DOMINIOS DE REPRESENTACIÓN ALGEBRAICOS Y CONTINUOS: ARITMÉTICA REAL EXACTA

María Isabel David Otálvaro
Licenciada en Matemáticas
Código: 830069

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Matemáticas
Bogotá, D.C.
Junio de 2012

DOMINIOS DE REPRESENTACIÓN ALGEBRAICOS Y CONTINUOS: ARITMÉTICA REAL EXACTA

María Isabel David Otálvaro
Licenciada en Matemáticas
Código: 830069

Trabajo de grado para optar al título de
Magister en Matemáticas

Director
Rodrigo de Castro Korgi, Ph.D.
Doctor en Matemáticas

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Matemáticas
Bogotá, D.C.
Junio de 2012

Nota de aceptación

Trabajo de tesis

Jurado

Director
Rodrigo de Castro Korgi

Bogotá, D.C., Junio de 2012

Dedicado a

A Fernando Guerra por su apoyo incondicional.

Agradecimientos

Agradezco al profesor Rodrigo de Castro Korgi por su múltiples observaciones útiles, por su paciencia y sus enseñanzas.

Resumen

Los dominios son estructuras diseñadas para modelar computación vía aproximaciones. Este documento presenta una teoría efectiva de representación de los números reales por dominios continuos con su implementación en términos de LFT's, y por dominios algebraicos con su respectiva implementación en intervalos centrados diádicos.

Palabras clave: Dominios algebraicos, dominios continuos, efectividad, intervalos centrados diádicos, transformaciones lineales fraccionarias.

Abstract

Domains are ordered structures designed to model computation with approximations. This document present a framework to effective representation of real numbers by continuous domains with its implemented in terms of linear fractional transformations, and by algebraic domains with its implemented in dyadic centred intervals.

Keywords: Algebraic domains, continuous domains, effectivity, dyadic centred intervals, linear fractional transformations.

Índice general

Resumen **I**

Índice general **II**

Introducción **IV**

1. Preliminares **1**

1.1. Efectividad 1

1.2. Algunas aproximaciones a la computabilidad 2

1.3. Funciones Recursivas y Recursivamente Enumerables 2

1.3.1. Funciones recursivas primitivas 2

1.3.2. Funciones recursivamente enumerables 5

1.4. Conjuntos numerables 6

1.5. Predicados decidibles y parcialmente decidibles 9

1.6. Conjuntos Recursivos y Recursivamente Enumerables 10

1.7. Hasta ahora qué podemos decir de la computabilidad en $\mathbb{R}$? 11

1.8. Conjuntos Ordenados 12

1.8.1. Ordenes de Información 12

2. Teoría de Dominios **14**

2.1. Dominios Continuos y Dominios Algebraicos 14

2.2. La topología de Scott 22

2.3. Aproximaciones, Ordenamientos y Dominios 26

2.3.1. Aproximaciones y ordenamientos 26

2.3.2. Ideales y dominios 27

2.3.2.1. Construcción de Dominios Algebraicos 27

2.3.2.2. Construcción de Dominios Continuos 29

2.4. Una R.A. para los Números Reales. 30
2.4.1. Computabilidad 31
2.5. El dominio continuo de intervalos 32
2.5.1. Computabilidad en la recta real 33
2.6. Algunos Modelos para espacios clásicos 34
2.6.1. Una Representación de Dominio Algebraica en Espacios Topológicos 34
2.6.2. Otros Dominios Continuos de Representación 35
2.6.2.1. El Espacio Superior 35
2.6.2.2. El espacio de las bolas formales de un espacio métrico completo 36

3. Intervalos Centrados Diádicos 37
3.1. Aritmética Real Exacta 37
3.2. Preeliminares 38
3.3. Intervalos Centrados Diádicos como Aproximaciones 39
3.4. Estructura de las Aproximaciones 39
3.5. Computando una operación real exacta 40
3.5.1. Redondeo 40
3.5.2. Adición 42
3.5.3. Multiplicación 43

4. Transformaciones Lineales Fraccionarias 48
4.1. Introducción 48
4.2. Representación de los Números Reales 51
4.3. Semigrupos de LFTs de Refinamiento 54
4.4. Punto flotante exacto 56
4.4.1. Matrices Signo 57
4.4.2. Matrices Dígito 58
4.5. LFTs como operadores 59
4.6. LFTs de refinamiento en [-1,1] 61
4.7. Tensores 62
4.8. Tensores de refinamiento en [-1,1] 67
4.9. Escogiendo el intervalo base 67

Conclusiones y recomendaciones 69

Bibliografía 70

Introducción

La teoría de dominios fue introducida por Dana Scott en 1970 como una teoría matemática para la semántica de lenguajes de programación. Varias categorías cartesianas cerradas de dominios han sido empleadas en la semántica de los lenguajes de computación. En particular, para obtener modelos no triviales de los λ-cálculos y proveer una semántica denotacional para el lenguaje PCF (Programming Language for Computable Functions). Scott sugirió que el dominio de los intervalos compactos de la recta real puede usarse como un tipo de dato para números reales.

Los dominios algebraicos se caracterizan mediante un subconjunto de elementos llamados finitos o compactos, y representan resultados computacionales que pueden obtenerse en un número finito de pasos. Los elementos compactos del dominio forman una base (todo elemento del dominio es el límite de elementos de la base). Estos dominios se han utilizado para representar espacios clásicos en matemáticas. Stoltenberg-Hansen y Tucker han mostrado como representar espacios topológicos [13] en un escenario general. Jens Blanck [4] ha mostrado cómo hacer una inmersión de espacios métricos en dominios algebraicos haciendo énfasis en los dominios efectivos. Introduce una noción natural de espacio métrico efectivo considerando un espacio de las bolas formales.

Los dominios continuos son generalizaciones de los dominios algebraicos y comparten muchas de sus propiedades básicas. Los dominios continuos han sido estudiados por Edalat [9] para representar los reales como los elementos maximales del dominio de intervalos. Una base contable está dada por el conjunto de todos los intervalos con extremos racionales y un número real x se aproxima mediante una sucesión contractante de intervalos encajados de extremos racionales. En [8] el autor presenta la noción de un modelo computacional de teoría de dominios para un espacio Hausdorff localmente compacto con base contable. Esto muestra que los dominios continuos obtenidos tomando subconjuntos compactos no vacíos del espacio, ordenados por inclusión reversa, proveen un modelo simple para computación en estos espacios clásicos que están inmersos en el subespacio de los elementos maximales del dominio. El dominio *espacio superior* esta equipado con una base de subconjuntos compactos no vacíos que pueden ser enumerados para dar una teoría de efectividad. Un elemento del espacio localmente compacto identificado como un singleton puede ser obtenido como la intersección de una secuencia creciente encajada de elementos de la base.

Si se quiere extender la teoría de computabilidad (en el contexto de las funciones recursivas parciales) a estructuras tales como la de los números reales, se tiene que computar sobre las aproximaciones de los elementos de la estructura. Los dominios de representación de una estructura pueden verse como una forma de introducir una teoría de aproximaciones a la estructura. Desde los dominios se tiene una teoría de la computabilidad natural con una teoría de computaciones en las aproximaciones de la estructura usando la teoría de numeraciones de Mal'cev-Ershov para extender la computabilidad desde los números naturales a dominios.

Siguiendo la idea de Scott del dominio de intervalos como una representación de los números reales, varios autores han trabajado en la noción de un tipo de dato para números reales.

Blank [2] considera implementaciones de aritmética real exacta usando intervalos centrados diádicos (un número diádico es de la forma $k/2^n$ donde n es un número natural y k es un entero). Estos son intervalos representados por una pareja de numeros diádicos que dan el punto central y el radio del intervalo. El beneficio de esta representación comparado con intervalos de extremos diádicos es que hay menor costo en las operaciones básicas. La completación por ideales de los intervalos centrados diádicos es un dominio algebraico.

Un esquema práctico para computación exacta usando LFTs (Linear Fractional Transformations) y basado en teoría de dominios continuos ha sido introducida por Edalat y Potts [10]. Los IFS (Iterated Fractional Systems) pueden ser utilizados para representar números reales con números de punto flotante en una base b. Esta idea es luego generalizada para LFTs y juega un rol crucial en computación real exacta.

Una cuestión fundamental es que si un escenario factible para computación exacta puede ser desarrollado para cálculos numéricos básicos, estos puede ser ejecutado sin errores de redondeo.

El lenguaje de programación PCF resulta adecuado para definir los intervalos diádicos como un tipo de dato. En su trabajo doctoral, P. Di Gianantonio (A Functional Approach to Computability on Real Numbers, Ph.D. thesis, University of Pisa, 1993) presentó una extensión de PCF con un tipo de dato real como un dominio algebraico cuyos elementos compactos son isomorfos con el conjunto de los intervalos con extremos números diádicos ordenados por inclusión reversa. Un número real es entonces representado por una secuencia contractante encajada de intervalos diádicos que puede ser considerada como un real aproximado. Son incluidas nuevas constantes en el lenguaje para sumar y restar uno a los reales, multiplicar y dividir por 2 y un predicado para comparar reales con cero. Él mostró que cualquier función real computable puede definirse en este lenguaje.

La aproximación por LFT's ha sido implementada efectivamente en los lenguajes de programación Calm, C++ y PCF. M. H. Escardó en su tesis doctoral (PCF extended with real numbers, Ph.D. thesis, University of London Imperial College, Department of Computing, 1997) desarrolló una extensión de PCF con un tipo de dato real interpretado como un dominio continuo de intervalos. Más aún, Escardó mostró que la integración de Riemman puede ser introducida en este lenguaje.

Este trabajo es una monografía y se enfoca en estudio de dominios de representación algebraicos y continuos, una teoría de efectividad, y la implementación de las operaciones básicas como suma y multiplicación de números reales por medio de aritmética real exacta utilizando intervalos centrados diádicos y LFTs. Aunque los resultados teóricos no son propios se organizan los temas alrededor de la necesidad de establecer una teoría de la computabilidad para los números reales, y se trata de enriquecer los temas tratados dando algunos ejemplos para clarificar y reconocer las estructuras estudiadas.

CAPÍTULO 1

Preliminares

1.1. Efectividad

Un algoritmo o procedimiento efectivo es una regla mecánica, método automático o programa para ejecutar alguna operación matemática. Algunos ejemplos son:

1. Dado n encontrar el n-ésimo número primo.
2. Encontrar el máximo común divisor de dos números (Algoritmo de Euclides).
3. Dados dos números x y y, decidir si x es múltiplo de y.

Cuando un algoritmo o procedimiento efectivo es usado para calcular los valores de una función numérica entonces la función en cuestión es descrita por frases tales como *efectivamente calculable, algorítmicamente computable, efectivamente computable* o simplemente *computable*. Por ejemplo, las funciones $MCD(x, y)$ =el máximo común divisor de x y y, y $f(n)$ =el n-ésimo número primo, son computables en este sentido informal. Por otro lado considere la siguiente función:

$$g(n) = \begin{cases} 1 \text{ si existe un ciclo de } n \text{ 7's consecutivos} \\ \quad \text{en la expansión decimal de } \pi \\ 0 \text{ en otro caso} \end{cases}$$

¿g es computable? Existe un procedimiento mecánico para generar sucesivamente los dígitos en la expansión decimal de π, se sugiere el siguiente procedimiento para computar g.

Dado n empezar a generar la expansión decimal de π un dígito a la vez y observar los sietes. Si aparecen n sietes consecutivos el proceso para y $g(n) = 1$. Si tal secuencia de sietes no aparece $g(n) = 0$.

El problema de este procedimiento es que, si para algún n particular no existen n sietes consecutivos, no existe un estado en el proceso en el que podamos parar y concluir que este es el caso. En cualquier estado la secuencia de sietes podría aparecer en la parte de la

expansión de π que no ha sido examinada. El procedimiento no parará para las entradas n tal que $g(n) = 0$, así que el procedimiento no es efectivo. (Es concebible que exista un procedimiento efectivo para computar g, basado quizá en alguna propiedad teórica de π).

Este ejemplo pone de manifiesto dos hechos implícitos en la idea de procedimiento efectivo, tal procedimiento es llevado a cabo en una secuencia de etapas o pasos (cada uno completado en tiempo finito), y cualquier salida se obtiene en un número finito de pasos.

Hemos descrito informalmente la idea de algoritmo o procedimiento efectivo y la noción asociada de función computable. Estas nociones pueden ser precisadas en base a una teoría matemática de computabilidad.

1.2. Algunas aproximaciones a la computabilidad

Las siguientes son algunas plausibles caracterizaciones de computabilidad

1. Church (1936). Funciones λ-definibles.
2. Gödel-Kleene (1936). Funciones μ-recursivas y funciones parciales recursivas.
3. Turing (1936). Funciones computables por máquinas finitas conocidas como máquinas de Turing.
4. Shepherdson-Sturgis (1963). URM-funciones computables.

Cada una de estas propuestas para caracterizar la noción de computabilidad efectiva da origen a la misma clase de funciones. Aquí recordaremos algunos conceptos de la teoría de funciones recursivas.

1.3. Funciones Recursivas y Recursivamente Enumerables

1.3.1. Funciones recursivas primitivas

Una función $h : \mathbb{N}^k \to \mathbb{N}$ no necesariamente total (si h no esta definida en todas las k-tuplas de números naturales) se dice recursiva primitiva (**pr**) si puede obtenerse por una sucesión de aplicaciones de las reglas siguientes:

1. Las siguientes funciones llamadas básicas son recursivas primitivas:

$Z(x) = 0$	función constate cero,
$S(x) = x + 1$	función sucesor,
$P_k^n(x_1, \ldots, x_n) = x_k$	k-ésima proyección en n variables.

(Las funciones anteriores son “computables”)

2. *Composición*

   Si $f(y_1, \ldots, y_n)$ y $g_1(x_1, \ldots, x_k), \ldots, g_n(x_1, \ldots, x_k)$ son **p.r.** entonces $h(x_1, \ldots, x_k) = f(g_1(x_1, \ldots, x_k), \ldots, g_n(x_1, \ldots, x_k))$ es **p.r.**

   (Si f y g son "computables", h es "computable").

3. *Recursión Primitiva*

   La recursión es una forma de definir funciones especificando sus valores en términos de valores previamente definidos y posiblemente usando otras funciones ya definidas. Si $f(x_1, \ldots, x_k)$ y $g(x_1, \ldots, x_k, y, z)$ son **p.r.** entonces la función $h(x_1, \ldots, x_k, y)$ definida por

   $$\begin{cases} h(x_1, \ldots, x_k, 0) = f(x_1, \ldots, x_k) \\ h(x_1, \ldots, x_k, S(y)) = g(x_1, \ldots, x_k, y, h(x_1, \ldots, x_k, y)) \end{cases}$$

   es **p.r.**

   (Si f y g son "computables", h es "computable").

   Nota: En el esquema de recurrencia primitiva no es necesario que todas las variables aparezcan en f y g, más aún puede que no aparezca ninguna. La regla 3 contiene el caso particular

   $$\begin{cases} h(0) = n_0 \in \mathbb{N} \\ h(S(y)) = g(y, h(y)). \end{cases}$$

Ejemplo 1.3.1. Las siguientes funciones son recursivas primitivas:

1. La función identidad $I(x) = x = P_1^1(x)$ es **p.r.** por definición.

2. La función constante $C_n(x) = n$ para todo x es **p.r.**, pues $C_0(x) = Z(x)$ y $C_n(x) = S(S(\ldots S(Z(x))\ldots))$ para $n \geq 1$ (regla 2).

3. $Sum(x, y) = x + y$ es **p.r**. Sea $f(x) = P_1^1(x) = x$ y $g(x, y, z) = S(P_3^3(x, y, z)) = S(z)$, entonces

   $$\begin{cases} Sum(x, 0) = x + 0 = x = f(x) \\ Sum(x, S(y)) = g(x, y, x + y), \end{cases}$$

   pues $Sum(x, S(y)) = x + (y + 1)$ y $g(x, y, x + y) = S(x + y) = x + y + 1$.

4. $Prod(x, y) = x \cdot y$ es **p.r**. Sea $f(x) = Z(x)$, $g(x, y, z) = x + z$, entonces

   $$\begin{cases} Prod(x, 0) = x \cdot 0 = 0 = f(x) \\ Prod(x, S(y)) = g(x, y, x \cdot y), \end{cases}$$

   pues $Prod(x, S(y)) = x \cdot (y + 1)$ y $g(x, y, x \cdot y) = x + x \cdot y = x \cdot (y + 1)$

5. $h(x,y) = x^y$ es **pr**. Sea $f(x) = S(Z(x))$, $g(x,y,z) = x \cdot z$, entonces

$$\begin{cases} h(x,0) = x^0 = 1 = f(x) \\ h(x,S(y)) = g(x,y,x^y), \end{cases}$$

pues $h(x,S(y)) = x^{y+1}$ y $g(x,y,x^y) = x \cdot x^y = x^{y+1}$

6. La resta truncada se define

$$x \backsim y = \begin{cases} x - y & \text{si } x \geq y \\ 0 & \text{sino} \end{cases}$$

y es **p.r.** Primero se muestra que $h(x) = x \backsim 1$ es **p.r.**:

$$\begin{cases} h(0) = 0 \\ h(S(x)) = P_1^2(x, h(x)) = x. \end{cases}$$

Sea $f(x) = P_1^1(x)$ y $g(x,y,z) = h(P_3^3(x,y,z))$. Ahora definimos:

$$\begin{cases} H(x,0) = x \backsim 0 = x = f(x) \\ H(x,S(y)) = g(x,y,x \backsim y), \end{cases}$$

pues $H(x,S(y)) = x \backsim (y+1)$ y $g(x,y,x \backsim y) = h(x \backsim y) = (x \backsim y) \backsim 1 = x \backsim (y+1)$.

7. La función $|x-y| = (x \backsim y) + (y \backsim x)$ es **pr**.

8. La función

$$Sg(x) = \begin{cases} 0 & \text{si } x = 0 \\ 1 & \text{sino} \end{cases}$$

es **pr** pues

$$\begin{cases} Sg(0) = 0 \\ Sg(S(x)) = S(Z(Sg(x))) = 1. \end{cases}$$

9. La función

$$\overline{Sg}(x) = \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{sino} \end{cases}$$

es **pr** pues $\overline{Sg}(x) = 1 \backsim Sg(x)$.

10. Sea

$$Car_p(x) = \begin{cases} 1 & \text{si } x \text{ es par} \\ 0 & \text{sino.} \end{cases}$$

Entonces

$$\begin{cases} Car_p(0) = 1 \\ Car_p(S(x)) = 1 \backsim Car_p(x). \end{cases}$$

11. Sea
$$Car_{ab}(x) = \begin{cases} a & \text{si } x \text{ es par} \\ b & \text{sino.} \end{cases}$$
$g(x,y,z) = xz + y(1 \backsim z)$ es **p.r.** pues es composición de funciones **p.r.** $g = Sum\big(Prod(P_1^3, P_3^3), Prod(P_2^3, C_1 \backsim P_3^3)\big)$, entonces $Car_{ab}(x) = g(C_a(x), C_b(x), Car_p(x))$ también lo es.

1.3.2. Funciones recursivamente enumerables

Suponga que $f(x_1, \ldots, x_k, y)$ es una función (no necesariamente total) y que deseamos definir una función $g(x_1, \ldots, x_k)$ por

$$g(x_1, \ldots, x_k) = \text{el menor } y \text{ tal que } f(x_1, \ldots, x_k, y) = 0.$$

(Si f es "computable" entonces g también lo es). Se pueden presentar dos problemas. Primero que para algún $\mathbf{x} = (x_1, \ldots, x_k)$ no exista y tal que $f(\mathbf{x}, y) = 0$. Segundo asumiendo que f es "computable", considere el siguiente algoritmo natural para computar $g(x)$. "Compute $f(\mathbf{x}, 0)$, $f(\mathbf{x}, 1), \ldots$ hasta encontrar y tal que $f(\mathbf{x}, y) = 0$". Este procedimiento podría no terminar si f no es total, aún si tal y existe (por ejemplo, si $f(\mathbf{x}, 0)$ es indefinido pero $f(\mathbf{x}, 1) = 0$).

Para cualquier función $f(\mathbf{x}, y)$ definimos:

$$\mu y(f(\mathbf{x}, y) = 0) = \begin{cases} \text{el menor } y \text{ tal que} \\ \quad (i) \;\; f(\mathbf{x}, z) \text{ esta definido para todo } z \leq y \\ \quad (ii) \;\; f(\mathbf{x}, y) = 0, \text{ si un tal } y \text{ existe} \\ \text{indefinida, si no existe tal } y \end{cases}$$

donde $\mu y \ldots$ significa el "mínimo y tal que ..."

Definición 1.3.2. Una función f es recursivamente enumerable r.e. si es una de las funciones básicas o puede obtenerse de las funciones básicas por un número finito de aplicaciones de composición, recurrencia primitiva y la nueva regla:

4. *Minimización*

   Si $f(\mathbf{x}, y)$ es recursiva y regular para y entonces $g(\mathbf{x}) = \mu y(f(\mathbf{x}, y) = 0)$ es recursiva.

Por definición toda función p.r. es r.e. Existen funciones r.e. que no son p.r., un ejemplo conocido es la función de Ackermann definida por las siguientes ecuaciones

$$\begin{aligned} &\psi(0, y) = y + 1 \\ &\psi(x + 1, 0) = \psi(x, 1) \\ &\psi(x + 1, y + 1) = \psi(x, \psi(x + 1, y)). \end{aligned}$$

La clase $\mathcal{R}$ de funciones parciales recursivas es la clase más pequeña de funciones parciales que puede ser construida a partir de funciones básicas (Z, S, P_k^n) por un número finito de operaciones de composición, recursión y minimización.

A partir de funciones básicas se obtienen usando solo sustitución y recursión funciones computables totales. Un μ-operador puede generar una función computable no total desde una función computable total, por ejemplo tomando $f(x,y) = |x - y^2|$ y $g(x) = \mu y(f(x,y) = 0)$, tenemos que g es una función no total

$$g(x) = \begin{cases} \sqrt{x}, & \text{si } x \text{ es un cuadrado perfecto} \\ \text{indefinida}, & \text{en otro caso} \end{cases}$$

Gödel y Kleene originalmente confinaron su atención a las funciones totales, el operador μ solo era permitido si el resultado era una función total. Hoy en día el término **función recursiva** es usado para funciones r.e totales. El siguiente teorema es un resultado conocido:

Teorema 1.3.3. *Si $f(x)$ es una función r.e. total inyectiva, entonces f^{-1} es r.e.*

La importancia de la definición de r.e. es que trata de capturar el concepto de computabilidad. Una relación n-aria en los naturales es computable si existe un procedimiento efectivo (formalmente función computable, informalmente algoritmo) que dada una n-tupla arbitraria produzca la salida 1 o 0 dependiendo si la n-tupla pertenece o no a la relación. Diremos que es semi-computable si existe un procedimiento efectivo tal que dada una n-tupla que pertenezca a la relación, éste eventualmente lleve a la salida 1, y que cuando la n-tupla no pertenezca a la relación no lleve a la salida 1 (observe que no se exige que el procedimiento lleve a la salida 0 en éste caso). Una relación R es computable si y sólo si R y su complemento son semicomputables. Así formalmente decimos que una relación R es recursiva si ésta y su complemento son r.e.

