B(p, \varepsilon/2)$. $M \cup M'$ es compacto y como ambos contienen a p también tenemos que es conexo. Finalmente, por construcción tenemos que $M' \cup M'' \subset B(p'_k, \varepsilon) \cap B(p''_k, \varepsilon)$, lo que contradice nuestra hipótesis. **|**

La existencia de δ solo depende de ε , independientemente de los puntos, por eso diremos que un compacto, conexo y localmente conexo es uniformemente localmente

conexo.

| Teorema 2.13. *Un compacto, conexo y localmente conexo es conexo por caminos.*

Demostración. Sea M compacto, conexo y localmente conexo. Sabemos que M es uniformemente localmente conexo. Sean p', p'' dos puntos en M . Queremos demostrar que existe un camino continuo que los une.

Sea (ε_k) una sucesión de números positivos tales que $\sum_{k=1}^{\infty} \varepsilon_k < +\infty$. Por el teorema 2.12, dado $\varepsilon_1 > 0$ existe $\delta_1 > 0$ y una δ_1 -cadena que une p' con p'' . Podemos suponer que existen $2^{n_1} + 1$ puntos en la cadena, contando punto de inicio y de fin (si no es así los añadimos contando alguno varias veces). Denotemos la cadena así

$$p' = p_0^1, p_1^1, \dots, p_{2^{n_1}}^1 = p''.$$

Dos puntos consecutivos de la δ_1 -cadena pueden unirse mediante un conjunto compacto y conexo $M_i^1 \subset M$ de modo que $M_i^1 \subset B(p_i^1, \varepsilon_1) \cap B(p_{i+1}^1, \varepsilon_1)$. Esto termina el primer paso del proceso.

Ahora, para cada índice $i = 0, \dots, 2^{n_1} - 1$, podemos encontrar una δ_2 -cadena que une los puntos p_i^1 y p_{i+1}^1 contenida en el compacto y conexo M_i^1 . Podemos conseguir que todas las cadenas, para $i = 0, \dots, 2^{n_1} - 1$ tengan exactamente la misma longitud igual a $2^{n_2} + 1$ puntos siendo n_2 un cierto número natural. Denotemos las nuevas cadenas así

$$p_i^1 = p_{i2^{n_2}}^2, p_{i2^{n_2}+1}^2, \dots, p_{(i+1)2^{n_2}}^2 = p_{i+1}^1$$

para cada $i = 0, 1, 2, \dots, 2^{n_1} - 1$.

Dados dos puntos consecutivos de la colección de puntos anteriores es claro que existe un compacto y conexo $M_j^2 \subset M_k^1$ que pertenece a la intersección de las bolas de radio ε_2 centradas en los dos puntos considerados.

El proceso se puede repetir y podemos construir ahora δ_3 -cadenas entre puntos adyacentes de la segunda etapa, contenidas estas cadenas en los conjuntos M_j^2 y cada una de ellas con $2^{n_3} + 1$ puntos cada una.

Después de k pasos llegaremos a una cadena

$$p' = p_0^k, p_1^k, \dots, p_{2^{n_1+n_2+\dots+n_k}}^k = p''.$$

donde cada dos puntos consecutivos de la cadena son conectados por un compacto conexo $M_r^k \subset M_j^{k-1}$ que pertenece a la intersección de las bolas de radio ε_k centradas en los dos puntos considerados.

Comencemos definiendo una función g sobre aquellos $t \in \mathcal{I}$ que se puedan escribir en la forma

$$t_{i,k} = \frac{i}{2^{n_1+n_2+\dots+n_k}}$$

mediante

$$g(t_{i,k}) = p_i^k.$$

para todo $i = 0, 1, 2, \dots, 2^{n_1+n_2+\dots+n_k}$.

Por construcción cada cadena contiene a todas las cadenas precedentes y

$$p_i^{k-1} = p_{i2^{n_k}}^k$$

Vamos a demostrar que esta función es uniformemente continua en su dominio de definición

$$D = \{0_2 b_1 \dots b_s : b_i = 0 \text{ o } 1, s = 1, 2, 3, \dots\}$$

y por ello podremos extender g a todo $\mathcal{I} = \overline{D}$. Tengamos en cuenta que

$$0_2 b_1 b_2 \dots b_{n_1+n_2+\dots+n_k} = \frac{i}{2^{n_1+n_2+\dots+n_k}}$$

si $i = b_{n_1+n_2+\dots+n_k} + 2b_{n_1+n_2+\dots+n_k-1} + 4b_{n_1+n_2+\dots+n_k-2} + \dots + 2^{n_1+n_2+\dots+n_k-1}b_1$.

Sea $\varepsilon > 0$, entonces existe un k_0 lo suficientemente grande tal que

$$\sum_{k=k_0}^{\infty} \varepsilon_k < \frac{\varepsilon}{2}.$$

Entonces $|t' - t''| < 1/2^{n_1+n_2+\dots+n_{k_0}} \triangleq \delta$ y $t', t'' \in D$ implica que $\|g(t') - g(t'')\| < \varepsilon$.

Y como δ es independiente de t , la continuidad es uniforme. Vamos a demostrarlo. Si $|t' - t''| < \delta$, entonces t' y t'' se encontrarían, en el peor de los casos, en dos intervalos adyacentes

$$\left[\frac{j-1}{2^{n_1+n_2+\dots+n_{k_0}}}, \frac{j}{2^{n_1+n_2+\dots+n_{k_0}}} \right], \quad \left[\frac{j}{2^{n_1+n_2+\dots+n_{k_0}}}, \frac{j+1}{2^{n_1+n_2+\dots+n_{k_0}}} \right].$$

Podemos asumir sin pérdida de generalidad que t' se encuentra en el primer intervalo y t'' en el segundo. Entonces para algún r y j^* tendríamos

$$g(t') = p_{j^*}^{k_0+r}$$

Este punto se encuentra entre dos puntos $p_{j_0}^{k_0}$ y $p_{j_0+1}^{k_0}$, donde $j_0 = j - 1$, que por construcción están conectados por un conjunto compacto $M_{j_0+1}^{k_0} \subseteq N_{\varepsilon_{k_0}}(p_{j_0}^{k_0})$. A su vez, se encontrará entre los dos puntos

$$p_{j_1}^{k_0+1}, p_{j_1+1}^{k_0+1} \in M_{j_0+1}^{k_0}$$

que están conectados por el conjunto compacto $M_{j_1+1}^{k_0+1} \subseteq N_{\varepsilon_{k_0+1}}(p_{j_1}^{k_0+1}), \dots$

Tenemos que

$$\|p_{j_1}^{k_0+1} - p_{j_0}^{k_0}\| < \varepsilon_{k_0} \text{ y } \|p_{j_2}^{k_0+2} - p_{j_1}^{k_0+1}\| < \varepsilon_{k_0+1}$$

y por tanto,

$$\|p_{j_2}^{k_0+2} - p_{j_0}^{k_0}\| < \varepsilon_{k_0} + \varepsilon_{k_0+1}.$$

Continuando con este proceso llegamos finalmente a

$$p_{j^*}^{k_0+r} \in N_{\varepsilon_{k_0} + \dots + \varepsilon_{k_0+r-1}}(p_{j_0}^{k_0}),$$

es decir,

$$\|g(t') - g(t_{j_0, k_0})\| = \|p_{j^*}^{k_0+r} - p_{j_0}^{k_0}\| < \varepsilon_{k_0} + \dots + \varepsilon_{k_0+r-1} < \varepsilon/2$$

De forma análoga podemos probar

$$\|g(t'') - g(t_{j_0, k_0})\| < \varepsilon/2$$

y así llegamos al resultado que queríamos obtener $\|g(t') - g(t'')\| < \varepsilon$. Hemos llegado a que la función es uniformemente continua y por el teorema 2.6 del anexo podemos extender g a todo $\mathcal{I} = \bar{D}$. Por tanto M es conexo por caminos. |

| Teorema 2.14. *Sea M compacto, conexo y localmente conexo. Para todo $\varepsilon > 0$ existe $0 < \delta \leq \varepsilon$ tal que si $\|p' - p''\| < \delta$, $p', p'' \in M$, existe un camino continuo de p' a p'' contenido en $M \cap B(p', \varepsilon) \cap B(p'', \varepsilon)$.*

Demostración. Sea $\varepsilon > 0$ arbitrario pero fijo, y sea $0 < \lambda < \varepsilon/2$. Por el teorema 2.12, tenemos que existe un $\delta > 0$ tal que, si $\|p' - p''\| < \delta$ entonces hay un conjunto compacto y conexo $M' \subseteq M$ que contiene p', p'' y $M' \subseteq B(p', \lambda) \cap B(p'', \lambda)$. Sea

$\varepsilon_k = \lambda/2^k$ y escogemos la δ_1 -cadena de p' a p'' contenida en M' . Si denotamos la cadena

$$p' = p_0^1, p_1^1, \dots, p_{2^{n_1}}^1 = p''$$

entonces como vimos en la demostración del teorema anterior, dos puntos consecutivos de la cadena pueden unirse mediante un conjunto compacto y conexo $M_i^1 \subseteq B(p_i^1, \lambda/2)$. Como $p_i^1 \in B(p', \lambda) \cap B(p'', \lambda)$, entonces está claro que $M_i^1 \subset B(p', \lambda + \lambda/2) \cap B(p'', \lambda + \lambda/2)$. Continuando este procedimiento llegamos a que

$$M_i^k \subset B(p', \lambda + \lambda/2 + \dots + \lambda/2^k) \cap B(p'', \lambda + \lambda/2 + \dots + \lambda/2^k)$$

. Es claro que $\lambda + \lambda/2 + \dots = 2\lambda = \varepsilon$. Como la cadena que une p' con p'' contiene los puntos del conjunto M_i^k y los puntos de acumulación de dichos puntos están contenidos en $B(p', \varepsilon) \cap B(p'', \varepsilon)$, como queríamos probar. |

2.10. El teorema de Hahn-Mazurkiewicz

Ahora estamos listos para construir una aplicación continua y sobreyectiva de \mathcal{I} a cualquier conjunto $M \subset \mathbb{R}^n$ compacto, conexo y localmente conexo, al generalizar el proceso de extensión de $f : \mathcal{C} \rightarrow \mathcal{Q}$ a la curva que rellena el espacio de Lebesgue, usando el teorema de Hausdorff.

| **Teorema 2.15.** *Las siguientes condiciones son equivalentes:*

1. M es la imagen continua de \mathcal{I} .
2. M es compacto, conexo y localmente conexo.

Demostración. (1) \Rightarrow (2) Como \mathcal{I} es un compacto, por ser cerrado y acotado, y M es la imagen continua de \mathcal{I} , entonces por el teorema 2.5 tenemos que M es compacto. Por otro lado, el teorema 2.7 nos dice que como M es la imagen continua de \mathcal{I} e \mathcal{I} es un conexo, tenemos que M es un conexo. Por último nos faltaría ver que M es localmente conexo, pero también lo tenemos por el teorema 2.10.

(2) \Rightarrow (1) Supongamos que M es compacto, conexo y localmente conexo. Tenemos que probar que existe $f : \mathcal{I} \rightarrow M$ continua y sobreyectiva.

Por el teorema de Hausdorff sabemos que existe una función continua y sobreyectiva $f : \mathcal{C} \rightarrow M$ definida sobre el conjunto de Cantor \mathcal{C} .

También sabemos que \mathcal{C}^c , el complementario de \mathcal{C} en el intervalo $[0, 1]$, es una unión numerable de intervalos abiertos del tipo (a, b) . Entonces podemos extender f a una función sobre todo \mathcal{I} uniendo las imágenes $f(a), f(b)$ mediante un camino continuo contenido en M (puesto que todo compacto, conexo y localmente conexo es conexo por caminos). Estos caminos se construyen del siguiente modo.

Sea la sucesión $(\varepsilon_n) = (\frac{1}{n})$, en realidad nos vale cualquier sucesión que tienda a 0. Por el teorema 2.14, para cada $\varepsilon_n > 0 \exists \eta_n \leq \varepsilon_n$ de modo que si $p', p'' \in M$ y

$$\|p' - p''\| \leq \eta_n$$

entonces existe un camino continuo de p' a p'' contenido en $M \cap B(p', \varepsilon_n) \cap B(p'', \varepsilon_n)$. Podemos suponer siempre que (η_n) es una sucesión decreciente.

Para cada n , sea $m(n)$ el primer número natural que verifique

$$\frac{1}{3^{m(n)}} < \eta_n$$

entonces, todos los intervalos de la forma (a, b) que quitamos en la construcción del conjunto de Cantor en la etapa $m(n)$ y posteriores tienen longitud

$$b - a \leq \frac{1}{3^{m(n)}} \leq \eta_n.$$

Tenemos por ello una colección de etapas $m(1), m(2), m(3), \dots$. Entonces la extensión de $f : \mathcal{C} \rightarrow M$ a todo \mathcal{I} la hacemos de modo que todos los caminos construidos a partir de la etapa $m(n)$ cumplan la propiedad de que

$$\|f(t) - f(a)\| < \varepsilon \quad \forall t \in [a, b].$$

Entre las etapas $m(1)$ y $m(2) - 1$, los construimos de modo que $\|f(t) - f(a)\| < \varepsilon_1$ para todo $t \in [a, b]$, siendo (a, b) un intervalo de los que quitamos en estas etapas. Todos ellos cumplirán que

$$b - a < \frac{1}{3^{m(1)}} < \eta_1.$$

Entre las etapas $m(2)$ y $m(3) - 1$, construimos los intervalos de modo que $\|f(t) - f(a)\| < \varepsilon_2 \quad \forall t \in [a, b]$, siendo (a, b) uno de los intervalos que se quitan en

esas etapas. Todos ellos cumplirán que

$$b - a < \frac{1}{3^{m(2)}} < \eta_2.$$

Así tenemos claro como tenemos que ir definiendo f sobre \mathcal{C}^c , en \mathcal{I} .

Tomemos una extensión de f a todo el intervalo cerrado \mathcal{I} . Tenemos que demostrar que f es continua en \mathcal{I} . Por la definición en \mathcal{C}^c , está claro que ahí va a ser continua. Tenemos que demostrarlo para cualquier punto t_0 de \mathcal{C} .

Como vimos en el lema 2.2, t_0 puede ser punto de acumulación por la izquierda, por la derecha o por ambos lados. Vamos a estudiar el caso de punto de acumulación por la derecha, el resto de casos son similares. Que t_0 sea punto de acumulación por la derecha es equivalente a que sea el extremo derecho de algunos de los intervalos quitados en la construcción de \mathcal{C} . Por ello es claro que

$$\lim_{t \rightarrow t_0^-} f(t) = f(t_0).$$

Tenemos que probar que es continua por la derecha. Sea $\varepsilon > 0$ cualquiera, tomamos un ε_n tal que $2\varepsilon_n \leq \varepsilon$. Como f es uniformemente continua en \mathcal{C} , existe un cierto $\delta'_n > 0$ tal que

$$\delta'_n \leq \frac{1}{3^{m(n)}},$$

que nos asegura que

$$\|f(t_1) - f(t_2)\| < \varepsilon_n$$

siempre que $\|t_1 - t_2\| < \delta'_n$ y $t_1, t_2 \in \mathcal{C}$. Es seguro que hay algún punto de \mathcal{C} en $(t_0, t_0 + \delta'_n)$ pues t_0 es un punto de acumulación por la derecha. Tomamos un punto cualquiera t'_0 y tomamos δ_n como $|t'_0 - t_0| = \delta_n$. Con esta elección aseguramos que si tomamos un punto en $(t_0, t_0 + \delta_n)$ que no está en \mathcal{C} entonces está en un intervalo de los que se quitaron y ese intervalo queda seguro contenido en $(t_0, t_0 + \delta_n)$.