1.4. Conjuntos numerables

Definición 1.4.1. Un conjunto X es numerable si existe una biyección $f : X \to \mathbb{N}$

El término contable es usado para conjuntos finitos o numerables, es decir, para conjuntos infinitos, contable significa lo mismo que numerable.

Definición 1.4.2. Una numeración de X es una sobreyección $g : \mathbb{N} \to X$; esta es frecuentemente representada por $X = \{x_0, x_1, x_2, \ldots\}$, donde $x_n = g(n)$. Esta es una numeración sin repeticiones si g es inyectiva.

Sea X un conjunto de objetos finitos (por ejemplo un conjunto de enteros, o un conjunto de palabras, o un conjunto de instrucciones, o un conjunto de programas); entonces X es efectivamente numerable si existe una numeración $f : \mathbb{N} \to X$ "efectivamente computable" (existe un algoritmo que genera una sucesión de elementos de X en la cual aparece

eventualmente todo elemento de X). Esta noción no es formal, cuando necesitamos una definición formal aplicamos:

Definición 1.4.3. Un conjunto X es efectivamente numerable si $X = \emptyset$ ó existe una función biyectiva $f : \mathbb{N} \to X$ tal que f y f^{-1} son r.e.

Teorema 1.4.4. *Los siguientes conjuntos son efectivamente numerables*

1. $\mathbb{N} \times \mathbb{N}$
2. $\mathbb{N}^+ \times \mathbb{N}^+ \times \mathbb{N}^+$
3. $\bigcup_{k \geq 0} \mathbb{N}^k$ *el conjunto de todas las secuencias finitas de números naturales*

Demostración.

1. $\pi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ definida por $\pi(m, n) = 2^m(2n + 1) - 1 = 2^m(2n + 1) \backsim 1$ claramente es recursiva pues es composición de funciones p.r.

   La función π es sobre pues cualquier número $x + 1$ en caso de ser impar se puede escribir de la forma $2y+1$ tomando $x = 0$ y si es par se puede dividir entre dos hasta obtener un número impar. Es uno a uno pues $2^{x_1}(2y_1 + 1) = 2^{x_2}(2y_2 + 1)$ implica $x_1 = x_2$, luego $y_1 = y_2$. Como la función es biyectiva existe $\pi^{-1}(x) = (\pi_1(x), \pi_2(x))$ y dado que π es total y computable por el Teorema 1.3.3. π^{-1} también es computable.

2. Una biyección explícita $\xi : \mathbb{N}^+ \times \mathbb{N}^+ \times \mathbb{N}^+ \to \mathbb{N}$ utilizando la función π esta dada por $\xi(m, n, q) = \pi(\pi(m - 1, n - 1), q - 1)$

   Entonces

   $\xi^{-1}(x) = (\pi_1(\pi_1(x)) + 1, \pi_2(\pi_1(x)) + 1, \pi_2(x) + 1)$

   Dado que π, π_1 y π_2 son efectivamente computables, ξ y ξ^{-1} también lo son.

3. Una biyección $\tau : \bigcup_{k>0} \mathbb{N}^k \to \mathbb{N}$ esta definida por

   $$\tau(a_1, \ldots, a_k) = 2^{a_1} + 2^{a_1+a_2+1} + 2^{a_1+a_2+a_3+2} + \ldots + 2^{a_1+a_2+\ldots+a_k+k-1} - 1$$

   Claramente τ es recursiva. Para ver que τ es una biyección se usa el hecho que un número natural tiene una única expresión binaria. Esto es dado x podemos encontrar "efectivamente" números únicos $k \geq 1$ y $0 \leq b_1 < b_2 \ldots < b_k$ tales que

   $$x + 1 = 2^{b_1} + 2^{b_2} + \ldots + 2^{b_k}$$

   de donde se obtiene $\tau^{-1}(x) = (a_1, a_2, \ldots, a_n)$ con $a_1 = b_1$ y $a_{i+1} = b_{i+1} - b_i - 1$ $(1 \leq i < k)$. Dado que τ es inyectiva, total y computable por el Teorema 1.3.3. τ^{-1} también es computable.

$\square$

La noción de computabilidad en $\mathbb{N}$, podemos extenderla a otros conjuntos usando el recurso de codificación. Una codificación de un conjunto D es una numeración efectiva $\alpha : D \to \mathbb{N}$. Decimos que un objeto $d \in D$ es codificado por el número natural $\alpha(d)$. Suponga que $f : D \to D$ es codificada naturalmente por la función $f^* : \mathbb{N} \to \mathbb{N}$ que envía el código de un objeto $d \in \mathrm{Dom}(f)$ en el código de $f(d)$. Explícitamente tenemos

$$f^* = \alpha \circ f \circ \alpha^{-1}$$

$$\begin{tikzcd}
\mathbb{N} \arrow[r, "\alpha^{-1}"] \arrow[d, "f^*"'] & D \arrow[d, "f"] \\
\mathbb{N} & D \arrow[l, "\alpha"]
\end{tikzcd}$$

Ejemplo 1.4.5. Considere el conjunto $\mathbb{Z}$. Un código explícito esta dado por la función α donde

$$\alpha(n) = \begin{cases} 2n & \text{si } n \geq 0 \\ -2n-1 & \text{si } n < 0 \end{cases}$$

Entonces α^{-1} esta dada por

$$\alpha^{-1}(n) = \begin{cases} \frac{1}{2}n & \text{si } n \text{ es par} \\ -\frac{1}{2}(n+1) & \text{si } n \text{ es impar} \end{cases}$$

Considere $f(x) = x - 1$ en $\mathbb{Z}$, entonces la función f^* en $\mathbb{N}$ esta dada por

$$f^*(x) = \begin{cases} 1 & \text{si } x = 0 \\ x-2 & \text{si } x > 0 \text{ y } x \text{ es par} \\ x+2 & \text{si } x \text{ es impar} \end{cases}$$

f^* es **p.r.** En efecto, sea $g(x,y) = \big(Car_{40}(P_1^2(x,y)) \backsim (1 \backsim P_1^2(x,y))\big) + (P_1^2(x,y) \backsim 1)$, entonces

$$\begin{cases} f^*(0) = 1 \\ f^*(S(x)) = g(x, f^*(x)) \end{cases}$$

Entonces $f(x) = x - 1$ es una función computable en $\mathbb{Z}$.

En contraste podemos dar la definición de computabilidad en un nuevo conjunto directamente en términos de sus objetos y su estructura intrínseca.

Ejemplo 1.4.6. Sea $D = \Sigma^*$, donde $\Sigma = \{a, b\}$. La clase $\mathcal{R}^D$ de funciones parciales recursivamente enumerables en Σ^* es la clase más pequeña de funciones parciales tal que

1. Las funciones básicas

$$\begin{aligned} f(\sigma) &= \lambda \\ f(\sigma) &= \sigma a \\ f(\sigma) &= \sigma b \\ P_k^n(\sigma_1, \ldots, \sigma_n) &= \sigma_k \end{aligned}$$

   están en $\mathcal{R}^D$

2. $\mathcal{R}^D$ es cerrado bajo composición

3. $\mathcal{R}^D$ es cerrado bajo recursión primitiva de la siguiente forma:

$$\begin{aligned} h(\boldsymbol{\sigma}, \lambda) &= f(\boldsymbol{\sigma}) \\ h(\boldsymbol{\sigma}, \tau a) &= g_1(\boldsymbol{\sigma}, \tau, h(\boldsymbol{\sigma}, \tau)) \\ h(\boldsymbol{\sigma}, \tau b) &= g_2(\boldsymbol{\sigma}, \tau, h(\boldsymbol{\sigma}, \tau)) \end{aligned}$$

   donde f, g_1, $g_2 \in \mathcal{R}^D$

4. $\mathcal{R}^D$ es cerrado bajo minimización: si $f(\boldsymbol{\sigma}, \tau)$ esta en $\mathcal{R}^D$ también esta la función h dada por

$$h(\boldsymbol{\sigma}) = \mu\tau(f(\boldsymbol{\sigma}, \tau) = \lambda)$$

   donde $\mu\tau$ significa el primer τ en el orden lexicográfico

$$\lambda,\ a,\ b,\ aa,\ ab,\ ba,\ bb,\ aaa,\ aab,\ aba, \ldots$$

Es un resultado conocido que el conjunto de todas las funciones r.e. es efectivamente numerable. obteniendo códigos ó indices para funciones r.e. Por otro lado es fácil ver que las funciones con dominio $\mathbb{N}$ y recorrido $\{0, 1\}$ (funciones características en $\mathbb{N}$) tienen correspondencia biunívoca con el conjunto de partes de $\mathbb{N}$. Esto implica que hay muchas funciones que no son "computables".

1.5. Predicados decidibles y parcialmente decidibles

Suponga que $M(x_1, x_2, \ldots, x_n)$ es un predicado n-ario de números naturales y c_M su función característica:

$$c_M(x) = \begin{cases} 1, & \text{si } M(x) \text{ se tiene} \\ 0, & \text{si } M(x) \text{ no se tiene.} \end{cases}$$

Definición 1.5.1. El predicado $M(x)$ es decidible si la función c_M es computable.

Proposición 1.5.2. *Suponga que $R(\mathbf{x}, y)$ es un predicado decidible; entonces la función*

$$g(\mathbf{x}) = \mu y R(\mathbf{x}, y) = \begin{cases} \textit{el menor } y \textit{ tal que } R(\mathbf{x}, y) \textit{ se tiene}, & \textit{si tal } y \textit{ existe} \\ \textit{indefinida}, & \textit{en otro caso} \end{cases}$$

es r.e.

Demostración. $g(\mathbf{x}) = \mu y(\overline{Sg}(c_R(\mathbf{x}, y)) = 0)$ □

Definición 1.5.3. Un predicado $M(x)$ de los números naturales es parcialmente decidible si la función

$$c_M(x) = \begin{cases} 1, & \text{si } M(x) \text{ se tiene} \\ \text{indefinida}, & \text{en otro caso} \end{cases}$$

es r.e.

En este caso un algoritmo para c_M es un procedimiento que da la respuesta si, cuando $M(x)$ se satisface, y no da respuesta cuando no se tiene $M(x)$.

Nota: En la literatura los términos semi-computable, recursivamente enumerable son usados en el mismo sentido que parcialmente decidible.

1.6. Conjuntos Recursivos y Recursivamente Enumerables

Sea A un subconjunto de los números naturales y c_A su función característica:

$$c_A(x) = \begin{cases} 1, & \text{si } x \in A \\ \text{indefinido }, & \text{si } x \notin A \end{cases}$$

Definición 1.6.1. Un conjunto A es r.e si la función c_A es r.e.

Decimos que un conjunto $X \subseteq \mathbb{N}^k$ es recursivo si tanto él como su complemento son r.e. Por definición todo conjunto recursivo es r.e. Ahora, existen conjuntos r.e. que no son recursivos? La respuesta es si. El problema de la parada consiste en determinar dado un programa con código i y una entrada x para este programa, si el programa termina o no. El conjunto $K = \{(i, x) \in \mathbb{N}^2 \mid \text{el programa } i \text{ terminará eventualmente con la entrada } x\}$ es r.e., pero su complemento no. Se utiliza un argumento diagonal para ver que el complemento de K no es r.e.

1.7. Hasta ahora qué podemos decir de la computabilidad en $\mathbb{R}$?

Podemos extender la teoría de la computabilidad de los números naturales a los números reales de la siguiente manera: un número real es computable si existen procedimientos efectivos que producen su expansión decimal. Ahora bien sabemos que existen procedimientos efectivos para generar correctamente el decimal n-ésimo de números irracionales tales como $\sqrt{2}, \pi, e$, así que éstos números son computables. Existen números reales que no son computables, un ejemplo es la constante de Chaitin Ω que se define de la siguiente manera:

Sea P el conjunto de todos los programas que se detienen, y $|p|$ el número de bits de un programa p,

$$\Omega = \sum_{p \in P} 2^{-|p|}.$$

Para esta constante es posible conocer los primeros dígitos, pero a partir de cierto dígito decimal (que depende de la codificación elegida) no es posible obtener más.

Tomando la máquina de Turing como aproximación a la computabilidad, digamos que un número real es computable si existe una máquina de Turing que puede producir su expansión decimal.

Además una función $f : \mathbb{R}^n \to \mathbb{R}$ es computable , si y sólo si ésta puede ser evaluada en un computador con precisión arbitraria. Generalizando la idea de función computable en los naturales, diremos que una máquina de Turing M computa una función $f \subseteq \Sigma^\omega \to \Sigma^\omega$ si se tiene que

1. Si $p \in$dom(f), M computa la salida infinitamente larga y la escribe en la cinta de M.

2. Si $p \notin$dom(f), M no produce una salida.

Considere la función $f(x) = 3x$. Si la entrada x tiene prefijo 0.3333...3, la salida tiene prefijo 0.9999...9 lo cual es correcto si el siguiente dígito es menor o igual que 3, pero si el siguiente dígito es mayor que 3:

.	3	3	3	3	3	$\cdots$	4

.	9	9	9	9	9	$\cdots$	?

no se obtiene el prefijo correcto para este tipo de entradas. Aquí la multiplicación no es computable, así que el modelo no es apropiado.

Existen diferentes modelos donde la multiplicación en los reales es computable, en los siguientes capítulos trataremos dos modelos basados en teoría de dominios. Para introducir la teoría de dominios es necesario recordar algunas definiciones básicas de conjuntos ordenados. La teoría de dominios tiene aplicaciones a muchos campos, pero en éste trabajo nos enfocaremos en la teoría de dominios como una teoría de la aproximación, con miras a trabajar sobre una estructura de aproximación de los números reales.

1.8. Conjuntos Ordenados

Definición 1.8.1. Una relación $\propto$ que satisface las propiedades reflexiva, antisimétrica y transitiva se denomina relación de orden. Si $\propto$ solo satisface las propiedades transitiva y reflexiva se denomina relación de preorden. Un conjunto S dotado de una relación de orden $\propto$ se denomina conjunto parcialmente ordenado y dotado de una relación de preorden se denomina conjunto preordenado. Tales conjuntos se denotan $(S, \propto)$.

Definición 1.8.2. Una relación de orden $\propto$ sobre un conjunto A recibe el nombre de relación de orden total, si cualesquiera que sean los elementos a y b de A, a y b son comparables (se dice que a y b son comparables si $a \leq b \quad \vee \quad b \leq a$). Se dice que entonces que $(A, \propto)$ es un conjunto totalmente ordenado, linealmente ordenado o cadena.

Definición 1.8.3. Sea P un conjunto ordenado y sea $S \subseteq P$. Un elemento $x \in P$ es una cota superior de S si $s \leq x$ para todo $s \in S$. El conjunto de cotas superiores de S es denotado por $S^{\uparrow}$.

Dualmente, un elemento $y \in P$ es cota inferior de S si $y \leq s$ para todo $s \in S$. El conjunto de cotas inferiores de S es denotado por $S^{\downarrow}$.

Si $S^{\uparrow}$ tiene un menor elemento x (la menor de las cotas superiores), entonces x se denomina extremo superior de S y se denota sup S ó $\bigvee S$. Si existe en $S^{\downarrow}$ un mayor elemento (la mayor de las cotas inferiores), entonces y se denomina extremo inferior de S y se denota inf S ó $\bigwedge S$.

El siguiente lema que se puede encontrar como el lema 2.28 en [7] es un resultado que utilizaremos en el siguiente capítulo.

Lema 1.8.4. *Sea Q un subconjunto, con el orden inducido de algún conjunto ordenado P, y sea $S \subseteq Q$. Si $\bigvee P$ existe y pertenece a Q, entonces $\bigvee_Q S$ y es igual a $\bigvee_P S$ (y dualmente para $\bigwedge_Q S$).*

1.8.1. Ordenes de Información

Los conjuntos ordenados permiten representar la *cantidad de información* contenida en una expresión. $x \sqsubseteq y$ significa que que x posee igual o menos información que y.

Ejemplo 1.8.5. Sean $\Sigma = \{0, 1\}$ alfabeto binario.

Σ^* denota cadenas finitas de $0's$ y $1's$.

Σ^{**} denota cadenas finitas o infinitas de $0's$ y $1's$.

λ denota la cadena vacía.

$x \sqsubseteq y$ se define como x es un prefijo de y, por ejemplo:

$$101 \sqsubseteq 10100$$
$$010101 \sqsubseteq 0101010101010101010101...$$

$(\Sigma^{**}, \sqsubseteq)$ es un conjunto ordenado.

Ejemplo 1.8.6. Sea $(X \multimap\!\!\to Y)$ el conjunto de funciones parciales de X en Y (funciones cuyo dominio es un subconjunto de X y codominio es Y)

$\sigma \sqsubseteq \tau$ si y sólo si $\mathrm{dom}(\sigma) \subseteq \mathrm{dom}(\tau)$ y para todo $x \in \mathrm{dom}(\sigma)$, $\sigma(x) = \tau(x)$

$\big((X \multimap\!\!\to Y), \sqsubseteq\big)$ es un conjunto ordenado.

Ejemplo 1.8.7. Sea $\mathbb{IR}$ el conjunto de todos los intervalos $[\underline{x}, \overline{x}]$ en $\mathbb{R}^*$ los reales extendidos, donde

$$-\infty \leq \underline{x} \leq \overline{x} \leq \infty.$$

Sean $\mathbf{x} = [\underline{x}, \overline{x}]$ y $\mathbf{y} = [\underline{y}, \overline{y}]$. Se define $\mathbf{x} \sqsubseteq \mathbf{y}$ si y sólo si $\underline{x} \leq \underline{y}$ y $\overline{y} \leq \overline{x}$ con el orden usual de los números reales, es decir, el intervalo $[\underline{x}, \overline{x}]$ contiene al intervalo $[\underline{y}, \overline{y}]$.

$(\mathbb{IR}, \sqsubseteq)$ es un conjunto ordenado.

CAPÍTULO 2

Teoría de Dominios

2.1. Dominios Continuos y Dominios Algebraicos

Definición 2.1.1. Sea $P = (P, \sqsubseteq)$ un preorden, un subconjunto $D \subseteq P$ es dirigido si $D \neq \emptyset$ y siempre que $x, y \in D$ existe $z \in D$ tal que $x \sqsubseteq z$ y $y \sqsubseteq z$.

Definición 2.1.2. $(P, \sqsubseteq, \perp)$ es un orden parcial completo para dirigidos dcpo (directed-complete partial order), si P tiene elemento mínimo $\perp$ y cualquier conjunto dirigido $D \subseteq P$ tiene supremo.

El supremo de un conjunto D se denota $\bigvee D$. Cuando el conjunto es dirigido de denota $\bigsqcup D$.

Ejemplo 2.1.3. $(\Sigma^{**}, \sqsubseteq)$ es un dcpo.

Se caracterizan primero los conjuntos dirigidos. Sea $D = \{x_i\}_{i \in I}$ dirigido, es decir, para todo $x_i, x_j \in D$ existe $x_k \in D$ tal que x_i prefijo de x_k y x_j prefijo de x_k.

Si alguna de las cadenas x_i o x_j es una cadena infinita (elemento maximal), sin perdida de generalidad suponga que x_j es infinita, entonces $x_k = x_j$ y por tanto x_i es prefijo de x_j, es decir, x_i y x_j son comparables. Note que D puede tener a lo más una cadena infinita.

Si x_i y x_j son cadenas finitas

$$x_i = a_{i_1} a_{i_2} \dots a_{i_n}$$
$$x_j = a_{j_1} a_{j_2} \dots a_{j_m},$$

$x_k \sqsupseteq x_i$ implica $a_{i_1} = a_{k_1}$, $a_{i_2} = a_{k_2}$, $a_{i_n} = a_{k_n}$ y $x_k \sqsupseteq x_j$ implica $a_{j_1} = a_{k_1}$, $a_{j_2} = a_{k_2}$, $a_{j_m} = a_{k_m}$. Si $m \leq n$, $x_j \sqsubseteq x_i$; si $n \leq m$, $x_i \sqsubseteq x_j$. Luego x_i y x_j son comparables.

En cualquier caso D es totalmente ordenado, es decir, cualquier par de elementos de D es comparable. Ahora se hallará el supremo de un conjunto dirigido.

Si D contiene una cadena infinita x_j entonces $x_j = \bigsqcup D$ pues para cualquier $x_i \in D$, $x_j \sqsupseteq x_i$ como se mostró antes.

Si D no contiene cadenas infinitas $D = \{x_i\}_{i\in I}$ donde x_i cadena finita para todo $i \in I$ entonces $D \subseteq \Sigma^*$, es decir, D es finito ó infinito enumerable. Si D es finito claramente tiene máximo. Si D es infinito enumerable $D = \{x_i\}_{i\in\mathbb{N}}$, ordenando D de menor a mayor $D = \{x_1, x_2, \ldots\}$ con $x_1 \sqsubseteq x_2 \sqsubseteq \cdots$ se define $x = \bigsqcup D$ como sigue:

$$x = a_1 a_2 \ldots$$

donde a_i es la i-ésima componente de x_i. Note que no importa la longitud de la cadena x_i, dado que D es infinito se garantiza que $x \sqsupseteq x_i$ $\forall i \in \mathbb{N}$, con x cadena infinita.

Ejemplo 2.1.4. $((X \multimap Y), \sqsubseteq)$ es un dcpo.

Sea $\mathcal{H} = \{f_i\}_{i\in I}$ dirigido, entonces para todo f_i, $f_j \in \mathcal{H}$, existe $f_k \in \mathcal{H}$ tal que $f_k \sqsupseteq f_i$ y $f_k \sqsupseteq f_j$, es decir

$$\begin{aligned} &\text{Dom } f_i \subseteq \text{ Dom } f_k \text{ y } f_i(x) = f_k(x)\ \forall x \in \text{ Dom } f_i \\ &\text{Dom } f_j \subseteq \text{ Dom } f_k \text{ y } f_j(x) = f_k(x)\ \forall x \in \text{ Dom } f_j \end{aligned}$$

Así $f_i(x) = f_j(x), \forall x \in$ Dom $f_i \cap$ Dom f_j es condición necesaria para la existencia de f_k.

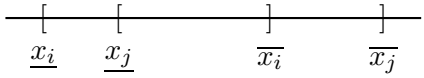
$((X \multimap Y), \sqsubseteq)$ es un sub-dcpo de $(\wp(X \times Y), \subseteq)$, y como $\bigcup \mathcal{H} \in (X \multimap Y)$ entonces por el Lema1.8.4 $\bigsqcup \mathcal{H} = \bigcup \mathcal{H}$.

Veamos que en efecto $\bigcup \mathcal{H}$ es una función. Si (x, y) y $(x, z) \in \bigcup \mathcal{H}$, $y = f_i(x)$ y $z = f_j(x)$ para algunos $i, j \in I$, con $x \in$ Dom$f_i \cap$ Domf_j. Luego $f_i(x) = f_j(x)$

Ejemplo 2.1.5. $(\mathbb{IR}, \sqsubseteq)$ es un dcpo.

Sea $D = \{x_i\}_{i\in I} = \{[\underline{x_i}, \overline{x_i}]\}_{i\in I}$ dirigido. Una condición necesaria para la existencia de x_p tal que $x_p \sqsupseteq x_i$ y $x_p \sqsupseteq x_j$ es

$$\underline{x_j} \leq \overline{x_i} \text{ para todo } i, j \in I \tag{2.1}$$



Sean $\underline{D} = \{\underline{x_i}\}$ y $\overline{D} = \{\overline{x_i}\}$ subconjuntos de $\mathbb{R}^*$. Como $D \neq \emptyset$ entonces $\underline{D} \neq \emptyset$, además $\underline{D}$ es superiormente acotado por 2.1, luego $\bigvee \underline{D}$ existe en $\mathbb{R}^*$. Análogamente $\bigwedge \overline{D}$ existe en $\mathbb{R}^*$.

Afirmamos que $\bigsqcup D = [\bigvee \underline{D}, \bigwedge \overline{D}]$. Primero para ver que $\bigsqcup D$ esta bien definido se demostrará que $\bigvee \underline{D} \leq \bigwedge \overline{D}$. Por contradicción, suponga que no, es decir, suponga que $\bigwedge \overline{D} < \bigvee \underline{D}$, entonces existen $\underline{x_j} \in \underline{D}$, $\overline{x_k} \in \overline{D}$ tales que $\bigwedge \overline{D} < \underline{x_j} \leq \bigvee \underline{D}$ y $\bigwedge \overline{D} \leq \overline{x_k} < \underline{x_j}$, esta última desigualdad contradice (1).

$[\bigvee \underline{D}, \bigwedge \overline{D}]$ es cota superior pues claramente $[\bigvee \underline{D}, \bigwedge \overline{D}] \sqsubseteq [\underline{x_i}, \overline{x_i}]$, $\forall i \in I$. Ahora suponga que $[\underline{y}, \overline{y}]$ es cota superior de D, es decir, $[\underline{y}, \overline{y}] \sqsupseteq [\underline{x_i}, \overline{x_i}]$, $\forall i \in I$ entonces

$$\underline{y} \geq \underline{x_i}\ \forall i \in I \quad \text{y} \quad \overline{y} \leq \overline{x_i}\ \forall i \in I$$

$$\underline{y} \geq \bigvee_{i\in I} \underline{x_i} \quad \text{y} \quad \overline{y} \leq \bigwedge_{i\in I} \overline{x_i}$$

$$\underline{y} \geq \bigvee \underline{D} \quad \text{y} \quad \overline{y} \leq \bigwedge \overline{D}.$$

Luego $[\bigvee \underline{D}, \bigwedge \overline{D}] \sqsubseteq [\underline{y}, \overline{y}]$. Así $[\bigvee \underline{D}, \bigwedge \overline{D}]$ es la menor cota superior.

Definición 2.1.6. Si $(P, \sqsubseteq)$ es un dcpo, y x, $y \in P$, entonces decimos que x aproxima a y, y escribimos $x \ll y$ si para todo subconjunto dirigido $D \subseteq P$ con $y \sqsubseteq \bigsqcup D$ existe algún $d \in D$ con $x \sqsubseteq d$. Decimos que x es compacto si se aproxima a sí mismo.

Nota 1. Observe la analogía entre la definición de elemento compacto en un dcpo y conjunto compacto en una topología

Tomando $D = \{y\}$ en la definición se deduce

$$x \ll y \Rightarrow x \sqsubseteq y. \tag{2.2}$$

También se tiene directamente de la definición

$$x' \sqsubseteq x \ll y \sqsubseteq y' \Rightarrow x' \ll y'. \tag{2.3}$$

De 2.3 se deduce que si x es compacto, aproxima todo elemento por encima de el.

La propiedad más importante del orden de aproximación $\ll$ en un dominio continuo es la *propiedad de interpolación*: Si $x \ll y$ entonces existe $z \in P$ tal que $x \ll z$ y $z \ll y$. (Para obtener este resultado tome en la definición $z = d$). En general

$$\text{Si } x_i \ll y \text{ para } i = 1 \ldots n \Rightarrow \exists z \in P \text{ con } x_i \ll z \text{ y } z \ll y. \tag{2.4}$$

El significado intuitivo de $x \ll y$ es que "la información contenida en x es esencial para y". Es frecuente referirse a esta relación como "x esta muy por debajo de y".

Introducimos la siguiente notación

$$\begin{aligned}
\downdownarrows x &= \{y \in P \mid y \ll x\} \\
\upuparrows x &= \{y \in P \mid x \ll y\} \\
\upuparrows A &= \bigcup_{a\in A} \upuparrows a \\
K(P) &= \{x \in P \mid \text{x compacto }\}.
\end{aligned}$$

Definición 2.1.7. Sean P y Q dcpo's. Una función $f : P \to Q$ es continua si es monótona y

$$f\big(\bigsqcup D\big) = \bigsqcup f(A)$$

para todo D subconjunto dirigido de P.

No toda función monótona es continua. Por ejemplo $f : \wp(\mathbb{N}) \to \wp(\mathbb{N})$, $f(S) = \begin{cases} \emptyset, \text{ si } S \text{ es finito} \\ \mathbb{N}, \text{ si } S \text{ es infinito.} \end{cases}$

Tomando $D = \{S \subseteq \mathbb{N},\ S \text{ finito}\}$ se tiene $f\big(\bigsqcup D\big) = \mathbb{N} \not\subseteq \bigsqcup f(D) = \emptyset$.

Ejemplo 2.1.8. Los elementos compactos de $(\Sigma^{**}, \sqsubseteq)$ son exactamente las cadenas de longitud finita:

Sea k compacto tal que $k \sqsubseteq \bigsqcup D$. Se recuerdan algunas observaciones importantes del Ejemplo 2.1.3

D es dirigido, si y sólo, si D es totalmente ordenado.

Si D es dirigido, entonces D tiene a lo más una cadena infinita.

Caso 1. Si D tiene una cadena infinita x_j, $\bigsqcup D = x_j$, así $k \sqsubseteq x_j$. Se construye $\overline{D}$ dirigido tal que $\bigsqcup \overline{D} = x_j$ y $\overline{D}$ no tenga cadenas infinitas. Sea $x_j = a_1 a_2 \ldots$

$$\overline{D} = \{x_1, x_2, \ldots\} \text{ donde } x_n = a_1 a_2 \ldots a_n \forall n \in \mathbb{N}.$$

Luego k compacto y $k \sqsubseteq \bigsqcup \overline{D}$, entonces $k \sqsubseteq \overline{d}$ para algún $\overline{d} \in \overline{D}$, entonces k cadena finita.

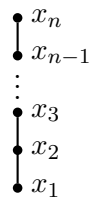
Caso 2. Si D no contiene cadenas infinitas entonces $k \sqsubseteq d$ para algún $d \in D$, entonces k cadena finita.

Sea k cadena finita y $k \sqsubseteq \bigsqcup D$. Se recuerdan algunas observaciones del ejemplo 2.1.3.

Si D es dirigido, entonces D es totalmente ordenado y es finito ó infinito enumerable.

Caso 1. Si D es finito

$\bigsqcup D = x_n$. Luego $k \sqsubseteq x_n \in D$, es decir, k es compacto.



Caso 2. Si D infinito, entonces $\bigsqcup D$ cadena infinita. Como k es cadena finita $k = a_1 a_2 \ldots a_m$ y dado que $k \sqsubseteq \bigsqcup D$, $\bigsqcup D = a_1 a_2 \ldots a_m \ldots$ Si se consideran los primeros $m+1$ elementos de D ordenados de menor a mayor $D = \{x_1, x_2, \ldots, x_{m+1}\}$, se tiene que x_{m+1} es una cadena finita que tiene por lo menos m componentes, entonces $k \sqsubseteq x_{m+1} \in D$, es decir, k es compacto.

Ejemplo 2.1.9. Los elementos compactos de $((X \multimap Y), \sqsubseteq)$ son exactamente las funciones parciales cuyo dominio es finito.

Sean g compacto y $D = \{f_i : \text{Dom } f_i \Subset \text{Dom } g \text{ y } f_i(x) = g(x)\ \forall x \in \text{Dom } g\}$, D es dirigido y $g = \bigsqcup D$, como g compacto $g \sqsubseteq f_i$ para algún $f_i \in D$, entonces Dom g es finito.

Sea $g \in ((X \multimap Y), \sqsubseteq)$ tal que Dom g finito y $g \sqsubseteq f = \bigsqcup\{f_i\}_{i\in I}$ donde $f : \bigcup D_i \to Y$ y $f(x) = f_i(x)\ \forall x \in D_i$ (D_i denota el dominio de f_i, es decir, $f_i : D_i \to Y$, $D_i \subseteq X$), entonces Dom $g \Subset \bigcup D_i$. Se observa que $\{D_i\}_{i\in I}$ es dirigido en $(\wp(X), \subseteq)$. Luego Dom $g \subseteq D_i$ para algún $i \in I$ pues los elementos compactos en $(\wp(X), \subseteq)$ son exactamente los conjuntos finitos. Así $g \sqsubseteq f_i : D_i \to Y$, pues Dom $g \subseteq D_i$ y $g(x) = f_i(x)\ \forall x \in$ Dom g.

Ejemplo 2.1.10. El único elemento compacto de $(\mathbb{IR}, \sqsubseteq)$ es $[-\infty, \infty]$

Sea $[\underline{x}, \overline{x}] \in \mathbb{IR}$ tal que $[\underline{x}, \overline{x}] \neq [-\infty, \infty]$, $D = \{[\underline{x} - \frac{1}{n}, \overline{x} + \frac{1}{n}]\}_{n\in\mathbb{N}}$ es dirigido

$\bigsqcup D = [\underline{x}, \overline{x}]$, pero $[\underline{x}, \overline{x}] \not\sqsubseteq [\underline{x} - \frac{1}{n}, \overline{x} + \frac{1}{n}]$, $\forall n \geq 1$, entonces $[\underline{x}, \overline{x}]$ no es compacto.

Ejemplo 2.1.11. Sean P y Q dcpos, se define

$$[P \to Q] = \{f : f \text{ es función continua de } P \text{ en } Q\}.$$

Los elementos compactos de $[P \to Q]$ son las funciones

$$\begin{aligned} f_{ab} :& P \to Q \\ & x \mapsto f_{ab}(x) = \begin{cases} b \text{ si } x \sqsupseteq a \\ \bot \text{ en caso contrario,} \end{cases} \end{aligned}$$

donde $a \in K(P)$ y $b \in K(Q)$.

Primero se mostrara que f_{ab} es continua. Sea D subconjunto dirigido de P.

Si $\bigsqcup D \sqsupseteq a$, entonces $f_{ab}(\bigsqcup D) = b$. Como D dirigido y a compacto, existe $x \in D$ tal que $a \sqsubseteq x$, entonces $f_{a,b}(x) = b$, es decir, $b \in f_{ab}(D)$. Además b es cota superior de $f_{ab}(D)$, luego $\bigsqcup f_{ab}(D) = b$. Así se concluye $f_{ab}(\bigsqcup D) = b = \bigsqcup f_{ab}(D)$.

Si $\bigsqcup D \not\sqsupseteq a$ entonces $f_{ab}(\bigsqcup D) = \bot$. $\bigsqcup D \not\geq a$ implica $\forall x \in D,\ x \not\geq a$ (contrareciproco de $\exists x \in D$ tal que $x \sqsupseteq a$ implica $\bigsqcup D \sqsubseteq a$), entonces $f_{ab}(x) = \bot$, $\forall x \in D$, es decir, $\bigsqcup f_{ab}(D) = \bot$. Por tanto, $f_{ab}(\bigsqcup D) = \bot = \bigsqcup f_{ab}(D)$.

Ahora se mostrara que f_{ab} es elemento compacto de $[P \to Q]$. Sea $\{\varphi_i\}_{i\in I}$ familia dirigida en $[P \to Q]$ tal que $f_{ab} \sqsubseteq \bigsqcup_{i\in I} \varphi_i$, entonces

$$\begin{aligned} & f_{ab}(a) \sqsubseteq \Big(\bigsqcup_{i\in I} \varphi_i\Big)(a) \\ \Rightarrow & b \sqsubseteq \Big(\bigsqcup_{i\in I} \varphi_i\Big)(a) \\ \Rightarrow & \exists j \in I \text{ tal que } b \sqsubseteq \varphi_j(a) \text{ (porque } b \text{ es compacto)} \end{aligned}$$

Afirmamos que $f_{ab} \sqsubseteq \varphi_j$. En efecto, sea $x \sqsupseteq a$, entonces $f_{ab}(x) = b \sqsubseteq \varphi_j(a) \sqsubseteq \varphi_j(x)$ (pues φ_j es monótona). De esta manera se concluye f_{ab} es elemento compacto de $[P \to Q]$.

Recíprocamente, todo elemento compacto de $[P \to Q]$ es de la forma f_{ab}. Veamos que

$$f = \bigsqcup\{f_{ab} \mid f_{ab} \sqsubseteq f\}$$

Es inmediato que f es cota superior de $\{f_{ab} \mid f_{ab} \sqsubseteq f\}$, hay que probar que es la menor de las cotas superiores. Sean $x \in P$ y $E = \{b' \in K(Q) \mid b' \sqsubseteq f(x)\}$, entonces $f(x) = \bigsqcup E$. Sea $b' \in E$, entonces $b' \sqsubseteq f(x) = f(\bigsqcup\{a' \in K(P) \mid a' \sqsubseteq x\}) = \bigsqcup\{f(a') \mid a' \in K(P), a' \sqsubseteq x\}$ pues f es continua. Como b' es elemento compacto de Q entonces $b' \sqsubseteq f(a')$ para algún $a' \in K(P),\ a' \sqsubseteq x$. Luego $b' = f_{a'b'}(x) \sqsubseteq \bigsqcup_{b \sqsubseteq f(a)} f_{ab}(x)$, es decir, $\bigsqcup_{b \sqsubseteq f(a)} f_{ab}(x)$ es cota superior de E. $\bigsqcup E \sqsubseteq \bigsqcup_{b \sqsubseteq f(a)} f_{ab}(x)$, entonces $\forall x \in P\ f(x) = \bigsqcup E \sqsubseteq \bigsqcup_{b \sqsubseteq f(a)} f_{ab}(x) = \bigsqcup\{f_{ab}(x) \mid f_{ab} \sqsubseteq f\}$ ($f \sqsupseteq f_{ab}$ cuando $b \sqsubseteq f(a)$). Así f es la menor de la cotas superiores de $\bigsqcup\{f_{ab} \mid f_{ab} \sqsubseteq f\}$.

Dado que $f \sqsubseteq \bigsqcup\{f_{ab} \mid f_{ab} \sqsubseteq f\}$ y suponiendo que f es elemento compacto de $[P \to Q]$, entonces $f \sqsubseteq f_{ab}$ para algunos $a, b \in K(P)$, pero $f_{ab} \sqsubseteq f$. Luego $f = f_{ab}$.

Definición 2.1.12. Decimos que un subconjunto B de un dcpo P es una base para P, si para todo elemento $x \in P$ el conjunto $B_x = \Downarrow x \cap B$ es un conjunto dirigido con supremo x. Llamamos a los elementos de B_x aproximaciones de x relativas a B.

Nota 2. Observe que $\perp \in B$. En efecto $B_\perp = \{\perp\} \cap B$ es vacío o es el conjunto $\{\perp\}$, como $B_\perp$ es dirigido entonces $B_\perp = \{\perp\}$, así que $\perp \in B$.

Una base para $\mathbb{R} \cup \{\infty\}$ (los reales necesitan elemento máximo para ser dcpo) es $\mathbb{Q} \cup \{\infty\}$. Existen otras posibles bases como los números diádicos y los números irracionales añadiendo el elemento máximo (se define un número diádico como un número de la forma $m/2^n$, con $m \in \mathbb{Z}$ y $n \in \mathbb{Z}^+$).

Definición 2.1.13. Un dcpo es llamado continuo si este tiene una base. Este es llamado algebraico si tiene una base de elementos compactos. Decimos que P es ω-continuo si existe una base contable y decimos que es ω-algebraico si $K(P)$ es una base contable.

Una base no sólo da una aproximación de los elementos, sino que también aproxima la relación de orden:

Proposición 2.1.14. *Sea D un dominio continuo con base B y sean x e y elementos de D. Entonces $x \sqsubseteq y$, $B_x \subseteq B_y$ y $B_x \subseteq \downarrow y$ son todos equivalentes.*

Se sigue de la definición 2.1.13 que todo dcpo algebraico es continuo. Otra forma de definir dcpo algebraico es la siguiente:

Proposición 2.1.15. *Un dcpo P es algebraico si y sólo si para todo $x \in P$, $x = \bigsqcup(K(P) \cap \downarrow x)$. El orden de aproximación esta caracterizado por $y \ll x$ si y sólo si existe $c \in K(P)$ tal que $y \sqsubseteq c \sqsubseteq x$.*

Demostración. Sean

$$A_x = K(P) \cap \downarrow x = \{\ y \in K(P) | y \sqsubseteq x\}$$
$$B_x = K(P) \cap \Downarrow x = \{y \in P | \exists c \in K(P), y \sqsubseteq c \sqsubseteq x\}.$$

Si D es un dcpo algebraico $x = \bigsqcup B_x$, es inmediato que x es cota superior de A_x, razonando por contradicción suponga que existe u cota superior de A_x tal que $u \sqsubseteq x$.

Como u es mayor o igual que todos los elementos compactos menores o iguales que x, entonces para $c \in K(P)$ con $y \sqsubseteq c \sqsubseteq x$, $c \sqsubseteq u$. Luego $y \sqsubseteq u$ para cualquier $y \in B_x$, contradiciendo que x es la mínima cota superior de B_x.

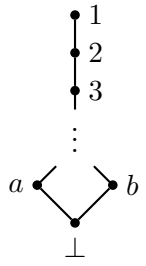
Ahora suponga que $x = \bigsqcup A_x$, es inmediato que x es cota superior de B_x, razonando por contradicción suponga que existe u cota superior de B_x tal que $u \sqsubseteq x$. Note que $B_x \supseteq A_x$, luego u también es cota superior de A_x, contradiciendo que x es la mínima cota superior de A_x. $\square$

Nota 3. $A_x = \{\ y \in K(P) | y \sqsubseteq x\}$ se conoce como el conjunto de aproximaciones de x. Así un dcpo D es algebraico si para cada $x \in D$, sus aproximaciones forman un conjunto dirigido y x es el supremo del sus aproximaciones.

Ejemplo 2.1.16. $(\Sigma^{**}, \sqsubseteq)$ es algebraico.

Ejemplo 2.1.17. $((X \multimap Y), \sqsubseteq)$ es algebraico.

Ejemplo 2.1.18. $P = 1 \oplus \overline{2} \oplus \mathbb{N}^{\mathrm{OP}}$ es un dcpo algebraico pero $[P \to P]$ no lo es.



$D \subseteq P$ es dirigido si y sólo si $D \neq \{a, b\}$ y $D \neq \{a, b, \perp\}$. Como P satisface la condición de cadena ascendente, entonces todo subconjunto no vacío de P tiene elemento maximal. En el caso de los subconjuntos dirigidos de P, todo subconjunto dirigido tiene máximo. Así para todo $p \in P$, $p \sqsubseteq \bigsqcup D$ implica $p \sqsubseteq \bigsqcup D \in D$. Luego p es compacto, es decir, $K(P) = P$. Para todo $x \in P$, $x = \bigsqcup\{p \in P \mid p \sqsubseteq x\} = \bigsqcup\{p \in K(P) \mid p \sqsubseteq x\}$, es decir, P es algebraico.

Los elementos compactos de $[P \to P]$ son las funciones

$$f_{k\widetilde{k}} = \begin{cases} \widetilde{k} \text{ si } x \sqsupseteq k \\ \perp \text{ en caso contrario} \end{cases} \quad , \qquad \forall k, \widetilde{k} \in K(P) = P$$

Sea

$$f(x) = \begin{cases} x \text{ si } x \neq a \\ b \text{ si } x = a \end{cases}$$

$f \in [P \to P]$, en efecto, sea D subconjunto dirigido de P:

Si $D \cap \mathbb{N} \neq \emptyset$, $\bigsqcup D = n$ y $\bigsqcup f(D) = n = f(n) = f(\bigsqcup D),\ n \in \mathbb{N}$

Si $a \notin D$, $\bigsqcup f(D) = \bigsqcup D = f(\bigsqcup D)$

Si $D \cap \mathbb{N} = \emptyset$ y $a \in D$, entonces $D = \{a\}$ ó $D = \{\perp, a\}$. En cualquier caso $\bigsqcup D = a$ y $\bigsqcup f(D) = b = f(a) = f(\bigsqcup D)$.

Pero f no es aproximable por compactos. Sea $D = \{f_{k\widetilde{k}} \mid f_{k\widetilde{k}} \sqsubseteq f\}$. Suponga D dirigido (en caso contrario $[P \to P]$ no es algebraico). $f_{k\widetilde{k}} \sqsubseteq f$ si y sólo si $f_{k\widetilde{k}}(x) \sqsubseteq f(x),\ \forall x \in P$. En particular $f_{k\widetilde{k}}(a) \sqsubseteq f(a) = b$, entonces $\widetilde{k} \sqsubseteq b$ si $k \sqsubseteq a$.

Por tanto si $k \sqsubseteq a$, $f_{k\widetilde{k}} \in \{f_{\perp\perp}, f_{\perp b}, f_{a\perp}, f_{ab}\}$.

Si $k > a$, $f(k) = k$. Ahora bien, $f_{k\widetilde{k}} \sqsubseteq f$ si y sólo si $\widetilde{k} \sqsubseteq f(k)$. Luego para $k > a$ $f_{k\widetilde{k}} \sqsubseteq f$ si y sólo si $\widetilde{k} \sqsubseteq k$.