Sea $t \in (t_0, t_0 + \delta_n)$ cualquiera. Si $t \in \mathcal{C}$ entonces es claro que $|t - t_0| < \delta_n \leq \delta'_n$ y por ello $\|f(t) - f(t_0)\| < \varepsilon_n$.

Si $t \in \mathcal{C}^c$ entonces debe estar en uno de los intervalos quitados pero sabemos que ese intervalo (a, b) estaría contenido en $(t_0, t_0 + \delta_n)$ y por ello

$$|b - a| \leq \delta_n \leq \delta'_n \leq \frac{1}{3^{m(n)}} < \eta_n.$$

Por ello el camino construido sobre ese intervalo cumple que

$$\|f(t) - f(a)\| < \varepsilon_n \quad t \in (a, b)$$

Por lo tanto

$$\|f(t) - f(t_0)\| \leq \|f(t) - f(a)\| + \|f(a) - f(t_0)\| \leq 2\varepsilon_n \leq \varepsilon.$$

Luego hemos demostrado que f es continua por la derecha de t_0 . |

Comentario 2.1. En el libro [14] podemos encontrar la demostración de una generalización de este teorema: Un espacio de Hausdorff es la imagen continua de \mathcal{I} si y solo si dicho espacio es compacto, conexo y localmente conexo.

Capítulo 3

Curva de Jordan con medida de Lebesgue positiva

Sabemos por el Teorema de Netto que una curva inyectiva no puede tener contenido de Jordan positivo, pero vamos a ver en este capítulo que si es posible que una curva inyectiva tenga medida de Lebesgue positiva. Veamos primero la definición de curva de Jordan:

| Definición 3.1. Diremos que C es una curva de Jordan, si es la imagen de una aplicación $f : \mathcal{I} \rightarrow \mathbb{R}^n$ continua e inyectiva.

Como ya hemos dicho, en este capítulo veremos algunos ejemplos de curvas de Jordan cuya medida de Lebesgue es positiva. De hecho para cualquier número dado entre 0 y 1 daremos una curva de Jordan cuya medida bidimensional de Lebesgue será igual a dicho número.

3.1. Curva de Jordan de Osgood

La construcción que da Osgood para una familia de curvas de Jordan con medida de Lebesgue positiva, consiste en la eliminación sucesiva de regiones abiertas, comenzando en el cuadrado unidad \mathcal{Q} y procediendo como se indica en la Fig. 3.1, donde los cuadrados cerrados sombreados son lo que quedan después de cada iteración y como podemos ver los cuadrados están conectados por segmentos. Para

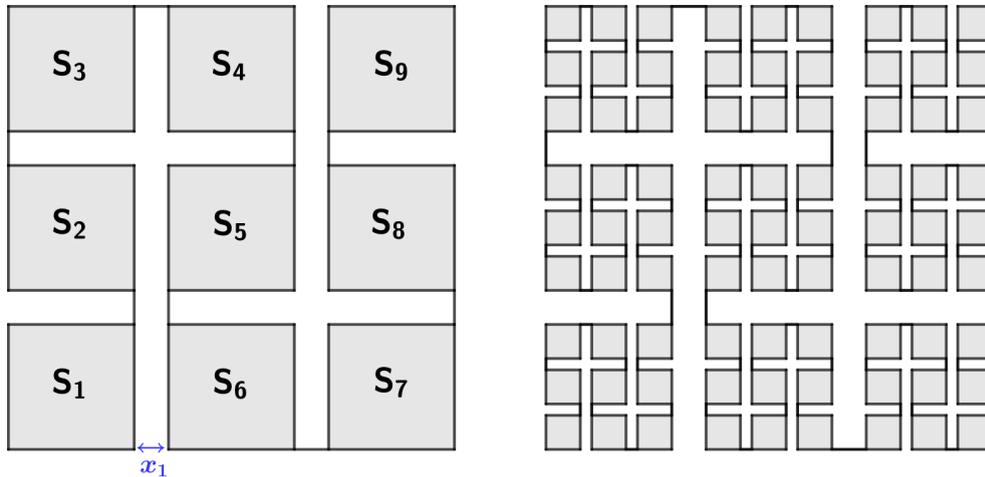


Figura 3.1. Construcción de la curva de Osgood

construir una curva con medida de Lebesgue λ , con $\lambda \in (0, 1)$, elegimos las regiones eliminadas tal que la suma de sus áreas sea $1 - \lambda$. Esto lo conseguimos de la siguiente manera: En el primer paso, elegimos x_1 tal que la distancia entre los cuadrados S_k y S_{k+1} , con $k = 1 \dots 8$ sea $x_1 = \frac{1-\sqrt{\lambda}}{4}$. En el siguiente paso, la elegimos tal que la distancia sea $x_2 = \frac{1-\sqrt{\lambda}}{24}$, en el siguiente paso $x_3 = \frac{1-\sqrt{\lambda}}{144}$, \dots . Entonces para el n -ésimo paso elegimos que la distancia entre los nuevos subcuadrados sea $x_n = \frac{1-\sqrt{\lambda}}{4 \cdot 6^{n-1}}$. Si denotamos por $m(A)$ a la medida de Lebesgue del conjunto A podemos ver que se verifica:

Teorema 3.1. *Sea A_n la región que contiene los 9^n cuadrados y los 9^{n-1} segmentos de línea que quedan después de la n -ésimo paso. Entonces*

$$m(A_n) = [\sqrt{\lambda} + (1 - \sqrt{\lambda})/2^n]^2$$

Demostración. Denotamos por l_n a la longitud lateral de los cuadrados en el n -ésimo paso, entonces $m(A_n) = l_n^2 \cdot 9^n$. Además también es fácil ver que (ver Fig. 3.1),

$$l_n = 3l_{n+1} + 2x_n$$

Vamos a probar el teorema por inducción:

Para $n = 1$, $l_1 = \frac{1-2x_1}{3}$ y

$$\begin{aligned} m(A_1) &= l_1^2 \cdot 9 = \left(\frac{1-2x_1}{3}\right)^2 9 \\ &= (1-2x_1)^2 = \left(1-2\frac{1-\sqrt{\lambda}}{4}\right)^2 \\ &= \left(\frac{1+\sqrt{\lambda}}{2}\right)^2 = \left(\sqrt{\lambda} + \frac{1-\sqrt{\lambda}}{2}\right)^2 \end{aligned}$$

vemos que efectivamente para $n=1$ se verifica el teorema. Supongamos que también es cierto para n y comprobemos que se verifica para $n + 1$. De la ecuación $l_n = 3l_{n+1} + 2x_n$ obtenemos que

$$l_{n+1} = \frac{l_n - 2x_{n+1}}{3}$$

, entonces

$$m(A_{n+1}) = l_{n+1}^2 \cdot 9^{n+1} = \left(\frac{l_n - 2x_{n+1}}{3}\right)^2 9^{n+1}$$

Por inducción tenemos que se verifica

$$m(A_n) = \left(\sqrt{\lambda} + \frac{1-\sqrt{\lambda}}{2^n}\right)^2 = l_n^2 \cdot 9^n$$

luego

$$l_n = \left(\sqrt{\lambda} + \frac{1-\sqrt{\lambda}}{2^n}\right) \frac{1}{3^n}$$

entonces nos queda

$$\begin{aligned} m(A_{n+1}) &= \left[\left(\sqrt{\lambda} + \frac{1-\sqrt{\lambda}}{2^n}\right) \frac{1}{3^n} - 2x_{n+1}\right]^2 9^n \\ &= \left[\left(\sqrt{\lambda} + \frac{1-\sqrt{\lambda}}{2^n}\right) \frac{1}{3^n} - 2\frac{1-\sqrt{\lambda}}{4 \cdot 6^n}\right]^2 9^n \\ &= \left[\left(\sqrt{\lambda} + \frac{1-\sqrt{\lambda}}{2^n}\right) \frac{1}{3^n} - \frac{1-\sqrt{\lambda}}{2^{n+1}3^n}\right]^2 9^n \\ &= \left[\sqrt{\lambda} + \frac{2-2\sqrt{\lambda}-1+\sqrt{\lambda}}{2^{n+1}}\right]^2 \\ &= \left[\sqrt{\lambda} + \frac{1-\sqrt{\lambda}}{2^{n+1}}\right]^2 \end{aligned}$$

hemos llegado a que para $n + 1$ también se verifica el teorema, por tanto se verifica $\forall n$. |

Como los A_1, A_2, \dots son medibles Lebesgue y $A_1 \supset A_2 \supset \dots$ entonces, la medida de la intersección es el límite de la medida, luego el conjunto

$$C(\lambda) = \bigcap_{n=1}^{\infty} A_n(\lambda)$$

tiene medida le Lebesgue

$$m(C(\lambda)) = \lim_{n \rightarrow \infty} m(A_n(\lambda)) = \lambda > 0.$$

Para terminar nos queda demostrar que $C(\lambda)$ es una curva de Jordan.

Teorema 3.2. $C(\lambda)$ es una curva de Jordan, para todo $\lambda \in (0, 1)$.

Demostración. Primero tenemos que parametrizarlo, para ello construimos un conjunto de tipo Cantor \mathcal{C}_8 de la siguiente manera:

1. Eliminamos del intervalo unidad \mathcal{I} los ocho intervalos abiertos

$$\left(\frac{2k-1}{17}, \frac{2k}{17} \right) \quad k = 1, 2, \dots, 8$$

para quedarnos con los nueve intervalos cerrados

$$\left[\frac{2k}{17}, \frac{2k+1}{17} \right] \quad k = 0, 1, 2, \dots, 8$$

de longitud $1/17$ cada uno. Llamamos a la unión de estos nueve intervalos cerrados \mathcal{C}_8^1 .

2. A continuación, eliminamos de cada uno de los nueve intervalos cerrados, ocho intervalos abiertos de longitud $1/17^2$ cada uno. Esto nos dejaría un total de 81 intervalos cerrados de longitud $1/17^2$ cada uno, cuya unión llamamos \mathcal{C}_8^2 . Y seguimos repitiendo este proceso indefinidamente tal que

$$\mathcal{C}_8 = \bigcap_{j=1}^{\infty} \mathcal{C}_8^j(\lambda)$$

\mathcal{C}_8 es un conjunto de tipo Cantor, entonces todos sus puntos son puntos de acumulación por la izquierda, por la derecha o por ambos lados, dependiendo de si es un extremo izquierdo, derecho o uno de los puntos intermedios.

Ahora parametrizamos $C(\lambda)$ asignando de forma lineal el intervalo $\left(\frac{2k-1}{17}, \frac{2k}{17} \right)$ con $k = 1, 2, \dots, 8$ al segmento abierto que une S_k a S_{k+1} , $k = 1, 2, \dots, 8$ (ver

con una medida de Lebesgue tan cercana a 1 o a 0 como queramos. Si elegimos $\lambda \rightarrow 1$, entonces la curva de Jordan correspondiente tenderán hacia la curva de Peano (comparar Fig. 3.1 con Fig. 1.8). Aunque de esta forma la inyectividad se perdería ya que, en el límite, los cuadrados se tocarán y ciertos puntos de la curva pertenecerán a más de un cuadrado, y por lo tanto, conducirán a más de una preimagen.

Las curvas de Jordan en \mathbb{R}^n con una medida de Lebesgue n-dimensional positiva a menudo se denominan curvas de Osgood.

3.2. Las curvas de Jordan de Sierpinski y Knopp

Debido a la presencia de las uniones entre cuadrados, no todas las partes de la curva de Osgood son, por sí mismas, una curva de Osgood, y esto es lo que criticó Knopp. Por ello procedió a construir una curva de Osgood con la propiedad que cada parte de ella es, nuevamente, una curva de Osgood. Sierpinski ya había formulado una curva con estas propiedades, aunque algo más compleja.

Veamos primero la construcción que dió Sierpinski:

1. Partimos de un triángulo rectángulo isósceles T . Luego eliminamos un rectángulo abierto de área $(m(T))^2$ como se indica en la Fig. 3.3, quedándonos con los tres triángulos sombreados T_0, T_1, T_2 .
2. A continuación, eliminamos de cada uno de los triángulos T_0, T_1, T_2 , rectángulos abiertos de áreas $(m(T_0))^2, (m(T_1))^2, (m(T_2))^2$ respectivamente, y nos quedamos con los nueve triángulos sombreados que podemos ver en la Fig. 3.3. Repetimos el proceso indefinidamente.

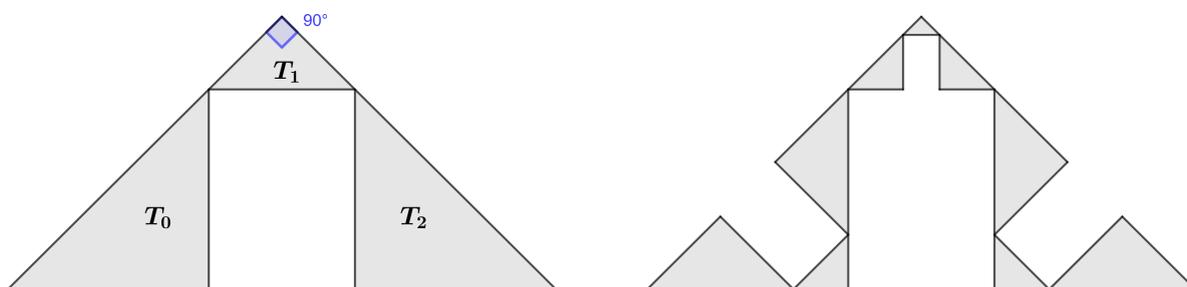


Figura 3.3. Generación de la curva de Osgood de Sierpinski.

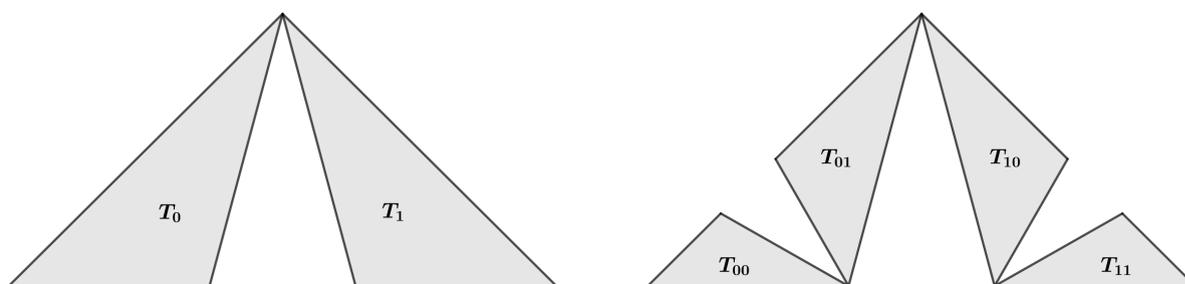


Figura 3.4. Generación de la curva de Knopp.

La intersección de todos estos triángulos sombreados resulta ser una curva de Jordan con medida de Lebesgue bidimensional positiva. La curva Osgood de Sierpinski tiene el inconveniente de que, para el triángulo inicial fijo, la medida de Lebesgue no se puede regular como en el caso del ejemplo de Osgood. Sin embargo, podemos observar que el uso de triángulos que se unen entre sí por vértices elimina la necesidad de esos segmentos de uniones entre los cuadrados que había en la curva de Osgood.

No estudiamos con detalle esta construcción ya que estudiaremos con más detalle la de Knopp, por ser una construcción más general. Knopp se basó en la idea de Sierpinski pero, en lugar de eliminar rectángulos en cada paso, eliminó triángulos y creó la siguiente construcción que le permitió crear curvas de Jordan con medida de Lebesgue $\lambda \in (0, 1)$.