Entonces $D \subseteq A_1 \cup A_2$, donde $A_1 = \{f_{\perp\perp}, f_{\perp b}, f_{a\perp}, f_{ab}\}$ y $A_2 = \{f_{k\widetilde{k}} \mid k > a \text{ y } \widetilde{k} \sqsubseteq k\}$. $A_2 \subseteq D$, pero tal vez algunas funciones de A no pertenecen a D. Verificando:

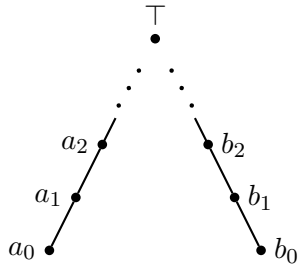
- $f_{\perp\perp}(x) = \perp = f_{a\perp}(x),\ \forall x \in P$. Además, $\perp \sqsubseteq f(x),\ \forall x \in P$. Entonces $f_{\perp\perp} = f_{a\perp} \in D$.
- $f_{\perp b}(x) = b,\ \forall x \in P$, $f_{\perp b}(\perp) = b \not\leq f(\perp) = \perp$. Entonces $f_{\perp b} \notin D$.
- $f_{ab}(x) = \begin{cases} b \text{ si } x \sqsupseteq a \\ \perp \text{ en caso contrario} \end{cases}$

  $f_{ab}(a) = b = f(a)$ (esto ya se había verificado); $f_{ab}(b) = \perp \sqsubseteq f(b) = b$; $f_{ab}(\perp) = \perp = f(\perp)$; si $x > a$, $f_{ab}(x) = b < f(x) = x$ pues $x > a$ implica $x > b$. Entonces $f_{ab} \in D$.

Luego $D = \{f_{\perp\perp}, f_{ab}\} \cup \{f_{k\widetilde{k}} \mid k > a \text{ y } \widetilde{k} \sqsubseteq k\}$. Si $f = \bigsqcup D$, entonces $b = f(b) = \bigsqcup\big(\{f_{\perp\perp}(b), f_{ab}(b)\} \cup \{f_{k\widetilde{k}}(b) \mid k > a \text{ y } \widetilde{k} \sqsubseteq k\}\big)$.

Ahora bien $f_{\perp\perp}(b) = f_{ab}(b) = f_{k\widetilde{k}}(b) = \perp$ para $k > a$. Por tanto, bajo la hipótesis $f = \bigsqcup D$ se llega a $b = \perp$. Así se concluye que $f \neq \bigsqcup D$, es decir, $[P \to P]$ no es algebraico.

Ejemplo 2.1.19. $(\mathbb{N} + \mathbb{N}) \oplus 1$ es un dcpo no continuo



Es fácil verificar que es un dcpo.

Ahora mostraremos que el orden de aproximación es vacío. Las parejas de la forma (a_j, b_j) y (b_j, a_j) no pueden estar relacionadas con el orden de aproximación pues son incomparables con la relación $\sqsubseteq$. Dos puntos $a_i \sqsubseteq a_j$ en la misma rama no están relacionados con el orden de aproximación pues $D = (b_n)_{n\in\mathbb{N}}$ es un conjunto dirigido con supremo $\top$, $a_j \sqsubseteq \bigsqcup D$ pero a_i no es menor o igual que ningún elemento en D. Análogamente se tiene que dos puntos $b_i \sqsubseteq b_j$ en la misma rama no están relacionados con el orden de aproximación.

Definición 2.1.20. Sea $P = (P, \sqsubseteq)$ un preorden, un subconjunto $E \subseteq P$ es consistente si $E \neq \emptyset$ y $\forall B \Subset E$ existe $y \in P$ tal que $\forall x \in B\ y \sqsupseteq x$ (todo subconjunto finito de E es acotado superiormente en P).

Definición 2.1.21. $(P, \sqsubseteq, \perp)$ es un orden parcial completo para consistentes ccpo si P tiene elemento mínimo $\perp$ y $\bigsqcup A$ existe en P siempre que $A \subseteq P$ sea consistente.

Ejemplo 2.1.22. $1 \oplus \overline{2} \oplus \mathbb{N}^{\mathrm{OP}}$ no es un ccpo pues $\{a, b\}$ es consistente pero $\bigvee\{a, b\}$ no existe.

Nota 4. Si P es un ccpo entonces $P \oplus 1$ es un retículo completo. En efecto, si $A \subseteq P$ y A acotado superiormente $\bigvee A$ existe; si A no esta acotado superiormente $\bigvee A = 1$. La reciproca es falsa por ejemplo $\mathbb{N} \oplus 1$ es un retículo completo pero $\mathbb{N}$ no es ccpo ($\mathbb{N}$ es consistente pero $\bigvee \mathbb{N}$ no existe).

Definición 2.1.23. Un dominio de Scott es un ccpo algebraico.

Definición 2.1.24. Un dominio continuo es un ccpo continuo.

Decimos $(P, \sqsubseteq, \perp)$ es un orden parcial completo para acotados bcpo si P tiene elemento mínimo $\perp$ y $\bigsqcup A$ existe en P siempre que $A \subseteq P$ sea acotado. Es fácil verificar que un conjunto es ccpo si y sólo si es dcpo y bcpo. Así que un dominio de Scott también se puede definir como un dcpo algebraico completo para acotados. Análogamente se puede definir un dominio continuo.

Ejemplo 2.1.25. $(\mathbb{IR}, \sqsubseteq)$ no es algebraico pues el único elemento compacto es $[-\infty, \infty]$. La relación "muy por debajo de" esta dada por $I \ll J$ sii $\mathrm{int}(I) \supseteq J$. $\mathbb{IR}$ es un dominio ω-continuo, una base contable esta dada por el conjunto de todos los intervalos con extremos racionales.

2.2. La topología de Scott

Definición 2.2.1. Considere un dcpo P. Un subconjunto $O \subseteq P$ es Scott abierto si se tiene lo siguiente:

1. $x \in O$ y $x \sqsubseteq y$ implica $y \in O$ (O es un conjunto superior).
2. Para cualquier conjunto dirigido $D \subseteq P$, $\bigsqcup D \in O$ implica que $D \cap O \neq \emptyset$ (O es inaccesible por dirigidos).

Para dcpo's algebraicos la definición de arriba es equivalente a la siguiente definición.

Definición 2.2.2. Sea P un dcpo algebraico, $O \subseteq P$ es Scott abierto si

1. $x \in O$ y $x \sqsubseteq y$ implica $y \in O$ (O es un conjunto superior).
2. $x \in O$ implica que existe $a \in \mathrm{aprox}(x) = \downarrow x \cap K(P)$ tal que $a \in O$

Suponga que O es abierto según la definición 2.2.1 y sea $x \in O$, $K(P) \cap \downarrow x$ es dirigido y $\bigsqcup K(P) \cap \downarrow x = x \in O$, entonces $K(P) \cap \downarrow x \cap O \neq \emptyset$ con lo que O satisface la definición 2.2.2. Recíprocamente suponga que O es abierto según la definición 2.2.2 y que $\bigsqcup D \in O$, existe $a \in K(P) \cap \downarrow \bigsqcup D$ con $a \in O$, entonces $a \in K(P)$ y $a \sqsubseteq \bigsqcup D$, luego existe $d \in D$ tal que $a \sqsubseteq d$ y como O es superior $d \in O$, así se tiene $d \in D \cap O$ con lo que O satisface la definición 2.2.1.

Una simple observación es que la topología de Scott en un dominio es T_0, sin embargo la topología no es T_1 excepto en el caso trivial que el dominio sea un conjunto unitario.

El conjunto $\Uparrow x$ es un conjunto abierto en cualquier dominio continuo. En efecto, suponga que D es subconjunto dirigido de $\Uparrow x$ con $\bigsqcup D \in \Uparrow x$, entonces $\bigsqcup D \gg x$ y por la *propiedad de interpolación* (2.4) existen $z_1, z_2 \in P$ tales que $x \ll z_1 \ll z_2 \ll \bigsqcup D$, por la propiedad (2.2) $z_2 \sqsubseteq \bigsqcup D$, luego $z_1 \sqsubseteq d$ para algún $d \in D$. Por último $x \sqsubseteq x \ll z_1 \sqsubseteq d$ de lo que se concluye por la propiedad (2.3) que $x \ll d$, es decir, $d \in D \bigcap \Uparrow x$.

Dado un elemento compacto a, $a \sqsubseteq y$ implica $a \ll y$. Entonces se tiene $\Uparrow a = \uparrow a$ para todo $a \in D_c$. Por tanto $\uparrow a$ es abierto para $a \in K(P)$.

Teorema 2.2.3. *Sea D un dcpo con una base B, entonces una base topológica para la topología de Scott en D esta dada por la familia $\{\Uparrow x \mid x \in B\}$*

Demostración. El conjunto $\Uparrow x$ es abierto para todo x en D como se acabo de demostrar.

Sea $w \in \Uparrow x \cap \Uparrow y$, es decir, $x, y \ll w$. Por interpolación escogemos $b \in B$ tal que $x, y \ll b \ll w$. Entonces $w \in \Uparrow b \subseteq y \Uparrow x \cap \Uparrow y$. Análogamente se muestra que para cada U subconjunto abierto de D,

$$U = \bigcup_{x \in U \cap B} \Uparrow x.$$

□

En el caso algebraico tenemos que $\{B_a : a \in K(P)\}$, donde $B_a = \uparrow a$ es una base para la topología de Scott.

Cualquier función continua entre dominios esta determinada por sus valores en una base del dominio. Sean D y E dominios continuos y sea B una base del dominio D. Una función monótona $f : B \to E$ tiene una única extensión a una función continua $g : D \to E$ tal que $f = g|_B$.

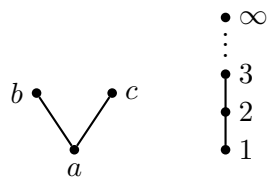
Proposición 2.2.4. *Sea P un dcpo, un subconjunto $F \subseteq P$ es Scott cerrado si se satisface lo siguiente:*

1. *$x \in F$ y $y \sqsubseteq x$ implica $y \in F$ (F es un conjunto inferior).*
2. *Si D subconjunto dirigido de F entonces $\bigsqcup D \in F$.*

Demostración. Primero se mostrará que F es inferior si y sólo si F^c es superior. Sean F inferior, $O = F^c$, $x \in O$ y $y \geq x$. Por contradicción suponga que $y \notin O$, entonces $y \in F$, luego $x \in F$ pues F es inferior, lo cual contradice que $x \in O = F^c$. Análogamente se tiene que si O es superior O^c es inferior.

Sean O Scott abierto y $F = O^c$ con $D \sqsubseteq F$. Si $\bigsqcup D \notin F = O^c$ entonces $\bigsqcup D \in O$ luego dado que O es abierto existe $d \in O \cap D$ lo cual es una contradicción pues $d \in D \subseteq F$ y $d \notin F$. Por tanto O^c es Scott cerrado.

Sean F Scott cerrado y $O = F^c$ con $D \sqsubseteq P$ tal que $\bigsqcup D \in O$. Dado que F es cerrado $\bigsqcup D \notin F \Rightarrow D \not\sqsubseteq F$ (contrarecíproca de 2.2.4-2) $\Rightarrow \exists d \in D - F \Rightarrow \exists d \in D \cap F^c = D \cap O$. □



Ejemplo 2.2.5. Para el dcpo $\overline{2}_\perp$ (Figura 2.2) se identifican $\mathcal{F}$ la familia de Scott cerrados

Los conjuntos inferiores son

$$F = \emptyset, \quad F = \{a\}, \quad F = \{a, b\}, \quad F = \{a, c\}, \quad F = \{a, b, c\}.$$

Si $F = \emptyset$, F no tiene subconjuntos dirigidos pues por definición si D es dirigido $D \neq \emptyset$. Todo subconjunto unitario es dirigido y $\bigsqcup\{x\} = x$, así que en adelante se consideraran sólo subconjuntos con más de dos elementos. Si $D = F = \{a, b\}$, $\bigsqcup D = b$. Si $D = F = \{a, c\}$, $\bigsqcup D = c$. Si $D = \{a, b\} \subseteq F = \{a, b, c\}$ ó $D = \{a, c\} \subseteq F = \{a, b, c\}$ claramente se tiene $\bigsqcup D \in F$.

Luego la familia de cerrados en $\overline{2}_\perp$ con la topología de Scott es

$$\mathcal{F} = \big\{\emptyset,\ \{a\},\ \{a, b\},\ \{a, c\},\ \{a, b, c\}\big\}.$$

Ejemplo 2.2.6. Para el dcpo $\mathbb{N} \oplus 1$ se identifican $\mathcal{F}$ la familia de Scott cerrados

Si F es inferior,

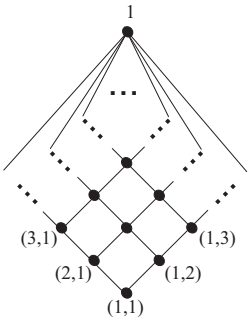
$$\begin{aligned} &F = \downarrow x,\ x \in \mathbb{N} \text{ (ideales principales)} \\ &F = \mathbb{N} \oplus 1 \\ &F = \emptyset. \end{aligned}$$

Cualquier subconjunto de $\mathbb{N} \oplus 1$ es dirigido pues sus elementos son comparables. Si $D \subseteq\downarrow x$, $(x \in \mathbb{N})$ entonces $\bigsqcup D = n \in \downarrow x$. Si $D \subseteq \mathbb{N} \oplus 1$ entonces $\bigsqcup D = n \in \mathbb{N}$ ó $\bigsqcup D = \infty$, en cualquier caso $\bigsqcup D \in \mathbb{N} \oplus 1$. Si $F = \emptyset$, F no tiene subconjuntos dirigidos.

Se concluye que la familia de Scott-cerrados en $\mathbb{N} \oplus 1$ es

$$\mathcal{F} = \{\downarrow x \mid x \in \mathbb{N}\} \cup \{\mathbb{N} \oplus 1, \emptyset\}.$$

Ejemplo 2.2.7. $(\mathbb{N} \times \mathbb{N}) \oplus 1$



Los conjuntos inferiores de $(\mathbb{N} \times \mathbb{N}) \oplus 1$ son de la forma

$$F_0 = \bigcup_{i \in I} \downarrow (a_i, b_i),\ \emptyset \neq I \Subset \mathbb{N},\ (a_i, b_i) \in \mathbb{N} \times \mathbb{N}$$

(unión finita de ideales principales en $\mathbb{N} \times \mathbb{N}$).

$$F_1 = \bigcup_{i \in I} \downarrow (a_i, b_i),\ (a_i, b_i) \in \mathbb{N} \times \mathbb{N},\ I \text{ subconjunto infinito de } \mathbb{N},\ (a_i, b_i) \neq (a_j, b_j)\ \forall i \neq j$$

(unión infinita enumerable de ideales principales en $\mathbb{N} \times \mathbb{N}$).

y por supuesto $F_2 = (\mathbb{N} \times \mathbb{N}) \oplus 1$, $F_3 = \emptyset$.

Afirmamos que $D_1 = \{(1, j) \mid j \in \mathbb{N}\} \subseteq F_1$ o $D^1 = \{(j, 1) \mid j \in \mathbb{N}\} \subseteq F_1$. En efecto, si se supone lo contrario existen R, $S \in \mathbb{N}$ tales que $(1, R)$ y $(S, 1) \in F_1$ pero $(1, R+1)$ y $(S+1, 1) \notin F_1$. Así (S, R) sería cota superior de F_1

$$\begin{aligned}&(a, b) \in F_1 \text{ conjunto inicial} \Rightarrow (a, 1) \in F_1 \text{ y } (1, b) \in F_1\\ &\Rightarrow a \sqsubseteq S \text{ y } b \sqsubseteq R \Rightarrow (a, b) \sqsubseteq (S, R),\end{aligned}$$

contradiciendo que F_1 es infinito.

Si $D_1 \subseteq F_1$, como D_1 dirigido y $\bigsqcup D_1 = 1 \notin F_1$, F_1 no es cerrado. Se llega a la misma conclusión si $D^1 \subseteq F_1$.

Son cerrados los conjuntos de la forma F_0 además de F_2 y F_3. Por tanto la familia de Scott-cerrados en $(\mathbb{N} \times \mathbb{N}) \oplus 1$ es

$$\mathcal{F} = \left\{ \bigcup_{i \in I} \downarrow (a_i, b_i) \;\middle|\; \emptyset \neq I \Subset \mathbb{N},\ (a_i, b_i) \in \mathbb{N} \times \mathbb{N} \right\} \cup \{(\mathbb{N} \times \mathbb{N}) \oplus 1, \emptyset\}.$$

Proposición 2.2.8. *La definición 2.1.7 de función continua es equivalente a la continuidad inducida por la topología de Scott en dcpos.*

La prueba de esta proposición la puede encontrar en [1].

Ejemplo 2.2.9. La topología de Scott es radicamente diferente de la topología Euclidiana en [0,1]. La función $f : [0, 1] \to [0, 1]$ definida por

$$f = \begin{cases} 0 \text{ si } 0 \leq x \leq \frac{1}{2} \\ 1 \text{ si } \frac{1}{2} < x \leq 1 \end{cases}$$

es Scott continua, pero no es continua con respecto a la topología usual.

Definición 2.2.10. Considere un espacio topológico (X, τ), entonces τ define un preorden de especialización $\sqsubseteq$ en X, $x \sqsubseteq y$ si $x \in O$ implica $y \in O$ para cualquier $O \in \tau$. Una

topología en un conjunto parcialmente ordenado se llama orden consistente si su orden de especialización coincide con el orden del poset.

2.3. Aproximaciones, Ordenamientos y Dominios

Deseamos computar en una estructura no contable como el campo de los números reales $\mathbb{R}$. Los objetos de $\mathbb{R}$ son en general objetos realmente infinitos (sucesiones de Cauchy ó cortaduras de Dedekind) sin descripción finita.

Ahora bien, los cómputos que un computador digital o una máquina de Turing pueden realizar se hacen sobre objetos finitos, es decir, objetos que son finitamente describibles o, equivalentemente codificados por un número natural. En particular, la estructura sobre la que los cómputos se realizan debe ser contable, así pues, no es posible computar directamente sobre $\mathbb{R}$, mejor computamos sobre aproximaciones finitas de elementos de $\mathbb{R}$ interpretando un computo de un número real como el "límite"de los cómputos de sus aproximaciones donde tal límite existe. De esto se sigue que los cómputos son procesos continuos.

2.3.1. Aproximaciones y ordenamientos

Consideremos el problema de aproximación abstractamente. Supongamos que X es un conjunto, o más generalmente una estructura. Decimos que un conjunto P es una aproximación para X (entendiendo que los elementos de P son aproximaciones para los elementos de X). Esto es, existe una relación $\prec$, la relación de aproximación de P a X,

$$p \prec x \Leftrightarrow \text{“}p \text{ aproxima } x\text{”}, \quad \text{donde } p \in P \text{ y } x \in X.$$

Ejemplo 2.3.1. Sea $P = \{[a, b] : a \leq b, a, b \in \mathbb{Q}\}$ y $X = \mathbb{R}$. Defina

$$[a, b] \prec x \Leftrightarrow x \in [a, b].$$

Note que P es contable y consiste de elementos finitos, en el sentido que un intervalo $[a, b]$ es finitamente describible a partir de la descripción finita de los números a y b y los símbolos "[ ] ,". Además, cada $x \in \mathbb{R}$ es el "límite"(intersección) de sus aproximaciones.

Ejemplo 2.3.2. Sea $P = \mathbb{Q}$ y $X = \mathbb{R}$. Para $a \in \mathbb{Q}$ y $x \in \mathbb{R}$ defina

$$a \prec x \Leftrightarrow a < x,$$

donde $<$ es el orden usual en $\mathbb{R}$.

Ejemplo 2.3.3. Sea X un espacio topológico con una base topológica $\mathcal{B}$. Para $B \in \mathcal{B}$ y $x \in X$ defina

$$B \prec x \Leftrightarrow x \in B.$$

Sean P y X y sea $\prec$ una relación de P a X. $\prec$ induce una preorden $\sqsubseteq$ en P:

$$p \sqsubseteq q \Leftrightarrow (\forall x \in X)(q \prec x \Rightarrow p \prec x).$$

Luego $p \sqsubseteq q$ expresa que q es mejor aproximación que p en el sentido que q aproxima menos elementos en X que p.

Agregamos algunos requerimientos sobre P y $\prec$ para obtener una estructura de aproximación en X.

Definición 2.3.4. Sean P y X conjuntos, $\prec$ una relación de P en X y $\sqsubseteq$ el preorden obtenido de $\prec$. Entonces $(P, \sqsubseteq)$ es una estructura de aproximación para X con respecto a $\prec$ si

(i) $\{p \in P : p \prec x\} = \{p \in P : p \prec y\} \Rightarrow x = y$ (unicidad);

(ii) $p \prec x$ y $q \prec x \Rightarrow (\exists r \prec x)(p \sqsubseteq r$ y $q \sqsubseteq r)$ (refinamiento);

(iii) $(\exists p \in P)(\forall x \in X)(p \prec x)$ (aproximación trivial).

Los ejemplos 2.3.1 y 2.3.2 producen estructuras de aproximación cuando agregamos la aproximación trivial. El ejemplo 2.3.3 produce una estructura de aproximación cuando el espacio es T_0.

2.3.2. Ideales y dominios

Los dominios son frecuentemente construidos como el completamiento de alguna estructura subyacente. Estudiaremos tales estructuras y comenzaremos con las estructuras desde las que se pueden construir los dominios de Scott.

2.3.2.1. Construcción de Dominios Algebraicos

Los elementos compactos $K(P)$ de un dominio de Scott forman un semiretículo superior condicional con elemento mínimo, abreviado *cusl*.

Definición 2.3.5. Un *cusl* es un conjunto parcialmente ordenado donde la menor cota superior existe para toda pareja de elementos que tiene una cota superior.

Definición 2.3.6. Sea Q un cusl. Entonces $I \subseteq Q$ es un *ideal* si

(i) $I \neq \emptyset$

(ii) Si $a \in I$ y $b \sqsubseteq a$ entonces $b \in I$, y

(iii) Si $a, b \in I$ entonces existe una cota superior de a y b en I

Nota 5. (i) y (iii) es equivalente a que I es dirigido

De (ii) y (iii) y de la definición de cusl se tiene que si a y $b \in I$ entonces $a \sqcup b \in I$. $\downarrow a$ denota el ideal principal generado por a. La completación ideal sobre un cusl Q es el conjunto de todos los ideales sobre Q, denotado por $\mathrm{Idl}(Q)$. La completación ideal de

un cusl ordenada por inclusión forma un dominio de Scott. Los elementos compactos de $\mathrm{Idl}(Q)$ son los ideales principales $\downarrow a$, para $a \in Q$.