1. Partimos de un triángulo T (no necesariamente rectángulo) y quitamos un triángulo abierto de área $r_1 m(T)$ donde $r_1 \in (0, 1)$. Por tanto nos quedan los dos triángulos sombreados T_0, T_1 cuya unión tiene área igual a $m(T)(1 - r_1)$. (Ver Fig. 3.4).
2. Luego de T_0 eliminamos un triángulo de área $r_2 m(T_0)$ y de T_1 eliminamos un triángulo de área $r_2 m(T_1)$ para algún $r_2 \in (0, 1)$. Nos quedamos con cuatro triángulos $T_{00}, T_{01}, T_{10}, T_{11}$, cuya área total es $m(T)(1 - r_1)(1 - r_2)$ y continuamos el proceso infinitamente.

En el límite obtendríamos el conjunto de puntos definido por la intersección

$$C = (T_0 \cup T_1) \cap (T_{00} \cup T_{01} \cup T_{10} \cup T_{11}) \cap \dots$$

dicho conjunto es no vacío por el teorema de la intersección de Cantor y tiene medida de Lebesgue claramente igual a

$$m(C) = m(T) \prod_{j=1}^{\infty} (1 - r_j).$$

| Teorema 3.3. *Sea una sucesión $\{a_n\}$. Si $0 < a_n < 1$, entonces $\prod_{n=1}^{\infty} (1 - a_n) > 0$*

si y solo si $\sum_{n=1}^{\infty} a_n$ es convergente.

Demostración.

$$\begin{aligned} \prod_{n=1}^{\infty} (1 - a_n) > 0 &\iff \exists \lim_{m \rightarrow \infty} \prod_{n=1}^m (1 - a_n) > 0 \\ &\iff \exists \lim_{m \rightarrow \infty} -\log \left(\prod_{n=1}^m (1 - a_n) \right) > 0 \\ &\iff \exists \lim_{m \rightarrow \infty} \sum_{n=1}^m (-\log(1 - a_n)) > 0 \\ &\iff \sum_{n=1}^{\infty} (-\log(1 - a_n)) \text{ es convergente} \end{aligned}$$

Por el criterio de comparación por paso al límite

$$\lim_{n \rightarrow \infty} \frac{a_n}{-\log(1 - a_n)} = \lim_{n \rightarrow \infty} (1 - a_n) \in (0, 1)$$

tenemos que ambas series tienen el mismo comportamiento, es decir ambas convergen o ambas divergen. Por tanto

$$\sum_{n=1}^{\infty} (-\log(1 - a_n)) \text{ converge} \iff \sum_{n=1}^{\infty} a_n \text{ converge}$$

Como queríamos demostrar. |

Por este último teorema, si elegimos los r_j de forma que $\sum_{j=1}^{\infty} r_j$ sea convergente, entonces $m(C) > 0$.

Si además en cada paso elegimos correctamente los triángulos que eliminamos entonces el resto de triángulos se contraen a puntos, aunque dicha afirmación se vea claramente en las figuras, no hemos encontrado un algoritmo que nos indique como tenemos que ir quitando los triángulos para que efectivamente se contraigan a puntos.

Knopp dió una idea de como podemos eliminar los triángulos para que esto ocurra, pero no ofreció una prueba clara. Veámos dicha idea. Partamos de un triángulo obtuso cuyos ángulos menores valgan α , y en cada paso eliminemos un triángulo de forma que los triángulos restantes sigan siendo obtusos y con ángulos mayores o iguales a α . Si denotamos c a la longitud de la base del triángulo inicial y extendemos dicho triángulo por un de sus lados como se indica en la Fig. 3.5, tenemos que $c \cos(\alpha) = b$. Vemos claramente que la base de los triángulos de la etapa siguiente son menores que b y como hemos quitado los triángulos de forma que los ángulos restantes no fueran menores que α también tenemos que el coseno de cualquiera de los ángulos será menor que el $\cos(\alpha)$ pues a mayor ángulo, menor coseno. Luego si seguimos este razonamiento llegamos a que la base de los triángulos obtenidos en la etapa n no puede exceder a $c \cos^n(\alpha)$, que tiende a cero cuando $n \rightarrow \infty$.

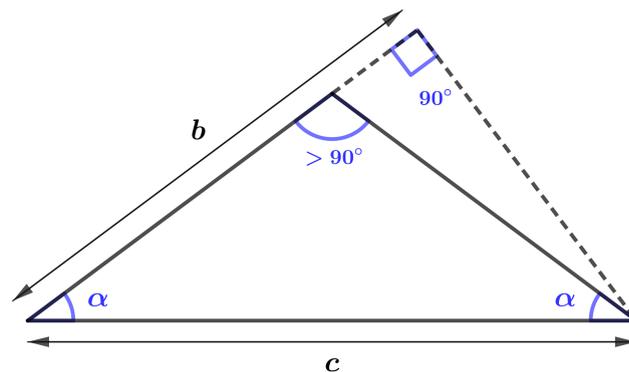


Figura 3.5. Extensión triángulo inicial.

Si elegimos como triángulo inicial un triángulo rectángulo con una base de longitud 2, por tanto $m(T) = 1$, y ponemos que $r_j = r^2/j^2$ para $r \in (0, 1)$, entonces tendríamos

$$m(C) = \prod_{j=1}^{\infty} (1 - r^2/j^2).$$

Por el teorema de factorización de Weierstrass

$$\text{sen}(\pi z) = \pi z \prod_{j=1}^{\infty} (1 - z^2/j^2).$$

tenemos,

$$m(C) = \prod_{j=1}^{\infty} (1 - r^2/j^2) = \frac{\text{sin}(\pi r)}{\pi r}.$$

Observemos que para todo $\lambda \in (0, 1)$, la ecuación $\frac{\sin(\pi r)}{\pi r} = \lambda$ tiene una solución $r \in (0, 1)$ por la continuidad de $\frac{\sin x}{x}$ en $[0, \pi]$. Por tanto, si elegimos r adecuadamente podemos obtener un conjunto $C(\lambda)$ con medida de Lebesgue λ para cualquier $\lambda \in (0, 1)$.

| Teorema 3.4. *La curva $C = (T_0 \cup T_1) \cap (T_{00} \cup T_{01} \cup T_{10} \cup T_{11}) \cap \dots$ es una curva de Jordan con medida de Lebesgue positiva.*

Demostración. Ya hemos visto que C tiene medida de Lebesgue perteneciente a $(0, 1)$, con lo cual tiene medida positiva. Para demostrar que C es una curva de Jordan, primero tenemos que parametrizar el conjunto, para ello construimos una aplicación $f : \mathcal{I} \rightarrow C$ que nos lleve el intervalo $[0, 1/2]$ en el triángulo T_0 y el intervalo $[1/2, 1]$ en T_1 , tal que $t = 1/2$ vaya al punto que T_0 y T_1 tienen en común. De modo que el intervalo $[0, 1/4]$ lo lleve al triángulo T_{00} , el $[1/4, 1/2]$ al T_{01} , el $[1/2, 3/4]$ al T_{10} y el $[3/4, 1]$ al T_{11} , donde el punto $t = 1/4$ está relacionado con el punto en común de T_{00} y T_{01} , y el punto $t = 3/4$ con el punto en común de los triángulos T_{10} y T_{11} .

Si representamos cada punto $t \in (0, 1]$ por un número binario infinito

$$t = 0_2 b_1 b_2 b_3 \dots$$

donde cada binario finito de la forma $0_2 b_1 b_2 b_3 \dots b_n 1$ lo reemplazamos por el binario infinito $0_2 b_1 b_2 b_3 \dots b_n \bar{1}$. Luego, por la aplicación f que hemos definido anteriormente tenemos que $f(t)$ se encuentra en el triángulo T_{b_1} , en el $T_{b_1 b_2}$, en el $T_{b_1 b_2 b_3} \dots$. Dado que

$$T_{b_1} \supset T_{b_1 b_2} \supset T_{b_1 b_2 b_3} \supset \dots$$

es una secuencia de triángulos cerrados anidados que se contraen a puntos, tenemos que

$$T_{b_1} \cap T_{b_1 b_2} \cap T_{b_1 b_2 b_3} \cap \dots$$

define un único punto que, por construcción se encuentra en C :

$$f(0_2 b_1 b_2 b_3 \dots) = T_{b_1} \cap T_{b_1 b_2} \cap T_{b_1 b_2 b_3} \cap \dots$$

Faltaría probar que f es continua e inyectiva, veámoslo:

- f es continua: Si $|t_1 - t_2| < 1/2^n$, entonces $f(t_1), f(t_2)$ se encuentran en el peor de los casos en dos triángulos adyacentes y su distancia no puede superar la suma de las longitudes de los dos lados más largos, pero como los triángulos se reducen a puntos, entonces la longitud de sus lados también, por tanto hemos probado la continuidad de f .
- f es inyectiva: Cualquier punto de C tiene que estar en T_0 o T_1 , a su vez, debe de estar en uno de los cuatro triángulos T_{00}, T_{01}, T_{10} o T_{11} , que a su vez debe de estar en uno de los triángulos $T_{b_1 b_2 b_3}, \dots$. Tenemos que tener en cuenta que no puede estar en dos triángulos de la misma partición, excepto cuando es el único punto que es común a dos triángulos adyacentes, en cuyo caso, tiene como única imagen previa en $[0, 1]$ el punto que es común a dos subintervalos cerrados adyacentes, es decir, si por ejemplo el punto pertenece tanto a T_0 como a T_1 entonces la imagen previa sería el $1/2$ que es el punto común de sus intervalos. Entonces a la secuencia de triángulos cerrados anidados le corresponde una única secuencia de intervalos cerrados anidados

$$\left[\frac{b_1}{2}, \frac{b_1}{2} + \frac{1}{2} \right] \supset \left[\frac{b_1}{2} + \frac{b_2}{4}, \frac{b_1}{2} + \frac{b_2}{4} + \frac{1}{4} \right] \supset \left[\frac{b_1}{2} + \frac{b_2}{4} + \frac{b_3}{8}, \frac{b_1}{2} + \frac{b_2}{4} + \frac{b_3}{8} + \frac{1}{8} \right] \supset \dots$$

que define un único punto en $[0, 1]$.

■

Comentario 3.1. Si comenzamos con un triángulo isósceles cuyo ángulo superior es de 120° , tomamos $r_j = 1/3$ y en cada paso eliminamos triángulos isósceles de área $\frac{1}{3}m(T_j)$ tal y como se muestra en la Fig. 3.6. Entonces, como podemos ver, los triángulos restantes son similares a los triángulos iniciales ya que vuelven a tener un ángulo obtuso de 120° . Con estos valores se obtiene una curva de Jordan con medida de Lebesgue nula

$$m(C) = \prod_{j=1}^{\infty} (1 - 1/3) = \lim_{n \rightarrow \infty} (2/3)^n = 0.$$

Esta curva en el límite es conocida como la curva de von Koch, que no es diferenciable en ningún punto, ver Fig. 3.7.

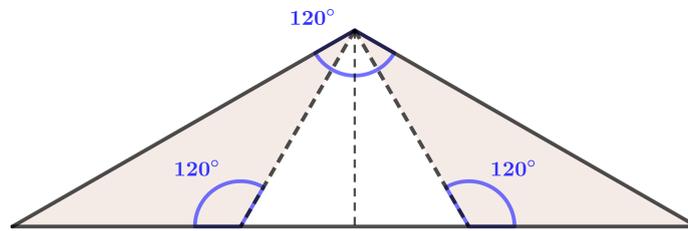


Figura 3.6. Eliminación de los triángulos para generar la curva de von Koch.

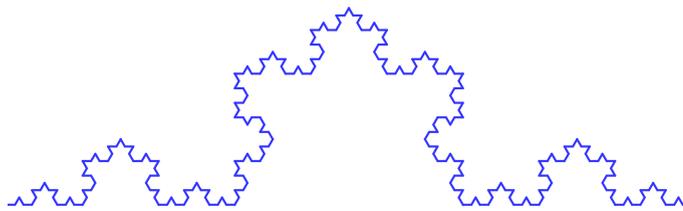


Figura 3.7. Quinta iteración de la curva de von Koch.

3.3. Otras curvas de Osgood

En las secciones anteriores discutimos las familias de las curvas de Osgood, una de las cuales tenía como límite la curva de Peano, mientras que la otra tenía como límite la curva de Sierpinski-Knopp. Gelbaum y Olmsted estudiaron como se podría obtener una familia de curvas de Osgood cuyo límite fuera la curva de Hilbert. En la Fig. 3.8 hemos indicado los dos primeros pasos. Podemos observar que esta construcción también requiere el uso de segmentos de unión, como en el primer caso que estudiamos, la curva de Osgood.

Por otro lado, Lance y Thomas propusieron una familia de curvas de Osgood en cuyo límite coincide con la curva de Lebesgue. Podemos ver los dos primeros pasos en la Fig. 3.9. Si en el primer paso elegimos que la distancia entre los cuadrados sea $x_1 = \frac{1 - \sqrt{\lambda}}{3}$, en el segundo paso elegimos $x_2 = \frac{1 - \sqrt{\lambda}}{3^2}$, en el tercero $x_3 = \frac{1 - \sqrt{\lambda}}{3^3}$, etc. Entonces de forma general tenemos $x_n = \frac{1 - \sqrt{\lambda}}{3^n}$. Si denotamos A_n a la región que nos queda en el paso n , es decir a los 4^n cuadrados y los 4^{n-1} segmentos de línea que quedan después del n -ésimo, llegamos a que

$$m(A_n(\lambda)) = \{1 - [1 - (2/3)^n][1 - \sqrt{\lambda}]\}^2$$

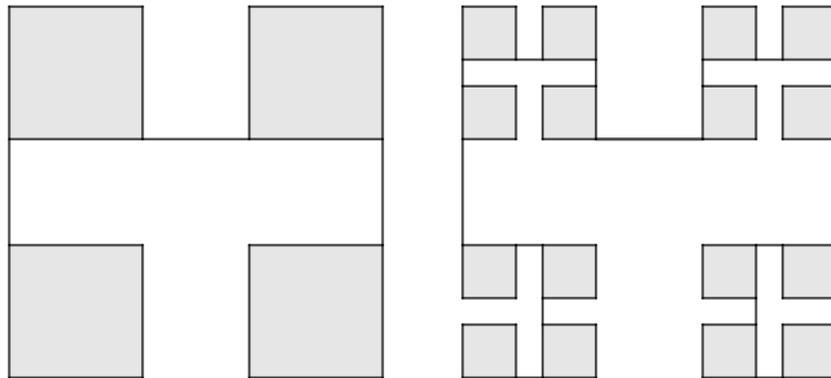


Figura 3.8. Generación de la curva de Osgood cuyo límite nos da la curva de Hilbert.

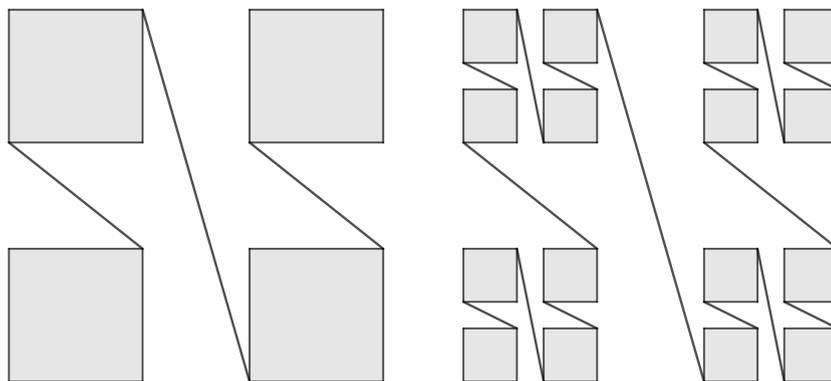


Figura 3.9. Generación de la curva de Osgood cuyo límite nos da la curva de Lebesgue.

cuya demostración es análoga al teorema 3.1. Por tanto, como $C(\lambda) = \bigcap_{n=1}^{\infty} A_n(\lambda)$, tenemos que

$$m(C(\lambda)) = \lim_{n \rightarrow \infty} [\sqrt{\lambda} + (2/3)^n(1 - \sqrt{\lambda})]^2 = \lambda.$$

Para probar que $C(\lambda)$ es una curva de Jordan usamos la misma idea que en la demostración del teorema 3.2, pero en este caso asignamos de forma lineal cada intervalo eliminado en la construcción del conjunto de Cantor \mathcal{C} al segmento de unión de los cuadrados, y los intervalos restantes a los cuadrados restantes.