El teorema de representación de dominios de Scott nos lleva a que cualquier dominio de Scott es la completación ideal de un cusl. Denotamos $\overline{Q} = Idl(Q, \sqsubseteq)$, el conjunto de todos los ideales sobre $(Q, \sqsubseteq)$. Si P preorden, $Idl(P, \sqsubseteq)$ es un dcpo algebraico no necesariamente consistentemente completo.

Teorema 2.3.7. *Sea D dominio de Scott, entonces $D \simeq \overline{K(D)}$.*

Para su prueba ver [14].

Teorema 2.3.8. *Sea $P = (P, \sqsubseteq)$ un preorden y $\overline{P} = (Idl(P), \subseteq)$ su completación ideal ordenada por inclución.*

(i) *$\subseteq$ es un orden parcial.*

(ii) *Si $\mathcal{F} \subseteq \overline{P}$ es una familia dirigida de ideales, entonces $\bigcup \mathcal{F}$ es un ideal y $\bigcup \mathcal{F}$ es el supremo de $\mathcal{F}$.*

(iii) *Si $p \in P$, $\mathcal{F} \subseteq \overline{P}$ es dirigido y $\downarrow p \subseteq \bigcup \mathcal{F}$ entonces existe $I \in \mathcal{F}$ tal que $\downarrow p \subseteq I$.*

(iv) *Sea $I \in \overline{P}$. Suponga que para cada dirigido $\mathcal{F}$ tal que $I \subseteq \bigcup \mathcal{F}$ existe $J \in \mathcal{F}$ tal que $I \subseteq J$. Entonces $I = \downarrow p$ para algún $p \in P$.*

(v) *Sea $I \in \overline{P}$. Entonces el conjunto $\{\downarrow p \mid \downarrow p \subseteq I\}$ es dirigido e $I = \bigcup\{\downarrow p \mid \downarrow p \subseteq I\}$.*

La prueba de este teorema es sencilla, para (iv) sea $\mathcal{F} = \{\downarrow p \mid \downarrow p \subseteq I\}$.

Los elementos compactos de $Idl(P)$ son precisamente los ideales principales. Si $(P, \sqsubseteq)$ una estructura de aproximación para X con respecto a $\prec$, entonces cada $x \in X$ puede ser identificado de manera única con el conjunto $\{p \in P : p \prec x\}$. Note que si $p \sqsubseteq q \prec x$ entonces $p \prec x$. Junto con (ii) y (iii) de la definición 2.3.4 vemos que $\{p \in P : p \prec x\}$ es un ideal sobre $(P, \sqsubseteq)$. Dada una estructura de aproximación $(P, \sqsubseteq)$ de X con respecto a $\prec$, existe entonces una inyección de X en $\mathrm{Idl}(P)$. Así, $\mathrm{Idl}(P)$ contiene tanto a X como a sus aproximaciones.

Note que para D dcpo algebraico, $K(D)$ es una estructura de aproximación para D con respecto a $\prec$, donde para $a \in K(D)$ y $x \in D$,

$$a \prec x \Leftrightarrow a \sqsubseteq x.$$

Luego tenemos:

Corolario 2.3.9. *Los dcpos algebraicos son precisamente los completamientos ideales de estructuras de aproximación.*

Nota 6. Recuerde que $\mathbb{IR}$ no es algebraico, pero $\mathrm{Idl}(\mathbb{IR})$ si lo es.

2.3.2.2. Construcción de Dominios Continuos

Empezamos haciendo la pregunta ¿qué información es necesaria para construir un dominio desde una de sus bases?. La respuesta es que sólo requerimos el orden de aproximación $\ll$.

Ahora consideramos el tipo de estructura que nos lleva a construir dominios continuos.

Definición 2.3.10. Una base abstracta es una estructura $B = (B; \prec, \perp)$ donde $\prec$ es una relación transitiva en B tal que

$$M \prec x \Rightarrow \exists y \in B(M \prec y \prec x), \tag{2.5}$$

se tiene para todo elemento x y M subconjunto finito de B, donde $M \prec x$ significa $(\forall z \in M)(z \prec x)$. $\perp$ es el mínimo elemento de B, es decir, $(\forall x \in B)(\perp \prec x)$.

La relación $\prec$ no es otra cosa que el orden de aproximación $\ll$ restringido a B, el axioma 2.5 se satisface por la propiedad de interpolación mencionada en la página 16.

La relación $\prec$ en una base abstracta no necesita satisfacer la propiedad reflexiva ni antisimétrica. Un ideal A en una base tiene la propiedad que para todo $x \in A$ existe otro elemento $y \in A$ con $x \prec y$.

Extendemos la notación $\downarrow a$ para la relación transitiva $\prec$ por $\downarrow a = \{b \mid b \prec a\}$.

Definición 2.3.11. Para una base $\langle B, \prec\rangle$ sea $\mathrm{Idl}(B)$ el conjunto de todos los ideales ordenados por inclusión. Esta es llamada la completación ideal de B. Más aún $i : B \to \mathrm{Idl}(B)$ denota la función tal que $b \mapsto \downarrow b$. Si queremos hacer énfasis en la relación con que esta dotada B escribimos $\mathrm{Idl}(B, \prec)$ para la completación ideal.

Las demostraciones de las siguientes proposiciones se encuentran en [1].

Proposición 2.3.12. *Sea $(B, \prec, \perp)$ una base abstracta.*

(i) La completación ideal $Idl(B)$ ordenada por inclusión es un dominio continuo. La relación "muy por debajo de" esta caracterizada por $I \ll I'$ si y sólo si existe $a \prec b$ en B tal que $I \subseteq \downarrow a \subseteq \downarrow b \subseteq I'$. Una base para $Idl(B)$ esta dada por $i(B)$.

(ii) Si $\prec$ es reflexiva entonces $Idl(B)$ es un dcpo algebraico.

(iii) Si $\langle B, \prec\rangle$ es un poset entonces $B, K(Idl(B))$, e $i(B)$ son todos isomorfos.

Proposición 2.3.13. *Sea D un dominio continuo con base B. Asumiendo $\langle B, \ll\rangle$ como una base abstracta, tenemos lo siguiente*

1. *$Idl(B)$ es isomorfo a D. El isomorfismo $\sigma : Idl(B) \to D$ es la extensión de la inmersión de B en D. Su inversa β le asigna al elemento $x \in D$, B_x.*

2. *Para todo dcpo E y función continua $f : D \to E$ tenemos que $f = \hat{f} \circ \beta$, donde $\hat{f}$ es la restricción de f a B.*

Corolario 2.3.14. *Una función continua de un dominio D a un dcpo E esta completamente determinada por su comportamiento en una base de D.*

Podemos trabajar entonces sólo con los elementos de la base. Sin embargo hay un problema más por superar, el hecho de que las funciones continuas no preservan el orden de aproximación. La única forma de salir de esto es cambiar de funciones a relaciones, donde relacionamos un elemento c de la base con todos los elementos de la base que aproximan $f(c)$. Esto se puede axiomatizar como sigue.

Definición 2.3.15. Una relación R entre bases abstractas B y C es llamada aproximable si las siguientes condiciones se satisfacen:

1. $\forall x, x' \in B\ \forall y, y' \in C.\ (x' \succ xRy \succ y' \Rightarrow x'Ry')$.
2. $\forall x \in B, \forall M \Subset C.\ (\forall y \in M.\ xRy \Rightarrow \exists z \in C.\ xRz$ y $z \succ M)$.

2.4. Una Representación de Dominio Algebraica para los Números Reales

Del ejemplo 2.3.1 tomamos la estructura de aproximación para $\mathbb{R}$, de los intervalos cerrados finitos con extremos racionales, adicionando $\mathbb{R}$ y definiendo $\mathbb{R} \prec r$ para cada $r \in \mathbb{R}$ (aproximación trivial). Denotamos esta estructura de aproximación como $\mathbb{IQ}$. El orden de refinamiento inducido es $[a, b] \sqsubseteq [c, d]$ si y sólo si $[a, b] \supseteq [c, d]$. Sea $r \in \mathbb{R}$.

El ideal

$$I^r = \{[a, b] \in \mathbb{IQ} : a \leq r \leq b\} \cup \{\mathbb{R}\}.$$

representa a r. Note que I^r tiene la propiedad $\bigcap I^r = \{r\}$. Ahora considere el ideal

$$I_r = \{[a, b] \in \mathbb{IQ} : a < r < b\} \cup \{\mathbb{R}\}.$$

También en este caso $\bigcap I_r = \{r\}$ pero $I^r \neq I_r$ en el caso en que r es racional, e I_r también representa r. Se dice que $I \in \overline{\mathbb{IQ}}$ representa un número real r si $\bigcap I = \{r\}$. Sea $\overline{\mathbb{IQ}}^R$ el conjunto de todos los ideales cuya intersección es un singleton y defina una función $v : \overline{\mathbb{IQ}}^R \to \mathbb{R}$ por

$$v(I) = r \Leftrightarrow \bigcap I = \{r\}.$$

Proposición 2.4.1. *La función $v : \overline{\mathbb{IQ}}^R \to \mathbb{R}$ es continua y sobre con respecto a las topologías de Scott y Euclidiana.*

La prueba es simple y depende del hecho que v se asume como una función cociente. Este es un resultado topológico que no tiene nada de especial en el caso de estos dominios.

Tenemos entonces el siguiente diagrama

$$\mathbb{IQ} \hookrightarrow \overline{\mathbb{IQ}} \hookleftarrow \overline{\mathbb{IQ}}^R \xrightarrow{v} \mathbb{R}.$$

La tupla $(\overline{\mathbb{IQ}}, \overline{\mathbb{IQ}}^R, v)$ es un ejemplo de un dominio de representación de $\mathbb{R}$. Es cerrado hacia arriba en el sentido que si $I \in \overline{\mathbb{IQ}}^R$ e $I \subseteq J \in \overline{\mathbb{IQ}}$ entonces $J \in \overline{\mathbb{IQ}}^R$.

Lema 2.4.2. $(\mathbb{IQ}, \sqsubseteq)$ *es un cusl.*

Demostración. El elemento mínimo es $(-\infty, \infty)$. El supremo de un conjunto consistente de intervalos es su intersección. □

2.4.1. Computabilidad

Por el Lema 2.4.2 sabemos que los intervalos racionales forman un cusl. La completación ideal de este cusl es el dominio algebraico de intervalos. Ahora nos restringimos a los intervalos racionales cuyos extremos son números diádicos. Un número diádico es un número de la forma $z/2^n$ con $n \in \mathbb{N}$ y $z \in \mathbb{Z}$.

Lema 2.4.3. *Los intervalos diádicos forman un sub-cusl de los intervalos racionales.*

Demostración. Solo necesitamos verificar que el supremo de dos intervalos diádicos consistente es de nuevo un intervalo centrado diádico. El supremo es la intersección de los dos intervalos. El extremo izquierdo de la intersección es uno de los extremos izquierdos de los dos intervalos, luego el extremo izquierdo es diádico. De manera similar se concluye que el extremo derecho es diádico. □

Un número real x es representado por una secuencia de enteros $< s_0, \ldots, s_i, \ldots >$ tal que

(i) $\forall n,\ 2s_n - 1 \leq s_{n+1} \leq 2s_n + 1$

(ii) $x = \bigcap_{n\in\mathbb{N}} \left[\frac{s_n-1}{2^n}, \frac{s_n+1}{2^n}\right]$

Una secuencia de enteros se utiliza para describir una secuencia de intervalos racionales. Cada intervalo en la secuencia esta contenido en el anterior.

Comenzamos a construir nuestro dominio algebraico a partir de la representación de los números reales como una secuencia de enteros. Sea $< s_i >_{i\in\mathbb{N}}$ una secuencia de enteros que define un número real x. Para cualquier secuencia finita $< s_i >_{i<n}$ asociamos el intervalo $[a, b]$ que contiene los números reales que pueden ser representados por éste. El intervalo $[a, b]$ representa la información contenida en la secuencia $< s_i >_{i<n}$.

Sea $(ID, \sqsubseteq)$ el orden parcial formado por los intervalos racionales:

$$\left[\frac{s_n - 1}{2^n}, \frac{s_n + 1}{2^n}\right]$$

$\forall n$, donde s_n satisface (i). Ahora consideremos su completación ideal $(\overline{ID}, \sqsubseteq)$. Definimos una numeración efectiva de los elementos de $(\overline{ID}, \sqsubseteq)$ de la siguiente forma:

$$I_0 = \mathbb{R}$$
$$I_{\xi(m,m',n)+1} = \downarrow [(m - m' - 1)/2^n, (m - m' + 1)/2^n]$$

donde ξ es la numeración de $\mathbb{N}^+ \times \mathbb{N}^+ \times \mathbb{N}^+$ mostrada en el Teorema 1.4.4.

Los elementos de $\overline{ID}$ pueden pensarse como las clases de equivalencia de secuencias de enteros. Cada clase de equivalencia esta compuesta por secuencias que contienen información idéntica sobre el valor real que ellas aproximan.

Definición 2.4.4. Sea $q : \overline{ID} \to \wp(\mathbb{R})$ definida por

$$q(d) = \bigcap_{[a,b]\in d} [a, b],$$

y sean $I, I^-, I^+ : \mathbb{R} \to \overline{ID}$ que a cada $x \in \mathbb{R}$ le asigna I_x, I_x^-, I_x^+ respectivamente, donde:

$$\begin{aligned} I_x &= \{[a, b] \in ID \mid x \in (a, b)\} \cup \{\mathbb{R}\} \\ I_x^- &= \{[a, b] \in ID \mid x \in (a, b]\} \cup \{\mathbb{R}\} \\ I_x^+ &= \{[a, b] \in ID \mid x \in [a, b)\} \cup \{\mathbb{R}\} \end{aligned}$$

Proposición 2.4.5. *Se tienen las siguientes proposiciones:*

(i) Para todo elemento maximal en $\overline{ID}$ existe un número real x tal que $q(d) = \{x\}$

(ii) Para todo número real x, $\{x\} = q \circ I(x) = q \circ I^-(x) = q \circ I^+(x)$.

(iii) Para todo número diádico x, $I_x \subset I_x^-$, $I_x \subset I_x^+$ y $\{I_x^-, I_x^+\}$ no es consistente.

Observamos que el conjunto de elementos maximales de $\overline{ID}$ es similar a la recta real, excepto porque cada número diádico es triplicado.

Definición 2.4.6. Un elemento $x \in D$ dominio algebraico es computable si el conjunto de sus aproximaciones $A_x = \{\ y \in K(P) | y \sqsubseteq x\}$ es efectivamente numerable.

Definición 2.4.7. Un número real x es computable si existe un elemento computable $d \in \overline{ID}$ tal que $x = q(d)$.

2.5. El dominio continuo de intervalos

Recuerde $\mathbb{IR}$ es un dominio ω-continuo (no es algebraico, pues el único elemento compacto es $(-\infty, \infty)$ y por tanto los elementos del dominio no pueden ser aproximados por compactos). El sup de cualquier subconjunto dirigido en este dcpo (es decir, cualquier conjunto filtrado de intervalos compactos) es la intersección de los intervalos. La relación muy por debajo de, esta dada por $I \ll J$ sii $\mathrm{int}(I) \supseteq J$. Una base contable esta dada por el conjunto de todos los intervalos con extremos racionales.

Un número real x es aproximado por una cadena creciente, es decir, una secuencia contractante de intervalos encajados de extremos racionales. Un Scott abierto básico de $\mathbb{IR}$, esta dado para cualquier abierto $O \subseteq \mathbb{R}$, por la colección $\Box O = \{a \in \mathbb{IR} \mid a \subseteq O\}$. Los elementos maximales de $\mathbb{IR}$ son los intervalos $[a, a]$, es decir, los singletons. La función $s : \mathbb{R} \to \mathbb{IR}$ tal que $s(x) = [x, x]$, en una inmersión de la recta real sobre el conjunto de elementos maximales. Como $s^{-1}(\Box O) = O$, la topología euclidiana coincide con la topología de Scott en el subespacio de elementos maximales. Cualquier función continua

$f : \mathbb{R} \to \mathbb{R}$ extiende canónicamente a una función Scott continua $\mathbb{I}f : \mathbb{IR} \to \mathbb{IR}$ definida en cualquier intervalo compacto a por $\mathbb{I}f(a) = f(a)$. Esta es la extensión maximal de f sobre $\mathbb{IR}$, en otras palabras, si la función continua $g : \mathbb{IR} \to \mathbb{IR}$ satisface $g(\{x\}) = \{f(x)\}$ para todo $x \in \mathbb{R}$ entonces $g \sqsubseteq \mathbb{I}f$.

El dominio algebraico considerado para representar la recta real es la completación por ideales del conjunto de todos los intervalos con extremos racionales $\overline{\mathbb{IQ}}$, donde $\mathbb{R}$ es la inmersión como $\overline{\mathbb{IQ}}^R$. En contraste, el dominio continuo considerado para representar la recta real es $\mathbb{IR}$, donde $\mathbb{R}$ es simplemente la inmersión como el conjunto de sus elementos maximales.

2.5.1. Computabilidad en la recta real

Definición 2.5.1. Sea $(D, \sqsubseteq)$ un dominio continuo con base contable $D_0 = \{b_0, b_1, \ldots\}$. Un elemento $x \in D$ es computable, si el conjunto $\{n \in \mathbb{N} \mid b_n \sqsubseteq x\}$ es *r.e.* Sin perdida de generalidad en adelante asumimos que $b_0 = \bot$.

Una base contable para el dominio continuo $\mathbb{IR}$ de los intervalos, esta dada por el conjunto de todos los intervalos con extremos racionales junto con el elemento mínimo $\bot$. Este dominio será usado para dotar a los números reales de una estructura computable.

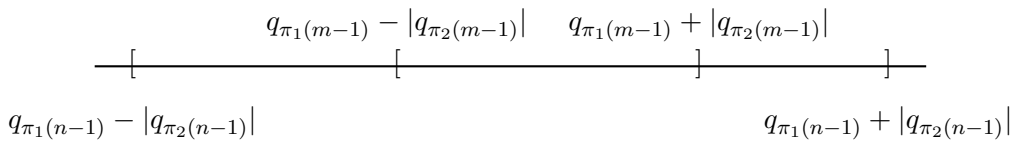
Sea $q_0, q_1, q_2, \ldots$ una numeración estándar de los racionales. Las operaciones aritméticas $+, -, \cdot, /$ como las comparaciones $<, \leq, =$ y la función valor absoluto $|\cdot|$ sobre los racionales son recursivas en sus índices.

Ahora se define una estructura efectiva para el dominio de intervalos:

$$\begin{aligned} I_0 &= \mathbb{R} \\ I_{\pi(n.m)+1} &= [q_n - |q_m|, q_n + |q_m|]. \end{aligned}$$

donde π es la numeración de $\mathbb{N} \times \mathbb{N}$ mostrada en el Teorema 1.4.4.

Claramente, esta enumera la base $\mathbb{IR}_0$. La relación "muy por debajo de" es r.e. Tenemos $I_n \ll I_m$ sii $I_m \subseteq \mathrm{int}(I_n)$ sii



$$n = 0 \text{ ó } \big(m, n > 0 \text{ y } |q_{\pi_1(n-1)} - q_{\pi_1(m-1)}| < |q_{\pi_2(n-1)}| - |q_{\pi_2(m-1)}|\big),$$

luego esta relación es siempre recursiva. La escogencia particular de la base y la enumeración para la base no es esencial para la teoría pues se puede pasar efectivamente de una base a otra.

Un número real x es *computable por izquierda*, si el conjunto $\{n \in \mathbb{N} \mid q_n < x\}$ es r.e. La *computabilidad por derecha* se define de manera similar.

Proposición 2.5.2. *Un intervalo $[x, y] \in \mathcal{I}$ es computable si y sólo si x es computable por izquierda y y es computable por derecha.*

Demostración. Por la Definición 2.5.1 el intervalo es computable si y sólo si $\{n | I_n \sqsubseteq [x, y]\}$ es r.e. Ahora $q_n < x$ si y sólo si existen m y k en $\mathbb{N}$ tal que $q_n = q_m - |q_k|$ y $I_{(m,k)+1} \sqsubseteq [x, y]$. La relación $q_n > y$ se puede caracterizar de manera similar, de donde se tiene la proposición. □

2.6. Algunos Modelos para espacios clásicos

En esta sección hablaremos de la inmersión de algunos espacios clásicos en los elementos maximales de sus respectivos dominios de aproximación. Aunque estos espacios no son el objeto de estudio de éste trabajo, se mencionan para dar una idea general de la teoría de dominios como teoría de la aproximación y teoría de la computabilidad de estructuras matemáticas via aproximaciones. También se mencionan para dar un panorama de temas a quienes deseen continuar su estudio.

2.6.1. Una Representación de Dominio Algebraica en Espacios Topológicos

A continuación generalizamos los resultados de la Sección 2.4 a un espacio topológico arbitrario.

Definición 2.6.1. Sea X un espacio topológico, D un dominio y D^R un subconjunto de D. Entonces (D, D^R, v) es una representación de dominio de X en el caso que $v : D^R \to X$ es una función continua y sobreyectiva cuando D^R esta dotado de la topología de Scott.

$$\begin{tikzcd}
 & D & \\
D_c \arrow[ur, hook] & & D^R \arrow[ul, two heads] \arrow[r, "v"] & X \\
\text{Teoría de la computación} \quad \mathbb{N} \arrow[u] & & &
\end{tikzcd}$$

Se utiliza el termino genérico "dominio" en el sentido de cualquier tipo de estructura ordenada, en el caso de los espacios topológicos. Comúnmente las estructuras ordenadas usadas son dcpo's algebraicos, dominios algebraicos, dominios continuos. D_c denota los elementos computables del dominio en el caso algebraico $D_c = K(D)$.

La inmersión de un espacio topológico en los elementos maximales de un dominio es un caso especial de dominio de representación. Esto es bastante restrictivo si solo consideramos inmersiones en dominios algebraicos. Los espacios métricos pueden ser inmersos en los elementos maximales de dominios continuos.

2.6.2. Otros Dominios Continuos de Representación

2.6.2.1. El Espacio Superior

En [8] el autor presenta la noción de un modelo computacional de teoría de dominios para un espacio Hausdorff localmente compacto con base contable:

Sea X un espacio Hausdorff. El espacio superior $(\mathbf{U}X, \supseteq)$ de X consiste de todos los subconjuntos compactos no vacios de X ordenados por inclusión reversa. La topología de Scott en $\mathbf{U}X$ tiene una base dada por las colecciones $\Box O = \{C \in UX | C \subseteq O\}$ para cualquier $O \in \Omega X$ (el retículo de los conjuntos abiertos de un espacio topológico X es denotado por ΩX). Esta topología es T_0 y su orden de especialización (Definición 2.2.10), denotado por $\sqsubseteq_u$, es la inclusión reversa.

$$A \sqsubseteq_u B \Leftrightarrow \forall O \in \Omega X[A \subseteq O \Rightarrow B \subseteq O] \Leftrightarrow A \supseteq B$$

$(\mathbf{U}X, \supseteq)$ es un dcpo en el cual el sup de un conjunto dirigido de subconjuntos compactos es su intersección.

Cuando X es localmente compacto $(\mathbf{U}X, \supseteq)$ es un dcpo continuo completo para acotados con $A \ll B$ sii B esta en el interior de A, y la topología superior coincide con la topología de Scott. Si X es también 2-contable (es decir, tiene una base contable) entonces $\mathbf{U}X$ es un dominio ω-continuo.

Obtenemos una base contable de $\mathbf{U}X$ como sigue. Empezamos con una base contable de X consistente de subconjuntos abiertos relativamente compactos (Un conjunto es relativamente compacto si su clausura es compacta). En efecto un espacio 2-contable localmente compacto es metrizable, podemos tomar como base de X el conjunto de todas las uniones finitas de bolas abiertas relativamente compactas con radio racional centrado en puntos de un subconjunto denso contable de X. Entonces la colección de las uniones finitas de las clausuras de estos abiertos básicos forman una base contable para $\mathbf{U}X$.

La función singleton

$$\begin{aligned} s : X &\to \mathbf{U}X \\ x &\mapsto \{x\}, \end{aligned}$$

es una inmersión de X en el conjunto $s(X)$ de elementos maximales de $\mathbf{U}X$ dado que $s^{-1}(\Box O) = O$. Como en el caso del dominio de intervalos, la topología Hausdorff en X coincide con la topología de Scott inducida en el subespacio de elementos maximales de el espacio superior. Cualquier función continua $f : X \to Y$ de espacios metricos compactos

induce una función Scott continua $\mathbf{U}f : \mathbf{U}X \to \mathbf{U}Y$ definida por $\mathbf{U}f(C) = f(C)$. De hecho $\mathbf{U}$ es un funtor de la categoría de los espacios localmente compactos y funciones continuas a la categoría de los dcpo's continuos y funciones continuas. La función $\mathbf{U}f$ es la extensión maximal de f. Para simplificar la notación escribiremos $\mathbf{U}f$ como f.

Esto muestra que los dominios continuos obtenidos tomando subconjuntos compactos no vacíos del espacio, ordenados por inclusión reversa, proveen un modelo simple para computación en estos espacios clásicos que están inmersos en el subespacio de los elementos maximales del dominio. El dominio *espacio superior* esta equipado con una base de subconjuntos compactos no vacíos que pueden ser enumerados para dar una teoría de efectividad. Un elemento del espacio localmente compacto identificado como un singleton puede ser obtenido como la intersección de una secuencia creciente encajada de elementos de la base.

2.6.2.2. El espacio de las bolas formales de un espacio métrico completo

Una bola formal de un espacio métrico (X, d) es una pareja (x, r) con $x \in X$ y $r \in \mathbb{R}^+$. El espacio $\mathbf{B}X$ es el conjunto de todas las bolas formales de X con el orden $(x, r) \sqsubseteq (y, s) \Leftrightarrow d(x, y) \leq r - s$. Si denotamos las bolas cerradas con centro en x y radio r por $C(x, r)$, entonces $(x, r) \sqsubseteq (y, s) \Rightarrow C(x, r) \supseteq C(y, s)$. El converso en general no es cierto, pero se tiene para espacios vectoriales normados. En otras palabras, en cualquier espacio vectorial normado, el cpo de las bolas formales y el conjunto de las bolas cerradas parcialmente ordenadas por inclusión reversa son isomorficos. En particular, esto se tiene para espacios de Banach y consecuentemente para espacios de Hilbert.

En cualquier espacio métrico X, tenemos $(x, r) \ll (y, s) \Leftrightarrow d(x, y) < r - s$. Más aún el cpo $\mathbf{B}X$ es continuo y los subconjuntos $\Uparrow (x, r) = \{(y, s) \mid d(x, y) < r - s\}$ forma una base de la topología de Scott. Un espacio métrico X es separable si y sólo si $\mathbf{B}X$ tiene una base contable (es decir, es ω-continuo). Si X es separable, entonces las bolas formales de la forma (x, r) donde x pertenece a un subconjunto denso contable de X y r es un número racional positivo da una base contable de $\mathbf{B}X$.

También existe una conexión interesante entre la completitud de un espacio métrico y la de un poset: un espacio métrico X es Cauchy completo sii $\mathbf{B}X$ es completo para dirigidos. Los elementos maximales de $\mathbf{B}X$ son precisamente los elementos $(x, 0)$ para todo $x \in X$. La función

$$\begin{aligned} i : X &\to \mathbf{B}X \\ x &\to (x, 0) \end{aligned}$$

es una inmersión de X en el conjunto de los elementos maximales del espacio de las bolas formales dado que $i^{-1}(\Uparrow a(x, r)) = \{y \mid d(x, y) < r\}$.

CAPÍTULO 3

Intervalos Centrados Diádicos

3.1. Aritmética Real Exacta

Una aritmética real exacta es un álgebra de operaciones computables definidas en un subcampo de los números reales. El fundamento teórico para la computación real exacta y este enfoque de un tipo de dato real exacto es el objeto del análisis computable. La razón más irresistible para estudiar análisis computable es que todas la operaciones serán implementadas en un computador ordinario.

Dado que un computador sólo podría representar un subconjunto contable de los números reales, aún si la memoria de la computadora pudiera extenderse indefinidamente, existirían números reales que no pueden ser representados en un computador. Dichos números reales son no representables pues no pueden existir procedimientos que simultáneamente den todas las cotas superiores e inferiores de un número real. Las sucesiones de Cauchy se emplean en distintos ámbitos de las matemáticas, pero ellas no pueden manejarse fácilmente como entrada o salida de una computadora. La interfaz para aritmética real exacta no es simplemente los números de entrada, sino que también hay que especificar una precisión para la aproximación resultante. De manera similar, la entrada no es vista como sucesiones que son leídas elemento por elemento, sino como funciones tomando una precisión y retornando un elemento de la sucesión de Cauchy con la precisión deseada.

Un número de punto flotante en una máquina de 32 bits, utiliza un bit para el signo S, 8 bits para el exponente s y 23 bits para la mantisa m normalizada sin el 1 inicial

$$(-1)^S \cdot 0{,}1m \cdot 2^{s-127}.$$

Las cadenas de longitud finita pueden representar con exactitud un subconjunto limitado de los números reales. La mayoría de los números reales es representada por números reales cercanos o intervalos reales que los contienen dando origen a los errores de redondeo. Es bien conocido que la acumulación de errores de redondeo debido al gran número de cálculos puede producir grandes errores o incluso resultados incorrectos.

Resaltamos que la aritmética de punto flotante no es una aritmética real exacta. Esto se debe a la cantidad limitada de precisión obtenida en números de punto flotante (causando errores de redondeo), y el conjunto de los números de punto flotante no es un subcampo de

los números reales (causando overflow). El análisis de intervalos de Moore ha sido usado para evadir parcialmente este problema manteniendo un par de cotas para el número en cuestión. Sin embargo, este intervalo puede ser injustificadamente grande y por tanto expresa muy poca información.

Existen varias implementaciones de aritmética real exacta. Están generalmente basadas en anidar o intersectar intervalos racionales. En este capitulo consideramos intervalos centrados diádicos cuyas implementaciones han sido desarrolladas por Jens Blanck, y que corresponden a la estructura de dominios algebraicos. Estos son intervalos racionales representados por una pareja de números diádicos, el punto central y el radio del intervalo. En el siguiente capítulo estudiaremos representaciones de los números reales por LFTs cuyas implementaciones han sido desarrolladas por Edalat y su estructura corresponde a los dominios continuos.

3.2. Preeliminares

La fundamentación teórica para la computación real exacta y para esta aproximación a un tipo de dato real, se conoce como análisis computable. Hay varias aproximaciones cuando se escogen representaciones concretas de números reales.

Una aritmética real exacta es un álgebra de operaciones exactas sobre un subconjunto A de los números reales. Para ser más especifico el álgebra podría ser de la forma

$$(A, +, \cdot, \ldots, 0, 1, \ldots).$$

El álgebra podría contener otras operaciones y constantes tales como $\sin, \log, \pi$ y e.

La representación que estudiaremos para los números reales es la intersección de intervalos racionales que disminuyen rápidamente de radio. Los puntos centrales de estas secuencias de intervalos son una secuencia de Cauchy convergente al número real. Estos números reales son secuencias de parejas de numeros racionales, o funciones de $\mathbb{N}$ en $\mathbb{Q}^2$.

Los elementos del álgebra de representación C son funciones, $C \subseteq (\mathbb{N} \to \mathbb{Q}^2)$. En el álgebra C hay representaciones de las operaciones en A, por ejemplo, $+ : C^2 \to C$ representa la operación $+ : A^2 \to A$.

$$C = (C, +, -, \cdot, \ldots, \hat{0}, \hat{1}, \ldots).$$

La noción de representación es modelado por un homomorfismo

$$\varphi : C \to A.$$

Asumimos que existe una función const:$\mathbb{Q} \to C$ que a cada número racional le asigna un número real, con el mismo valor. Para acceder a un número real asumimos que existe una operación

$$\text{aprox} : C \times \mathbb{N} \to \mathbb{Q}^2.$$

La operación aprox extrae el n-ésimo intervalo en la secuencia, que tiene radio menor o igual que 2^{-n}.

3.3. Intervalos Centrados Diádicos como Aproximaciones

Definición 3.3.1. Un intervalo centrado diádico es representado por una tripla (m, e, s) de la forma

$$(m \pm e)2^{-s},$$

donde la mantisa m y el exponente s son enteros y el término del error e es un número natural.

Un número real x es aproximado por $a = (m \pm e)2^{-s}$ si

$$|x - m2^{-s}| \leq e2^{-s}$$

o equivalentemente, si

$$x \in [(m - e)2^{-s}, (m + e)2^{-s}].$$

Fijando j, una j-aproximación centrada diádica es un intervalo centrado diádico donde el término del error esta estrictamente acotado por 2^j.

Estas aproximaciones pueden ser vistas como intervalos diádicos o números de punto flotante con cotas de error. La inclusión de términos de error en la aproximación es lo que nos permite asegurar que estamos haciendo aritmética real exacta. Se asume implícitamente que el valor de j de una j-aproximación centrada diádica es el mejor posible, es decir, una j-aproximación no es en general una $j + 1$-aproximación.

3.4. Estructura de las Aproximaciones

Por el Lema 2.4.2 sabemos que los intervalos racionales forman un cusl. La completación ideal de este cusl es el dominio algebraico de intervalos. Ahora nos restringimos a los intervalos racionales que son intervalos centrados diádicos.

Lema 3.4.1. *Los intervalos centrados diádicos forman un sub-cusl de los intervalos racionales.*

Demostración. Solo necesitamos verificar que el supremo de dos intervalos centrados diádicos consistente es de nuevo un intervalo centrado diádico. El supremo es la intersección de los dos intervalos. El extremo izquierdo de la intersección es uno de los extremos izquierdos de los dos intervalos, luego el extremo izquierdo es diádico. De manera similar se concluye que el extremo derecho es diádico. El punto medio entre dos números diádicos es diádico. La mitad de la distancia entre dos puntos diádicos es también diádica. Luego la intersección es un intervalo centrado diádico. □

Sin embargo, restringirnos a intervalos centrados diádicos implica una perdida de la propiedad del supremo.

Ejemplo 3.4.2. La menor cota superior de dos aproximaciones es la intersección de los intervalos, que de nuevo es un intervalo diádico. Sin embargo, en el caso de un término de error acotado, este intervalo diádico podría no ser representable con un término del error dentro de la cota. Por ejemplo considere las j-aproximaciones centradas diádicas para $j > 1$. La menor cota superior de $(0 \pm (2^j - 1))2^0$ y $(1 \pm (2^j - 1))2^0$ es el intervalo representado por $(1 \pm (2^{j+1} - 1))2^{-1}$, que claramente no puede ser representado con un término del error dentro de 2^j. Luego las aproximaciones no forman un cusl.

3.5. Computando una operación real exacta

Para llevar a cabo una computación exacta eficientemente debemos tener mucho cuidado en la elección de la estrategia a ser usada en el refinamiento de la computación y debemos notar la importancia de encontrar la aproximación más ajustada para cada paso durante la computación.

Definición 3.5.1. Sea $\mathbf{a}$ una k-tupla de intervalos centrados diádicos de números reales. Una operación $f : \mathbb{R}^k \to \mathbb{R}$ es una implementación óptima si existe una j-aproximación centrada diádica maximal (con respecto a $\sqsubseteq$) b que representa todo punto en $f(\mathbf{a})$.

Note que solo requerimos que la aproximación retornada sea maximal. Esto es porque no necesariamente existe un supremo entre las j aproximaciones centradas diádicas (ver sección 3.4).

Claramente, no todas las operaciones tienen implementaciones óptimas dado que el intervalo imagen no siempre es un intervalo centrado diádico.

Definición 3.5.2. Sea $\mathbf{a}$ una k-tupla de intervalos centrados diádicos de números reales. Una operación $f : \mathbb{R}^k \to \mathbb{R}$ es una implementación *casi óptima* si para cualquier número diádico $d > \mathrm{diam} f(\mathbf{a})$ puede ser computado un intervalo centrado diádico b tal que $f(\mathbf{a}) \subseteq b$, y $\mathrm{diam}(b) \leq d$, donde diam devuelve el diámetro de un intervalo.

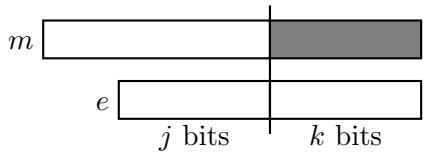
3.5.1. Redondeo

La operación de encontrar la mejor j-aproximación centrada diádica desde un intervalo centrado diádico será llamada redondeo. Esta es una aproximación importante dado que asumimos que todo resultado intermedio será redondeado después de cada operación.

Proposición 3.5.3. *Dado un intervalo centrado diádico* $(m \pm e)2^{-s}$ *existe una óptima* $j-$*aproximación centrada diádica.*

Demostración. Necesitamos dar una aproximación de la forma $(n \pm e')2^{-t}$ donde el término del error e es acotado estrictamente por 2^j, es decir, encaja en j bits. Necesitamos cortar m y e de manera que el nuevo término de error ajustado por el error de redondeo introducido (la zona gris) sea lo suficientemente pequeño. Sea q el número de bits en la representación de e, un corte de al menos $k = q - j$ bits es requerido. Sean

$$t = s - k$$
$$n = \mathrm{rnd}(m/2^k)$$



y r el error de redondeo introducido

$$r = |m - n2^k|.$$

El error de redondeo es a lo más 2^{k-1}. Creamos e' por redondeo hacia arriba del término de error total $e + r$ dividido por 2^k, es decir,

$$e' = \left\lceil \frac{e+r}{2^k} \right\rceil.$$

A partir de esta construcción es claro que $(m \pm e)2^{-s} \subseteq (n \pm e')2^{-t}$.

La representación de e' puede tener más que j bits. Si este es el caso el proceso debería ser repetido con los valores originales de m y e, y con k incrementado en 1. Debemos chequear que k no requiere ser incrementado indefinidamente, en efecto, a lo más dos incrementos serán necesarios, y esto sólo ocurrira cuando $j = 1$.

Considere primero cuando k ha sido incrementado una vez, entonces $k = q - j + 1$. Tenemos que $r \leq 2^{k-1}$ y $e < 2^q = 2^{j+k-1}$ y entonces

$$\begin{aligned} e' &= \left\lceil \frac{e+r}{2^k} \right\rceil \\ &\leq \left\lceil \frac{2^q + 2^{k-1}}{2^k} \right\rceil \\ &\leq \left\lceil \frac{2^{j+k-1}}{2^k} \right\rceil + \left\lceil \frac{2^{k-1}}{2^k} \right\rceil \\ &\leq 2^{j-1} + 1, \end{aligned}$$

que es menor que 2^j si $j > 1$. Para $j = 1$, considere entonces que k ha sido incrementado 2 veces, esto es $k = q - j + 2 = q + 1$, entonces

$$\begin{aligned} e' &= \left\lceil \frac{e+r}{2^k} \right\rceil \\ &\leq \left\lceil \frac{2^q + 2^{k-1}}{2^k} \right\rceil \\ &\leq \left\lceil \frac{2^{k-1} + 2^{k-1}}{2^k} \right\rceil \\ &\leq 1 \\ &< 2^j. \end{aligned}$$

La optimización se consigue mediante el cálculo de e' tan pequeño como sea posible y considerando el menor valor posible de k en primer lugar.

Tenga en cuenta que la aproximación devuelta no es única ya que n puede ser redondeado hacia arriba o hacia abajo si $m/2^k$ tiene parte fraccionaria exactamente $1/2$.

□

Ejemplo 3.5.4. Considere el redondeo del intervalo centrado diádico $(1280 \pm 257)2^{-10}$ donde $j = 1$. Tenemos $q = 9$, dado que la representación binaria de 257 es 100000001, entonces la primera opción para k es $k = 8$.

$k = 8$	$2^k = 256$	$n = 5$	$r = 0$	$e' = \lceil \frac{257+0}{256} \rceil = 2 \geq 2^1$
$k = 9$	$2^k = 512$	$n = 2$ ó 3	$r = 256$	$e' = \lceil \frac{257+256}{512} \rceil = 2 \geq 2^1$
$k = 10$	$2^k = 1024$	$n = 1$	$r = 256$	$e' = \lceil \frac{257+256}{1024} \rceil = 1 < 2^1$

Entonces el intervalo de aproximación resultante es $(1 \pm 1)2^0$. Esta es la mejor 1-aproximación centrada diádica que contiene el intervalo dado.

3.5.2. Adición

Proposición 3.5.5. *Existe una implementación óptima de la adición de j aproximaciones centradas diádicas.*

Demostración. Sean $a = (m \pm e)2^{-s}$ y $b = (n \pm f)2^{-t}$ intervalos centrados diádicos, y asumamos sin perdida de generalidad que $s \geq t$, la imagen intervalar exacta de su suma es

$$a + b = (m + n2^{s-t} \pm (e + f2^{s-t}))2^{-s}.$$

Una óptima j-aproximación centrada diádica se obtiene redondeando este intervalo centrado diádico usando la Proposición 3.5.3.

□

Para comparar la precisión de la entrada y la salida de las operaciones usamos la siguiente terminología.

Definición 3.5.6. Si una operación toma una q-aproximación como entrada y retorna una p-aproximación en la salida, entonces la operación resultante tiene una perdida de $q-p$ bits. Si $q-p$ es negativo es más natural decir que ganó $p-q$ bits.

Ejemplo 3.5.7. Calcularemos la perdida de dígitos en la suma, asumiendo que todas las aproximaciones tienen un término de error 1. Si p es la precisión deseada del resultado y q es la precisión de la entrada. El valor de q se calcula a partir de p. En efecto $q = p+2$ es suficiente. Suponga que las aproximaciones para x y y son $(m \pm 1)2^{-q}$ y $(n \pm 1)2^{-q}$ respectivamente. La suma de estas aproximaciones es

$$(m \pm 1)2^{-q} + (n \pm 1)2^{-q} = (m+n \pm 2)2^{-p-2}.$$

Si $m+n$ es par entonces esta expresión es igual a $(\frac{1}{2}(m+n) \pm 1)2^{-p-1}$ que es una $p+1$ aproximación de la suma.

Si $m+n$ es impar entonces la mejor aproximación posible con término del error 1 tiene mantisa $\frac{1}{4}(m+n)$ redondeado al entero más cercano. En efecto, si $m+n$ es impar entonces $m+n = 4k+1$ ó $m+n = 4k+3$ para algún $k \in \mathbb{Z}$. En el primer caso $\mathrm{rdn}(\frac{m+n}{4}) = k$ y $(m+n\pm 2)2^{-p-2} = (k-\frac{1}{4}, k+\frac{3}{4})2^{-p} \subset (k-1, k+1)2^{-p}$. En el segundo caso $\mathrm{rdn}(\frac{m+n}{4}) = k+1$ y $(m+n \pm 2)2^{-p-2} = (k+1-\frac{3}{4}, k+1+\frac{1}{4})2^{-p} \subset ((k+1)-1, (k+1)+1)2^{-p}$. Se puede verificar que no es posible hacer una $(p+1)$-aproximación con término de error 1. En conclusión la mejor aproximación posible con termino de error 1 es la p-aproximación $(\mathrm{rdn}(\frac{1}{4}(m+n)) \pm 1)2^{-p}$.

Así en la mitad de los casos la adición pierde un bit y en la otra mitad perderá dos bits de precisión.