Luego hemos encontrado otra familia de curvas de Jordan cuya medida de Lebesgue puede valer cualquier número dado entre 0 y 1.

Complemento

Durante todo el trabajo hemos estado estudiando diferentes curvas que rellenan superficies (space-filling curves). En este apartado vamos a mostrar algunas aplicaciones que tienen dichas curvas, viendo con más detalle la aplicación de la curva de Hilbert a procesamiento de imágenes.

Procesamiento de imágenes.

En 1991 la convención SIGGRAPH publicó una aplicación de la curva de Hilbert para procesamiento de imágenes. Dicha aplicación presenta una nueva técnica para la representación de fotos en blanco y negro en un dispositivo gráfico como una impresora láser. El problema radica en el hecho de que una impresora representa una matriz de píxeles en blanco y negro, mientras que los tonos de grises no se pueden representar a nivel de píxeles. Para solventar dicho problema se han utilizado varias técnicas llamadas de tramado, se basan en escanear una imagen en blanco y negro línea por línea o en pequeños bloques cuadrados y luego se hace una aproximación en blanco y negro de la imagen con el objetivo de minimizar el error. Pero en este caso se proporcionó un método diferente utilizando la curva de Hilbert, dicho método se basa en hacer pasar una curva de Hilbert por todos los píxeles de la imagen, lo que sería una alternativa al escaneo de la imagen línea por línea, pero veámoslo con más detalle.

En las imágenes de la Fig.1 podemos ver el principio del algoritmo de interpolación basado en cuatro etapas sucesivas del escaneo mediante la curva de Hilbert de una imagen. Ninguna secuencia de píxeles consecutivos a lo largo de este camino puede ser reemplazada por una aproximación en blanco y negro. La ventaja del es-

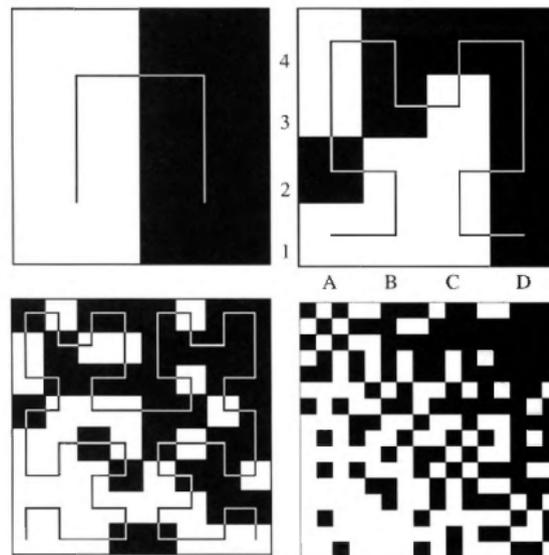


Figura 1. Algoritmo de escaneo mediante la curva de Hilbert.

Peitgen, H., Jürges, H., Saupe, D., *Chaos and Fractals*. Springer (2004).

caneeo de la imagen utilizando la curva de Hilbert es que de esta forma está libre de cualquier característica direccional presente en los métodos tradicionales, como por ejemplo el de línea a línea. De esta forma produce patrones aperiódicos de puntos agrupados con características similares a las estructuras de granulado fotográfico. La Fig.2 compara el enfoque tradicional con el nuevo método. A continuación vamos a describir los detalles de una versión simplificada del algoritmo de degradado con la curva de Hilbert.

Consideremos una imagen cuadrada de tonos grises que varían continuamente y la aproximamos por una imagen que puede contener solamente píxeles blancos y negros. La resolución de la imagen de salida debe ser una potencia de 2. Por ejemplo, consideremos imágenes con 2, 4, 8 y 16 píxeles por fila y columna (ver Fig. C.1), como podemos ver, podemos ajustar exactamente una curva de Hilbert al mosaico de la imagen. Esto introduce una forma de ordenar los píxeles. Para el ejemplo de 4 por 4 píxeles, donde etiquetamos las columnas con las letras A, B, C, D y las filas con 1, 2, 3, 4, los píxeles de la imagen se ordenan de la siguiente manera:

$$A1, B1, B2, A2, \dots, D2, C2, C1, D1$$

es decir, ordenamos los píxeles siguiendo el mismo recorrido de la curva de Hilbert.

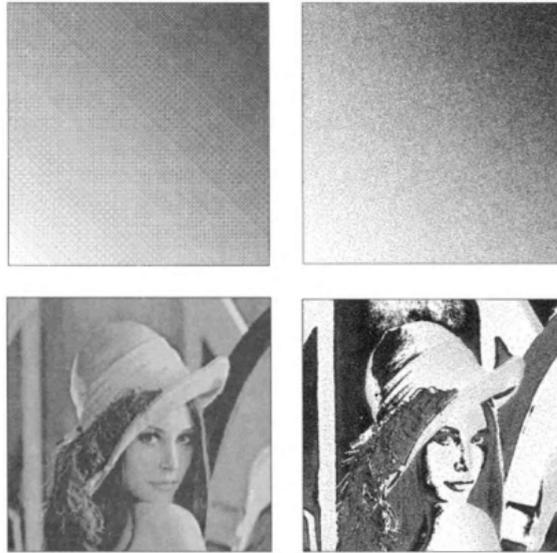


Figura 2. Comparación método mediante línea a línea y mediante la curva de Hilbert.

Peitgen, H., Jürges, H., Saupe, D., *Chaos and Fractals*. Springer (2004).

Si denotamos por

$$I_1, I_2, \dots, I_n$$

a los valores de intensidad de los píxeles correspondientes a la imagen de entrada (que van desde 0 para el negro al 1 para el blanco). Como el número total de píxeles es una potencia de 2, entonces n es una potencia de 2. Para la definición de la imagen de salida tenemos que calcular sus valores de intensidad correspondientes,

$$O_1, O_2, \dots, O_n \in 0, 1$$

podemos definir

$$O_1 = \begin{cases} 0 & \text{si } I_1 \leq 0,5 \\ 1 & \text{si } I_1 > 0,5 \end{cases}$$

Esta aproximación tiene un error

$$E_1 = I_1 - O_1$$

pero en lugar de ignorar este error lo pasamos al siguiente píxel de la secuencia, es decir,

$$O_2 = \begin{cases} 0 & \text{si } I_2 + E_1 \leq 0,5 \\ 1 & \text{si } I_2 + E_1 > 0,5 \end{cases}$$

Más precisamente, tendríamos

$$O_k = \begin{cases} 0 & \text{si } I_k + E_{k-1} \leq 0,5 \\ 1 & \text{si } I_k + E_{k-1} > 0,5 \end{cases}$$

y

$$E_k = I_k + E_{k-1} - O_k$$

En otras palabras, el error se difunde a lo largo de la secuencia de píxeles, el objetivo de esta difusión de errores es minimizar el error general. Por ejemplo, tenemos que los errores sumados en la imagen completa es igual a

$$\sum_{k=1}^n E_k = E_n$$

que se espera que sea relativamente pequeño. El punto crucial del algoritmo es que el error se difunde a lo largo de la curva de Hilbert que traza la imagen de una manera que se concibe como muy irregular para nuestro sistema sensorial, es decir, no percibiremos dicho error. Pero si reemplazamos la curva de Hilbert por ejemplo por una curva que escanea la imagen fila por fila, la difusión del error se percibiría más. El algoritmo propuesto en SIGGRAPH es una generalización de este método, considera bloques de píxeles consecutivos del escaneo de Hilbert a la vez, en lugar de píxeles individuales.