3.5.3. Multiplicación

Proposición 3.5.8. *Existe una implementación optima de la multiplicación de j-aproximaciones centradas diádicas.*

El intervalo imagen exacto de la multiplicación de los intervalos centrados diádicos $a = (m \pm e)2^{-s}$ y $b = (n \pm f)2^{-t}$ puede ser obtenido considerando 11 casos.

$$\begin{cases}
\textbf{(1.1)}\ (mn + ef \pm (|mf| + |ne|))2^{-s-t}, \text{ si } |m| \geq e,\ |n| \geq f,\ mn > 0 \\
\textbf{(1.2)}\ (mn - ef \pm (|mf| + |ne|))2^{-s-t}, \text{ si } |m| \geq e,\ |n| \geq f,\ mn < 0 \\
\textbf{(2.1)}\ (mn + mf \pm (|ne| + ef))2^{-s-t}, \text{ si } |m| < e,\ n \geq f \\
\textbf{(2.2)}\ (mn - mf \pm (|ne| + ef))2^{-s-t}, \text{ si } |m| < e,\ n \leq -f \\
\textbf{(3.1)}\ (mn + ne \pm (|mf| + ef))2^{-s-t}, \text{ si } m \geq e,\ |n| < f \\
\textbf{(3.2)}\ (mn - ne \pm (|mf| + ef))2^{-s-t}, \text{ si } m \leq -e,\ |n| < f \\
\textbf{(4.1)}\ (mn + |mf| \pm (|ne| + ef))2^{-s-t}, \text{ si } |m| < e,\ |n| < f,\ |mf| \leq |ne|,\ mn > 0 \\
\textbf{(4.2)}\ (mn - |mf| \pm (|ne| + ef))2^{-s-t}, \text{ si } |m| < e,\ |n| < f,\ |mf| \leq |ne|,\ mn < 0 \\
\textbf{(4.3)}\ (mn + |ne| \pm (|mf| + ef))2^{-s-t}, \text{ si } |m| < e,\ |n| < f,\ |mf| > |ne|,\ mn > 0 \\
\textbf{(4.4)}\ (mn - |ne| \pm (|mf| + ef))2^{-s-t}, \text{ si } |m| < e,\ |n| < f,\ |mf| > |ne|,\ mn < 0 \\
\textbf{(5)}\ (0 \pm (|mf| + |ne| + ef))2^{-s-t}, \text{ si } mn = 0.
\end{cases}$$

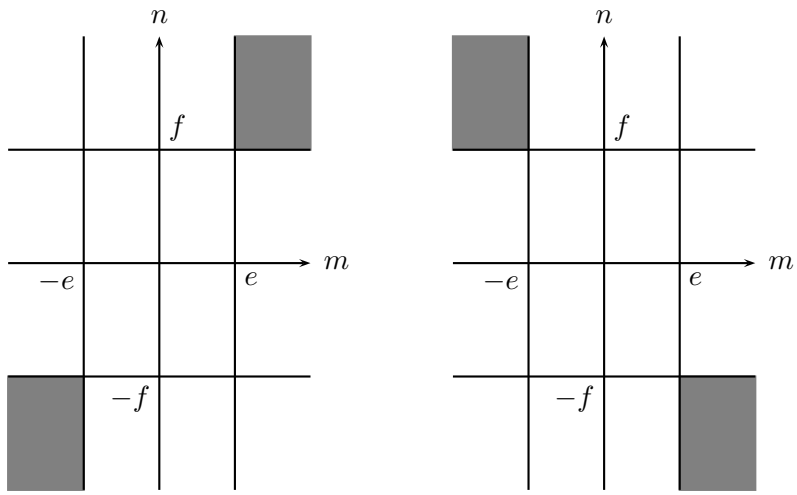
Demostración. Estos casos se obtienen a partir de la definición de multiplicación de intervalos de Moore

$$[\underline{x}, \overline{x}] \times [\underline{y}, \overline{y}] = [\text{mín}\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\}, \text{máx}\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\}].$$

En adelante denotaremos MIN y MAX la menor y mayor combinación posible respectivamente para el producto de los intervalos centrados diádicos a y b en términos de la mantisa y el error.

$$[m-e, m+e]2^{-s} \times [n-e, n+e]2^{-t} = [\text{MIN}, \text{MAX}]2^{-s-t}.$$

1. En **(1.1)** y **(1.2)** se tiene $|m| \geq e \Leftrightarrow (m-e \geq 0 \;\text{ ó }\; m+e \leq 0)$, luego los extremos del intervalo $[m-e, m+e]$ son ambos positivos ó ambos negativos. Análogamente para $|n| \geq f$.



caso (1.1) **caso (1.2)**

Si los signos de m y n son iguales, se tienen las dos siguientes opciones

a) Si $m-e \geq 0$ y $n-f \geq 0$ entonces MIN= $(m-e)(n-f) = mn-mf-ne+ef$ y MAX= $(m+e)(n+f) = mn+mf+en+ef$, como el punto medio es $mn+ef$ tenemos el intervalo centrado diádico $(mn+ef \pm (ne+mf))$.

b) Si $m+e \leq 0$ y $n-f \leq 0$ entonces MAX= $(m-e)(n-f) = mn-mf-ne+ef$ y MIN= $mn+mf+en+ef$, como el punto medio es $mn+ef$ tenemos el intervalo centrado diádico $(mn+ef \pm (-ne-mf))$.

Los intervalos obtenidos en *a*) y *b*) se pueden unificar en el intervalo $(mn+ef \pm (|ne|+|mf|))$ que resume el caso **(1.1)**, (tenga en cuenta que e y f son positivos por definición).

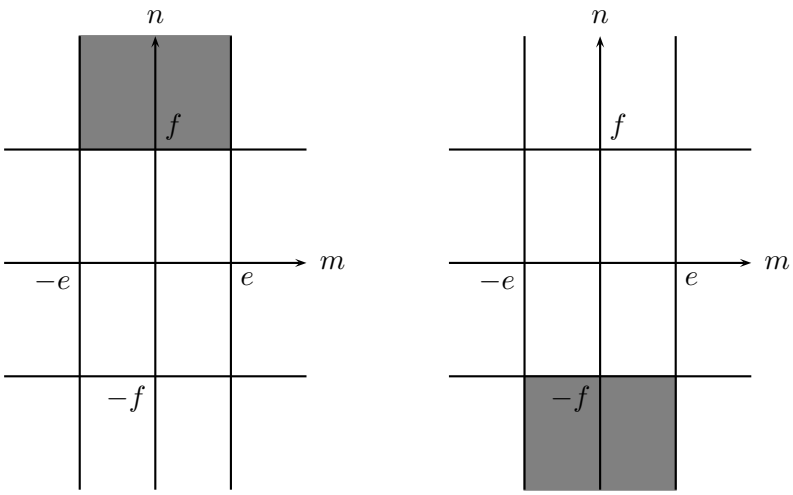
Para m y n con signo contrario tenemos

a) Si $m-e \geq 0$ y $n-f \leq 0$ entonces MIN= $(m+e)(n-f) = mn-mf+ne-ef$ y MAX= $(m-e)(n+f) = mn+mf-en-ef$, con punto medio $mn-ef$, luego el intervalo centrado diádico es $(mn-ef \pm (-ne+mf))$.

b) Si $m+e \leq 0$ y $n-f \geq 0$ entonces MAX= $(m+e)(n-f) = mn-mf+ne-ef$ y MIN= $(m-e)(n+f) = mn+mf-en-ef$, con punto medio $mn-ef$, luego el intervalo centrado diádico es $(mn-ef \pm (ne-mf))$.

Los intervalos obtenidos en *a*) y *b*) se pueden unificar en el intervalo $(mn-ef \pm (|ne|+|mf|))$ que resume el caso **(1.2)**.

2. En **(2.1)** y **(2.2)** $|m| < e$, lo cual significa que el extremo izquierdo del intervalo $[m-e, m+e]$ es negativo y su extremo derecho es positivo.

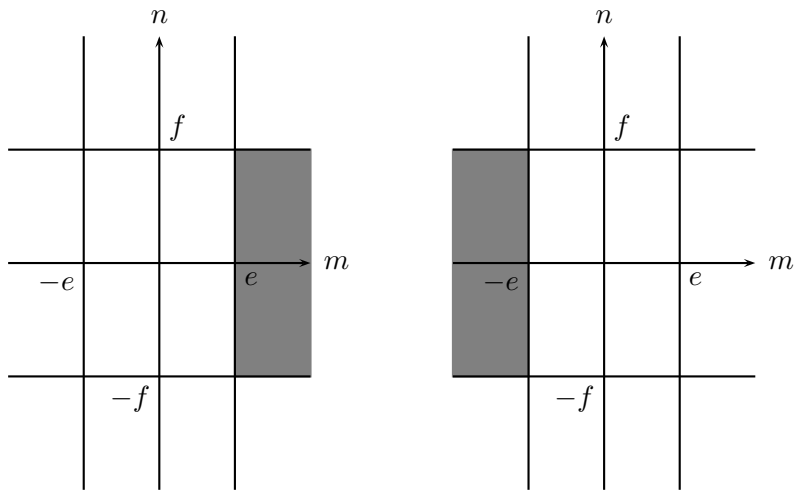


caso (2.1) **caso (2.2)**

Para **(2.1)** los extremos del intervalo $[n-f, n+f]$ son positivos, luego MIN= $(m-e)(n+f) = mn+mf-en-ef$ y MAX= $(m+e)(n+f) = mn+mf+en+ef$, que corresponde al intervalo centrado diádico $(mn+mf \pm (en+ef)) = (mn+mf \pm (|en|+ef))$.

Para **(2.2)** los extremos del intervalo $[n-f, n+f]$ son negativos, luego MIN= $(m+e)(n-f) = mn-mf+en-ef$ y MAX= $(m-e)(n-f) = mn-mf-en+ef$, que corresponde al intervalo centrado diádico $(mn-mf \pm (-en+ef)) = (mn-mf \pm (|en|+ef))$.

3. **(3.1)** y **(3.2)** son análogos a **(2.1)** y **(2.2)** intercambiando los papeles de los intervalos a y b.



caso (3.1) **caso (3.2)**

4. En **(4.1)**, **(4.2)**, **(4.3)** y **(4.4)** una de las condiciones es $|m| < e$ y $|n| < f$, esto se traduce en que los extremos izquierdos de ambos intervalos son negativos y los extremos derechos de ambos intervalos son positivos. Por tanto, en principio hay dos posibilidades para MAX y dos posibilidades para MIN:

$$\text{MAX} = \begin{cases} (m-e)(n-f) = mn - mf - ne + ef \\ (m+e)(n+f) = mn + mf + en + ef \end{cases} = mn + ef + u,$$

donde

$$u = \text{máx}\{-mf - en,\, mf + en\};$$

$$\text{MIN} = \begin{cases} (m-e)(n+f) = mn + mf - ne - ef \\ (m+e)(n-f) = mn - mf + en - ef \end{cases} = mn - ef + v,$$

donde

$$v = \text{mín}\{mf - ne,\, -mf + ne\}.$$

Se puede determinar u para los cuadrantes I y III, y v para los cuadrantes II y IV. En efecto:

u es $mf + en$ si $m > 0$ y $n > 0$, y $-mf - en$ si $m < 0$ y $n < 0$, es decir,

$$\text{MAX} = mn + ef + |mf| + |ne|, \text{ si } mn > 0. \tag{3.1}$$

v es $mf - en$ si $m < 0$ y $n > 0$, y $-mf + en$ si $m > 0$ y $n < 0$, es decir,

$$\text{MIN} = mn - ef - |mf| - |ne|, \text{ si } mn < 0. \tag{3.2}$$

Para determinar u en los cuadrantes II y IV se debe conocer el signo de $mf + ne$ y para determinar v en los cuadrantes I y III se debe conocer el signo de $mf - ne$:

a) En II, $-mf \leq ne$, luego u es $ne+mf$. En IV, $mf \leq -ne$, luego u es $-mf-ne$. Es decir,

$$\text{MAX} = mn + ef - |mf| + |ne|, \text{ si } |mf| \leq |ne| \text{ y } mn < 0. \tag{3.3}$$

En I, $mf < ne$, luego v es $mf - en$. En III, $-mf < -en$, luego v es $en - mf$. Es decir,

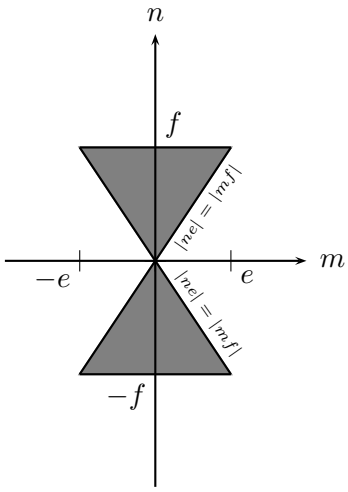
$$\text{MIN} = mn - ef + |mf| - |ne|, \text{ si } |mf| \leq |ne| \text{ y } mn > 0. \tag{3.4}$$

b) En II, $-mf > ne$, entonces u es $-mf - en$. En IV, $mf > -ne$, entonces u es $mf + en$. En conclusión

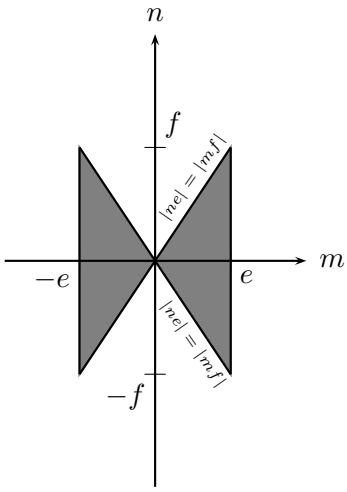
$$\text{MAX} = mn + ef + |mf| - |ne|, \text{ si } |mf| \geq |ne| \text{ y } mn < 0. \tag{3.5}$$

En I, $mf > ne$, entonces v es $-mf + ne$. En III $-mf > -ne$, entonces v es $mf - en$. Así

$$\text{MIN} = mn - ef - |mf| + |ne|, \text{ si } |mf| \geq |ne| \text{ y } mn > 0. \tag{3.6}$$



casos (4.1) y (4.2)



casos (4.3) y (4.4)

Unificando (3.1) y (3.4) se obtiene el intervalo centrado diádico $(mn+|mf|\pm(|en|+ef))$ (caso **(4.1)**). Unificando (3.2) y (3.3) se concluye $(mn-|mf|\pm(|ne|+ef))$ (el caso **(4.1)**). De (3.1) y (3.6) se tiene $(mn+|ne|\pm(|mf|+ef))$ (el caso **(4.3)**). Por último de (3.2) y (3.5) se obtiene el intervalo centrado diádico $(mn-|ne|\pm(|mf|+ef))$ (caso **(4.4)**).

5. Si $m = 0$, MAX= $|ne| + ef$ y MIN=$-|en| - ef$. Análogamente si $n = 0$, MAX= $|mf| + ef$ y MIN=$-|mf| - ef$. Unificando en un sólo caso, si $mn = 0$ el intervalo centrado diádico es $(0 \pm (|mf| + |en| + ef))$ (caso **(5)**).

De nuevo una j-aproximación centrada diádica se obtiene redondeando uno de los intervalos según el caso, usando la Proposición 3.5.3.

□

CAPÍTULO 4

Transformaciones Lineales Fraccionarias

4.1. Introducción

Un número x es computable si y sólo si el conjunto $\{x\}$ es el extremo superior de una cadena ascendente efectiva en $\mathbb{IR}$, es decir, si este es la intersección de una cadena efectiva contractante de intervalos anidados. En este capítulo mostraremos como obtener tal secuencia de intervalos contractantes.

Un sistema de funciones iteradas IFS (iterated functions system) en A es un conjunto finito de funciones computables $f_i : A \to A$, $i \in K$, donde K es un conjunto finito de indices. Asumiremos que todas las funciones son contracciones. Entonces para toda función total recursiva $h : \mathbb{N} \to K$, la secuencia

$$\begin{aligned}
J_0 &:= A \\
J_1 &:= f_{h(0)}(A) \\
J_2 &:= f_{h(0)}(f_{h(1)}(A)) \\
&\vdots \\
J_n &:= f_{h(0)}(f_{h(1)}(\cdots f_{h(n)}(A)\cdots)) \\
&\vdots
\end{aligned}$$

es una secuencia contractante de intervalos y define un elemento computable x_h de A con $\{x_h\} = \bigcap_{n\in\mathbb{N}} J_n$. El intervalo J_n contiene a x_h y este es considerado como una aproximación de este punto.

Para ser más específicos, si $A = [0, 1]$, $K = \{0, \ldots, 9\}$ y

$$f_i(x) = \frac{x+i}{10}$$

entonces la secuencia $h(0), h(1), \ldots$ es la expansión decimal del número real x_h. Como un ejemplo concreto tome h tal que $h(i)$ es el i-ésimo lugar decimal del número π. Entonces

$$
\begin{aligned}
J_0 &:= [0,1] \\
J_1 &:= f_3([0,1]) = [0.3, 0.4] \\
J_2 &:= f_3(f_1([0,1])) = f_3([0{,}1, 0{,}2]) = [0.31, 0.32] \\
J_3 &:= f_3(f_1(f_4([0,1]))) = f_3(f_1([0.4, 0.5])) = f_3([0.14, 0.15]) = [0.314, 0.315] \\
&\vdots
\end{aligned}
$$

Expansiones similares con respecto a bases diferentes pueden expresarse por un IFS. También se pueden permitir que los dígitos sean negativos para que los f_i tengan traslapamiento de rangos al aplicarse a intervalos.

Ejemplo 4.1.1. Sea $K = \{-1, 0, 1\}$ y

$$f_i(x) = \frac{x+i}{2},\ i \in K$$

para el intervalo $A = [-1, 1]$. Este IFS produce una expansión binaria signada (esto da origen a una representación redundante de los números reales, es decir, hay más de una representación para cada número real). Así, denotando -1 con $\overline{1}$, tenemos:

- Si el primer dígito es $\overline{1}$ y los siguientes son desconocidos, se tiene un número x de la forma $0.\overline{1}__\ldots$ que satisface la desigualdad

$$-1 = -\frac{1}{2} - \sum_{n=2}^{\infty} \frac{1}{2^n} \leq x \leq -\frac{1}{2} + \sum_{n=2}^{\infty} \frac{1}{2^n} = 0,$$

es decir, $x \in [-1, 0]$. El intervalo obtenido también se puede ver como la imagen del intervalo base $[-1, 1]$ por f_{-1}

- Si el primer dígito es 0 y los siguientes son desconocidos, se tiene un número x de la forma $0.0__\ldots$ que satisface la desigualdad

$$-\frac{1}{2} = -\sum_{n=2}^{\infty} \frac{1}{2^n} \leq x \leq \sum_{n=2}^{\infty} \frac{1}{2^n} = \frac{1}{2},$$

es decir, $x \in [-1/2, 1/2] = f_0[-1, 1]$

- Si el primer dígito es 1 y los siguientes son desconocidos, se tiene un número x de la forma $0.1__\ldots$ que satisface la desigualdad

$$0 = \frac{1}{2} - \sum_{n=2}^{\infty} \frac{1}{2^n} \leq x \leq \frac{1}{2} + \sum_{n=2}^{\infty} \frac{1}{2^n} = 1,$$

es decir, $x \in [0, 1] = f_1[-1, 1]$

Así, para un número como $x = 0.11\overline{1}0__\ldots$ su primer dígito indica que se encuentra en el intervalo $[0, 1]$. Sus primeros dos dígitos indican que satisface la desigualdad

$$\frac{1}{2} = \frac{3}{4} - \sum_{n=3}^{\infty} \frac{1}{2^n} \leq x \leq \frac{3}{4} + \sum_{n=3}^{\infty} \frac{1}{2^n} = 1,$$

es decir, $x \in [1/2, 1]$. Sus primeros tres dígitos indican que satisface la desigualdad

$$\frac{1}{2} = \frac{5}{8} - \sum_{n=4}^{\infty} \frac{1}{2^n} \leq x \leq \frac{5}{8} + \sum_{n=4}^{\infty} \frac{1}{2^n} = \frac{3}{4},$$

es decir $x \in [1/2, 3/4]$. Sus cuatro dígitos conocidos tienen la información

$$\frac{9}{16} = \frac{5}{8} - \sum_{n=5}^{\infty} \frac{1}{2^n} \leq x \leq \frac{5}{8} + \sum_{n=5}^{\infty} \frac{1}{2^n} = \frac{11}{16},$$

$x \in [9/16, 11/16]$.

Y también se puede expresar como composición de LFTs:

$$\begin{aligned}
J_0 &:= [-1, 1] \\
J_1 &:= f_1([-1, 1]) = [0, 1] \\
J_2 &:= f_1(f_1([-1, 1])) = f_1([0, 1]) = [1/2, 1] \\
J_3 &:= f_1(f_1(f_{-1}([-1, 1]))) = f_1(f_1([-1, 0])) = f_1([0, 1/2]) = [1/2, 3/4] \\
J_4 &:= f_1(f_1(f_{-1}(f_0([-1, 1])))) = f_1(f_1(f_{-1}([-\frac{1}{2}, \frac{1}{2}]))) = f_1(f_1([-\frac{3}{4}, -\frac{1}{4}])) \\
&= f_1([\frac{1}{8}, \frac{3}{8}]) = [\frac{9}{16}, \frac{11}{16}].
\end{aligned}$$

Ejemplo 4.1.2. Considere el computo de $y = 3 \cdot x$, donde x esta dado por la representación decimal de $1/3$, $\xi = 0.333\cdots$. Matemáticamente el resultado esta dado por $\eta = 0.999\cdots = 1$, pero es posible computar este resultado?. Recalcamos que en todo momento sólo se conoce un prefijo finito de ξ, y este prefijo finito es la única fuente de información disponible para producir un prefijo finito del resultado. Es imposible obtener $0.999\cdots$ en el resultado pues tiene una cantidad infinita de dígitos, ahora es posible obtener 1?

Asuma que conocemos el prefijo 0.333 de ξ. Esto es suficiente para determinar los primeros dígitos de η? Desafortunadamente no, porque el prefijo 0.333 denota el intervalo $[0.333, 0.334]$, que produce $[0.999, 1.002]$ cuando es multiplicado por 3. Ahora sabemos que η podría empezar con 0. ó 1., pero aún no sabemos cuál es correcto, dado que ni el intervalo [0,1] denotado por "0." ni el intervalo [1,2] denotado por "1." cubre el intervalo salida $[0.999, 1.002]$. Peor aún, es fácil ver que esto ocurre con todos los prefijos de la forma $0.333\cdots 3$. Por tanto si ξ es la cadena con infinitos '3's nunca podríamos tener el primer dígito de η, dado que ninguna cantidad finita de información de ξ es suficiente para decidir si η empezaría con 0 o 1. Esto significa que la multiplicación no es computable en el sistema decimal.

Una solución a este problema es admitir dígitos negativos ($-1, -2, \ldots, -9$ en base 10). Si sabemos que ξ comienza con 0.333 la salida podría ser 1. (incluso 1.00) como un prefijo

de η dado que nosotros podemos compensar los dígitos negativos si resultara después que el número representado por ξ es menor que 1/3, y entonces el resultado es más pequeño que 1. Formalmente el intervalo denotado por el prefijo 0.333 es ahora [0.332,0.334], dado que la extensión más pequeña posible de 0.333 ya no es 0.33300..., sino $0.333\overline{99}\ldots$ ($\overline{9}$ denota -9). Este intervalo nos lleva a [0.996,1.002] cuando es multiplicado por 3, el cual esta contenido en el intervalo [0.99, 1.01] que representa el prefijo 1.00, es decir, podemos poner con seguridad 1.00 como el comienzo de nuestra cadena de salida.

Cualquier fracción continua de la expansión de un número real puede ser expresada como una composición infinita de LFTs (lineal fractional transformations) de la forma

$$f(x) = \frac{ax+c}{bx+d}, \tag{4.1}$$

donde $a, b, c, d \in \mathbb{Z}$.

En efecto cualquier fracción continua

$$a_0 + \cfrac{b_0}{a_1 + \cfrac{b_1}{a_2 + \cfrac{b_2}{a_3 + \cfrac{b_3}{\ddots}}}}$$

expansión de un número r, puede expresarse como $r = \phi_0(r_0)$ con

$$r_0 = a_1 + \cfrac{b_1}{a_2 + \cfrac{b_2}{a_3 + \cfrac{b_3}{\ddots}}}$$

y $\phi_0(x) = a_0 + \frac{b_0}{x}$. Iterando este esquema obtenemos $r = \phi_0(\phi_1(\cdots \phi_n(r_n)))$ con

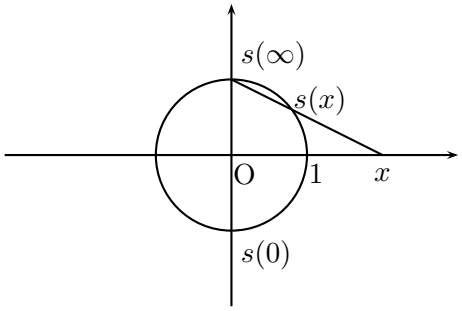
$$r_n = a_{n+1} + \cfrac{b_{n+1}}{a_{n+2} + \cfrac{b_{n+2}}{a_{n+3} + \cfrac{b_{n+3}}{\ddots}}}$$

y $\phi_i(x) = a_i + \frac{b_i}{x}$ $0 \leq i \leq n$.