Otras aplicaciones

Como ya hemos comentado, existen muchas aplicaciones de las curvas que rellenan superficies a una gran variedad de disciplinas diferentes, desde a las propias matemáticas [3], [9] como a la biología [12] o incluso al arte [5]. En esta sección daremos una idea de otros tipos de aplicaciones que nos han resultado interesante aunque en la bibliografía nombramos algunos más por si resulta de vuestro interés.

El problema del viajante

El problema del viajante, conocido como TSP (travelling salesman problem) es uno de los problemas más conocidos: un viajante quiere visitar n ciudades, pasando

una sola vez por cada una de ellas, empezando y terminando en la misma ciudad. ¿Cuál es el recorrido que debe seguir para que la distancia total que recorra sea mínima?

El artículo [6] estudia el recorrido dado por la heurística de una curva que rellena el espacio. Dicha heurística se basa en una aplicación sobreyectiva $f : \mathcal{I} \rightarrow \mathcal{Q}$ tal que para cada $x \in \mathcal{Q}$, se puede calcular rápidamente un $t \in \mathcal{I}$ tal que $f(t) = x$. Como hemos estado estudiando en capítulos anteriores, hay varias curvas que verifican esto, y cualquiera de ellas puede usarse para construirla. La idea esencial de dicha heurística es la sugerencia de que podemos encontrar un recorrido corto a través de un conjunto de n puntos $\{x_1 \dots x_n\} \subset \mathcal{Q}$ recorriendo los puntos en el orden de sus preimágenes en \mathcal{I} . Veamos el proceso más formalmente

1. Primero calculamos el conjunto de puntos $\{t_1 \dots t_n\} \subset \mathcal{I}$ tal que $f(t_i) = x_i$ para cada $i = 1 \dots n$.
2. Luego ordenamos los t_i tal que $t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(n)}$
3. Finalmente, definimos la permutación $\sigma : [1, n] \rightarrow [1, n]$ imponiendo que $x_{\sigma(i)} = f(t_{(i)})$.

A la ruta que pasa por los puntos $\{x_1 \dots x_n\}$ siguiendo el orden $x_{\sigma(1)}, \dots, x_{\sigma(n)}, x_{\sigma(1)}$ se le denominará recorrido de la curva que rellena el espacio.

En el artículo se prueba que la distancia del recorrido de la curva que rellena el espacio tienen una diferencia muy sutil de la solución óptima del problema TSP.

Búsqueda de información geográfica

Actualmente en cualquier disciplina a la mayor parte de la información que manejamos le podemos asignar una posición geográfica. Sin ir más lejos, algunas de las aplicaciones más útiles para la vida diaria dependen del acceso rápido a datos geoespaciales, como por ejemplo Google maps o Uber.

Pero veamos la complejidad que tiene manejar los datos geográficos: Normalmente representamos la Tierra como una superficie plana pero la computación de estos datos debe ser aún más simple, para ello registramos ordenadamente los datos mediante números, es decir, asociamos a cada punto de la Tierra un número el cual lo

podemos representar en una línea unidimensional. Para ello necesitamos una forma de asignar dichos números con sus respectivas ubicaciones en el plano.

El artículo [7] explica como se pueden asignar siguiendo la curva de Hilbert. Primero se introduce la Tierra en un cubo, cada cara de dicho cubo se rellena con la curva de Hilbert y se proyecta dicha curva sobre la superficie de la Tierra, teniendo en cuenta que la curva en cada cara está conectada a dos caras adyacentes, formando una sola línea.

De esta forma podemos representar una ubicación concreta como un punto específico de un intervalo. Las versiones menos precisas del mismo número representan la misma ubicación en escalas mayores, es decir, en vez de un punto concreto representaría un entorno, hasta llegar a la menos precisa que nos indicaría el lado del planeta en el que se encuentra.

Como podemos ver, proyectar la curva de Hilbert en la Tierra nos da una forma muy eficiente de representar un lugar como un solo número.

Almacenaje de Energía

Un condensador es un elemento capaz de almacenar energía eléctrica. Está compuesto por dos placas conductoras separadas por un material dieléctrico, es decir, un material que conduce mal la electricidad, lo que genera una diferencia de voltaje entre ambas placas. Cuando conectamos el condensador a una fuente de voltaje comienza a circular corriente por el circuito y una de las placas adquiere carga negativa y la otra carga positiva, al apagar la fuente de voltaje, si conectamos alguna carga al condensador, comenzará a circular corriente desde el condensador hacia la carga hasta descargarse. Los supercondensadores están formado por varios condensadores, generalmente colocados paralelamente, por tanto son dispositivos capaces de almacenar grandes cantidades de energía.

El artículo [13] propone un diseño alternativo para los supercondensadores realizados por espumas de grafeno inducidos por láser. Para ello introducen el concepto de electrodos anidados basados en la curva de Peano. Los condensadores son ordenados siguiendo bajas iteraciones de la curva de Peano, de esta forma consiguen una configuración simétrica que incorpora múltiples celdas de condensadores paralelos

que comparten electrodos comunes para una eficiencia máxima y una longitud óptima. Llegan a la conclusión de que este diseño de supercondensador tiene una mayor capacidad de almacenamiento de energía.

Bibliografía

- [1] AFTOSMIS, M.J., BERGER, M.J., MURMAN, S.M., *Applications of Space-Filling Curves to Cartesian Methods for CFD*. AIAA (Enero 2004)
- [2] BADER, M., *Space-Filling Curves an Applications in Scientific Computing*. Springer (2013)
- [3] BARTHOLDI, J.J., *A Routing System Based on Spacefilling Curves*. (11 abril 1995)
- [4] BORRELL, R., CAJAS, J.C., MIRAA, D., TAHA, A., KORIC, S., VÁZQUEZ, M., HOUZEAUXA, G., *Parallel mesh partitioning based on space filling curves*. Elsevier (31 enero 2018)
- [5] DON RELYEA, *Space Filling Curve Art 2004-2014*. (16 febrero 2013). Recuperado de <http://www.donrelyea.com/site2015/2013/02/16/space-filling-curve-art-2004-2014/>
- [6] GAO, J., STEELE, J.M., *General Spacefilling Curve Heuristics and Limit Theory for the Traveling Salesman Problem*. Academic Press (1994)
- [7] IVORRA, C., *La curva de Knopp* (3 agosto 2017) Recuperado de: <http://rinconmatematico.com/foros/index.php?action=printpage;topic=97476.0>
- [8] KREISS, S., *S2 cells and space-filling curves: Keys to building better digital map tools for cities*. SideWalkLabs (27 junio 2016). Recuperado de <https://sidewalklabs.com>
- [9] LAWDER, J.K., KING, P.J.H., *Using Space-Filling Curves for Multi-dimensional Indexing*. Springer (2000)

- [10] LERA, D., SERGEYEV. Y.D., *Deterministic global optimization using space-filling curves and multiple estimates of Lipschitz and Hölder constants*. Elsevier (2 diciembre 2014)
- [11] PEITGEN, H., JURGES, H., SAUPE, D., *Chaos and Fractals*. Springer (2004).
- [12] SAGAN, H., *Space-Filling Curves*. Springer-Verlag (1994)
- [13] SMREK, J., GROSBERG, A.Y., *A novel family of space-filling curves in their relation to chromosome conformation in eukaryotes*. Elsevier (20 agosto 2013)
- [14] TILIAKOS, A., TREFILOV, A., TANASA, E., BALON, A., STAMATIN, I., *Space-Filling Supercapacitor Carpets: Highly scalable fractal architecture for energy storage*. Elsevier (5 marzo 2018)
- [15] WILLARD, ST., *General Topology*. Addison-Wesley (1970)