4.2. Representación de los Números Reales

Trabajamos con los números reales extendidos ($\infty = \frac{r}{0},\ r \neq 0$ es un posible resultado). Un modelo simple de los reales extendidos con un único punto de compactificación $\mathbb{R}^* = \mathbb{R} \cup \{\infty\}$ es el circulo unitario S en el plano con centro en el origen equipado con la topología de subespacio euclidiana del plano. Dado cualquier punto $x \in \mathbb{R}$ sobre el eje

horizontal, la linea que une el punto máximo del circulo S y x intersecta a S en un único punto $s(x)$ como muestra la figura. Definimos $s(\infty)$ como el punto máximo del circulo. La función $s : \mathbb{R}^* \to S$ es un homeomorfismo conocido como la proyección estereográfica.



El orden usual en los reales induce una orientación anti-horaria en S. El intervalo $[a, b]$ es definido como el arco cerrado que va en sentido antihorario desde a hasta b. Una métrica en $\mathbb{R}^*$ obvia compatible con su topología sería tomar la distancia entre $x, y \in \mathbb{R}^*$ como la longitud del arco más corto que conecta $s(x)$ con $s(y)$ en S. Sin embargo, esta métrica involucra funciones trascendentales. Una métrica ρ en $\mathbb{R}^*$ que es facil de computar se define como sigue. Para reales extendidos x y y ambos no positivos ó ambos no negativos,

$$\rho(x, y) = \left| \frac{|x| - 1}{|x| + 1} - \frac{|y| - 1}{|y| + 1} \right|,$$

en otro caso, si x y y tienen signos diferentes

$$\rho(x, y) = \text{mín}\{\rho(x, 0) + \rho(0, y), \rho(x, \infty) + \rho(\infty, 0)\}.$$

Similar al termino $r/0$ con $r \neq 0$, podríamos tener expresiones como $\infty - \infty$, $0/0$ y 0^0, todas serán denotadas por $\perp = \mathbb{R}^*$, $\mathbb{R}^*_\perp = \mathbb{R}^* \cup \{\perp\}$ es $\mathbb{R}^*$ más un valor indefinido $\perp$. Esto nos lleva naturalmente al dominio $\mathbb{IR}^* = \{[a, b] \subset \mathbb{R}^*\} \cup \{\mathbb{R}^*\}$ de intervalos de $\mathbb{R}^*$ ordenados por inclusión reversa.

Usaremos la clase de LFTs o transformaciones de Möbius con coeficientes reales para codificar una secuencia de intervalos encajados contractantes, y por lo tanto de cualquier número real.

El conjunto de lodos los LFTs de la forma (4.1) con $a, b, c, d \in \mathbb{R}$ y $ad - bc \neq 0$ (para que f no sea una función constante) es denotado por $\mathbb{M}$. Un LFT es un homeomorfismo de $\mathbb{R}^*$ (un LFT es continuo en $\mathbb{R}^*$ aunque no en $\mathbb{R}$); este preserva la orientación si $ad - bc > 0$ y cambia la orientación si $ad - bc < 0$.

Recalcamos algunas propiedades elementales de $\mathbb{M}$. Bajo la composición de funciones $\mathbb{M}$ es un grupo de homeomorfismos de $\mathbb{R}^*$. Si $GL(2, \mathbb{R})$ denota el grupo lineal general de matrices no singulares 2×2 con coeficientes reales, la función

$$\begin{aligned} \Theta : GL(2, \mathbb{R}) &\to \mathbb{M} \\ \begin{pmatrix} a & c \\ b & d \end{pmatrix} &\mapsto \phi(x) = \frac{ax + c}{bx + d}, \end{aligned}$$

es un homomorfismo de grupos. El kernel K de Θ consiste de todas las matrices de la forma λI donde $\lambda \neq 0$ e I es la matriz identidad. Luego $\mathbb{M} \cong GL(2,\mathbb{R})/K$. Esto significa que podemos identificar cualquier LFT de la forma (1) con una matriz 2×2 salvo producto por escalar. Mas aún $\mathbb{R}^*$ puede ser identificado con el conjunto de subespacios unidimensionales de $\mathbb{R}^2$. En efecto, cualquier subespacio V es generado por un vector $v = \begin{pmatrix} k \\ l \end{pmatrix} \in V$ con $k, l \in \mathbb{R}$ no ambos cero. La razón $k/l \in \mathbb{R}^*$ es independiente del vector $v \in V$ que se escoja. Entonces uno puede identificar V con k/l. El vector $\begin{pmatrix} k \\ l \end{pmatrix}$ se dice que representa $x = \frac{k}{l} \in \mathbb{R}^*$. Podemos normalizar este vector dividiendo por $\sqrt{k^2+l^2}$, para obtener en coordenadas homogéneas $\begin{pmatrix} \sin\alpha \\ \cos\alpha \end{pmatrix}$ con $\tan\alpha = k/l$, α es el ángulo $s(0)\widehat{s(\infty)}s(x)$ entre los segmentos de línea $\overline{s(\infty)s(0)}$ y $\overline{s(\infty)s(x)}$, pues $\tan\alpha = \overline{Ox}/\overline{Os(\infty)}$. El ángulo $s(0)\widehat{Os}(x)$ entre los segmentos de linea $\overline{Os(0)}$ y $\overline{Os(x)}$ es 2α pues este describe el mismo arco que $s(0)\widehat{s(\infty)}s(x)$ pero su vértice esta en el centro y no sobre el circulo. Luego cuando α crece de 0 a π, x crece de 0 a ∞ y regresa a 0 a través de los números negativos, mientras que $s(x)$ da una vuelta a S desde $s(0)$ en el sentido anti-horario.

La acción de un LFT es equivalente a la multiplicación por una matriz. En efecto $\phi(\frac{k}{l}) = \frac{ak+cl}{bk+dl}$ puede escribirse como

$$\begin{pmatrix} k \\ l \end{pmatrix} \mapsto \begin{pmatrix} ak+cl \\ bk+dl \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} k \\ l \end{pmatrix}.$$

Luego podemos movernos libremente por un lado, entre $k/l \in \mathbb{R}^*$ y su representación homogénea $\begin{pmatrix} k \\ l \end{pmatrix}$, y por otro lado entre el LFT $x \to \dfrac{ax+c}{bx+d}$ y su matriz de representación $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$. En ambos casos la representación es única salvo multiplicación por un escalar.

Una propiedad básica del grupo $\mathbb{M}$ es que este es 3-transitivo, esto significa que para cualquier par de triplas de números distintos en $\mathbb{R}^*$ (x_1, x_2, x_3) y (y_1, y_2, y_3) $(i = 1, 2, 3)$, existe un único LFT $\phi \in \mathbb{M}$ con $y_i = \phi(x_i)$ $(i = 1, 2, 3)$. En efecto, para encontrar $\phi : x \to \dfrac{ax+c}{bx+d}$ que satisface $y_i = \phi(x_i)$ $(i = 1, 2, 3)$, se debe resolver el sistema de ecuaciones $ax_i + c = (bx_i + d)y_i$ $(i = 1, 2, 3)$ en las variables $a, b, c, d \in \mathbb{R}$ (si $x_i = \infty$ para algún i su ecuación correspondiente es $a = by_i$, si $y_i = \infty$ su correspondiente ecuación es $bx_i + d = 0$, si ambos son ∞ la ecuación es $b = 0$), que corresponde a la intersección de 3 hiperplanos independientes en $\mathbb{R}^4$, así que la solución es un conjunto afín de dimensión 1. Como el sistema es homogéneo (aunque la solución (0,0,0,0) esta restringida para el LFT), las soluciones en general están dadas por $(a, b, c, d) = \lambda(v_1, v_2, v_3, v_4)$ para (v_1, v_2, v_3, v_4) vector fijo y λ cualquier número real. Así que toda solución difiere sólo por un múltiplo escalar y por tanto el LFT es único.

Una consecuencia inmediata de la 3-transitividad de $\mathbb{M}$ es la siguiente propiedad.

Proposición 4.2.1. *Dados dos intervalos no triviales $[p, q]$ y $[r, s]$ con $p \neq q$ y $r \neq s$, existe un LFT $\phi \in \mathbb{M}$ tal que $\phi[p, q] = [r, s]$.*

De esto se sigue que si fijamos un intervalo base, entonces podemos expresar o codificar todos los intervalos no triviales como la imagen del intervalo base bajo un LFT. El intervalo base más eficiente es $[0, \infty]$ pues no es necesaria ninguna computación para determinar el LFT de la proposición. En efecto, $x \mapsto \frac{rx+s}{x+1}$ y $x \mapsto \frac{sx+r}{x+1}$ envían $[0, \infty]$ en $[r, s]$ el primero cambia la orientación, mientras que el segundo la preserva. Más aún, si $[r, s]$ es un intervalo racional $[\frac{a}{b}, \frac{c}{d}]$, entonces las funciones $x \mapsto \frac{ax+c}{bx+d}$ y $x \mapsto \frac{cx+a}{dx+b}$ con coeficientes enteros envían $[0, \infty]$ en $[\frac{a}{b}, \frac{c}{d}]$ cambiando y preservando la orientación respectivamente. Note que, si k y l son enteros positivos, las funciones $x \mapsto \frac{kax+lc}{kbx+ld}$ y $x \mapsto \frac{kcx+la}{kdx+lb}$ satisfacen la misma propiedad. Sin embargo el LFT será único bajo el cambio de orientación si pedimos que la suma de los valores absolutos de sus coeficientes sea minimal.

4.3. Semigrupos de LFTs de Refinamiento

Un LFT $\phi \in \mathbb{M}$ se dice que refina un intervalo $[p, q] \subset \mathbb{R}^*$ si $\phi[p, q] \subseteq [p, q]$ (ϕ puede preservar o cambiar la orientación del intervalo). Decimos que $S \subseteq \mathbb{M}$ refina $[p, q]$ si cada elemento de S refina $[p, q]$. Claramente, si S refina $[p, q]$, entonces el semigrupo generado por S también refina $[p, q]$. De esto se sigue que existe el más grande sub-semigrupo de $\mathbb{M}$ que refina un intervalo dado.

Considere el intervalo $[0, \infty]$. Sea $\mathbb{M}^+ \subseteq \mathbb{M}$ el conjunto de LFTs cuyos coeficientes son todos no negativos ó todos no positivos. Es fácil ver que $\mathbb{M}^+$ es el semigrupo de refinamiento de $[0, \infty]$. Más aún, tenemos:

Proposición 4.3.1. $[0, \infty]$ *es el único intervalo que tiene a* $\mathbb{M}^+$ *como su semigrupo de refinamiento.*

Demostración. Suponga que $\mathbb{M}^+$ además de ser el semigrupo de refinamiento de $[0, \infty]$, también es el semigrupo de refinamiento de un intervalo $A \in \mathbb{R}^*$, y sea $A = \phi[0, \infty]$ con $\phi : x \to \dfrac{ax + c}{bx + d}$. Por hipótesis, $\psi A \subseteq A$ $\forall \psi \in \mathbb{M}^+$. Todos los LFTs son monótonos en intervalos que no estén divididos por la asíntota vertical cuya imagen es ∞, y por tanto preservan la contenencia de conjuntos en dichos intervalos. Como $\psi A \subseteq A$, si $\psi A = A$ es inmediato $\phi^{-1}\psi\phi[0, \infty] = [0, \infty]$, y si la contenencia es propia $\psi[\frac{c}{d}, \frac{a}{b}] \subset [\frac{c}{d}, \frac{a}{b}]$ el conjunto $\psi[\frac{c}{d}, \frac{a}{b}]$ no contiene la asíntota vertical de $\phi^{-1} : x \to \dfrac{dx - c}{a - bx}$ que se encuentra en $\frac{a}{b}$ y la contenencia se preserva $\phi^{-1}\psi[\frac{c}{d}, \frac{a}{b}] \subset \phi^{-1}[\frac{c}{d}, \frac{a}{b}] = [0, \infty]$. Luego $\phi^{-1}\psi\phi[0, \infty] \subseteq [0, \infty]$ $\forall \psi \in \mathbb{M}^+$ ó equivalentemente $\gamma = \phi^{-1}\psi\phi \in \mathbb{M}^+$ $\forall \psi \in \mathbb{M}^+$. Tomando $\psi : x \to \dfrac{1}{x}$ tenemos

$$\gamma = \frac{(bd - ac)x + d^2 - c^2}{(a^2 - b^2)x + ac - bd},$$

lo cual implica que $ac - bd = 0$. Tomando $\gamma = x + 1$ tenemos

$$\gamma = \frac{(ad + bd - bc)x + d^2}{-b^2x + ad - bd - bc},$$

lo que significa que $b = 0$ o $d = 0$, y junto con la condición anterior $ac = 0$.

Si $b = 0$, entonces $a \neq 0$ y $d \neq 0$ pues esto contradice la definición de ϕ. Entonces $c = 0$ y $ad > 0$. Luego $\phi(x) = \dfrac{ax}{d}$ y por consiguiente $A = \phi[0, \infty] = [0, \infty]$.

Si $d = 0$, entonces $b \neq 0$ y $c \neq 0$ pues esto contradice la definición de ϕ.Entonces $a = 0$ y $bc > 0$. Luego $\phi(x) = \dfrac{c}{bx}$ y por consiguiente $A = \phi[0, \infty] = [\infty, 0]$.

□

¿Qué sucede si cambiamos el intervalo base?. Si B es cualquier intervalo no trivial, entonces $B = \phi[0, \infty]$ para algún LFT $\phi \in \mathbb{M}$ y tenemos:

Corolario 4.3.2. *El semigrupo de refinamiento de $B = \phi[0, \infty]$ esta dado por*

$$\phi \mathbb{M}^+ \phi^{-1} = \{\phi\psi\phi^{-1} \mid \psi \in \mathbb{M}^+\},$$

en otras palabras, $\gamma \in \mathbb{M}$ refina B si y sólo si $\phi^{-1}\gamma\phi \in \mathbb{M}^+$.

Este corolario implica que el semigrupo de refinamiento de $[0, \infty]$, denominado $\mathbb{M}^+$, se distingue por tener la caracterización más simple.

Proposición 4.3.3. *Para LFTs ϕ y ψ, tenemos $\phi[0, \infty] \supseteq \psi[0, \infty]$ si y sólo si $\psi = \phi\gamma$ con $\gamma \in \mathbb{M}^+$.*

Demostración. Si $\phi[0, \infty] = \psi[0, \infty]$, $[0, \infty] = \phi^{-1}\psi[0, \infty]$. Sea $\phi : x \to \dfrac{ax + c}{bx + d}$ y $\psi : x \to \dfrac{ex + g}{fx + h}$, entonces $\phi[0, \infty] \supset \psi[0, \infty]$ es equivalente a $[\frac{c}{d}, \frac{a}{b}] \supset [\frac{g}{h}, \frac{e}{f}]$, como $\psi[0, \infty]$ no contiene la asíntota vertical de ϕ^{-1}, entonces $[0, \infty] \supset \phi^{-1}\psi[0, \infty]$. Lo cual significa que $\phi^{-1}\psi \in \mathbb{M}^+$, o dicho en otras palabras, existe $\gamma \in \mathbb{M}^+$ tal que $\gamma = \phi^{-1}\psi$. □

De la proposición anterior se sigue que para cualquier secuencia encajada de intervalos contractantes $[p_0, q_0] \supseteq [p_1, q_1] \supseteq [p_2, q_2] \supseteq ...$, tenemos $[p_n, q_n] = \phi_0\phi_1 \cdots \phi_n[0, \infty]$ donde $\phi_0 \in \mathbb{M}$ y $\phi_i \in \mathbb{M}^+$ para $1 \leq i \leq n$. Luego, la secuencia puede expresarse como una composición infinita de LFTs, o equivalentemente como un producto infinito de matrices $\phi_0\phi_1\phi_2 \cdots$.

Un número racional puede representarse con un vector con coeficientes enteros. Un producto finito $\phi_0 \cdots \phi_n$ con $\phi_0 \in \mathbb{M}$, $\phi_i \in \mathbb{M}^+$ para $1 \leq i \leq n - 1$ y ϕ_n un vector con coeficientes no negativos, representa una secuencia de intervalos contractantes anidados que converge a un número racional.

Cualquier número real puede ser representado como la intersección $\bigcap_{n \geq 0} \phi_0\phi_1 \cdots \phi_n[0, \infty]$ con $\phi_0 \in \mathbb{M}$ y $\phi_i \in \mathbb{M}^+$ $(i \geq 1)$ tal que ϕ_n tiene coeficientes enteros para todo $n \geq 0$. Un número real puede expresarse como un producto infinito de matrices, llamado producto normal.

$$\begin{pmatrix} a_0 & c_0 \\ b_0 & d_0 \end{pmatrix} \begin{pmatrix} a_1 & c_1 \\ b_1 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & c_2 \\ b_2 & d_2 \end{pmatrix} \cdots$$

Productos parciales finitos dan aproximaciones del número real. Específicamente la primera matriz nos dice que el número se encuentra en el intervalo $[\frac{a_0}{b_0}, \frac{c_0}{d_0}]$ o $[\frac{c_0}{d_0}, \frac{a_0}{b_0}]$ de

acuerdo el signo del determinante de la matriz, el primer intervalo si el determinante es negativo y el segundo si el determinante es positivo. Las siguientes matrices refinarán sucesivamente este intervalo para dar una mejor aproximación del número real. La primera matriz se llama matriz signada mientras que las otras matrices son llamadas matrices dígito.

Por ejemplo la fracción continua para $\sqrt{2}$

$$1+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{\ddots}}}}$$

se puede expresar como el producto normal

$$\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 0\end{pmatrix}\cdots$$

Truncando el producto infinito y evaluando en el intervalo base $[0,\infty]$ obtenemos la siguiente secuencia de intervalos encajados anidados para aproximar $\sqrt{2}$ (teniendo en cuenta que $0=\frac{0}{1}\cong\begin{pmatrix}0\\1\end{pmatrix}$, $\infty=\frac{1}{0}\cong\begin{pmatrix}1\\0\end{pmatrix}$ y que los LFTs son monótonos, por tanto al aplicarlos a intervalos sólo es necesario aplicarlos a los extremos):

$$\begin{aligned}
\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}[0,\infty] &= [1,\infty]\\
\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 0\end{pmatrix}[0,\infty] &= [1,1.5]\\
\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 0\end{pmatrix}[0,\infty] &= [1.4,1.5]\\
\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 0\end{pmatrix}[0,\infty] &= [1.4,1.41\overline{6}]
\end{aligned}$$

4.4. Punto flotante exacto

Hasta el momento la representación de reales permite productos normales arbitrarios de matrices enteras $M_0M_1M_2\cdots$ con $M_0\in\mathbb{M}$ y $M_i\in\mathbb{M}^+$ para $i\geq 1$. Esto en la práctica presenta problemas, primero porque los intervalos podrían ser refinados en una razón arbitraria, haciendo el análisis de la complejidad de los algoritmos prácticamente imposible. En segundo lugar porque las matrices de multiplicación pueden producir rápidamente enteros grandes en una matriz absolutamente desproporcionada generando grandes errores en relación con la información contenida en esta al truncar el producto infinito. La información contenida en un LFT ϕ se define como $\text{info}(\phi)=\phi[0,\infty]$.

En analogía con los formatos de punto flotante, donde las representaciones de números en una base dada son generados por dos símbolos para el signo y un número finito de dígitos, restringimos las matrices signo y las matrices dígito a un conjunto finito de matrices especificas. Las matrices signo pueden ser rotaciones ϕ de S que permiten el cambio de signo de los elementos en la imagen del intervalo $[p, q] \subseteq [0, \infty]$ sin cambiar la longitud de este en S. Mientras que las matrices dígito son contracciones con respecto a la métrica ρ en $\mathbb{R}^*$ que permiten mejorar la información. Las matrices dígito son matrices de refinamiento y por tanto, aplicadas a una argumento en $[0, \infty]$ devuelve el resultado en $[0, \infty]$. Las matrices signo permiten representar los numeros en $\mathbb{R}^*$, no sólo los números positivos.

4.4.1. Matrices Signo

Empezamos con las matrices signo. La información de las matrices signadas debe traslaparse y cubrir S. Si más aún asumimos que tienen la misma longitud con respecto a ρ y son colocadas eventualmente en S, entonces ellas serán generadas por las rotaciones de S.

El LFT $\phi_{\exp i\theta} : x \mapsto \frac{x\cos\frac{\theta}{2}+\sin\frac{\theta}{2}}{-x\sin\frac{\theta}{2}+\cos\frac{\theta}{2}}$ rota S por θ. En efecto, asuma que $x \in \mathbb{R}^*$ esta representado en coordenadas homogéneas por $\begin{pmatrix}\sin\frac{\alpha}{2}\\ \cos\frac{\alpha}{2}\end{pmatrix}$ donde $s\widehat{(0)Os}(x) = \alpha$. Entonces su imagen $y = \phi_{\exp i\theta}(x)$ esta representada por el vector $\begin{pmatrix}\sin\frac{\alpha+\theta}{2}\\ \cos\frac{\alpha+\theta}{2}\end{pmatrix}$. Por tanto $s\widehat{(0)Os}(y) = \alpha + \theta$ y en consecuencia $s\widehat{(x)Os}(y) = \theta$.

Más aún $\phi_{\exp i\theta}$ genera un grupo cíclico si y sólo si θ es un múltiplo racional de 2π. Nuestra elección en el futuro esta restringida a que el LFT tenga coeficientes enteros.

Proposición 4.4.1. *Suponga que θ es un múltiplo racional no entero de 2π. Entonces el LFT $\phi_{exp\ i\theta}$ tiene coeficientes enteros si y sólo si $\theta = \frac{\pi}{2}$ ó $\theta = \pi$.*

Para $\theta = \pi$, el grupo cíclico de orden 2 consistente de $\phi_{\exp i\pi} : x \mapsto -\frac{1}{x}$ y la identidad $\mathrm{Id} : x \mapsto x$ produce los intervalos info$(\phi_{\exp i\pi}) = [\infty, 0]$ e info(Id)$= [0, \infty]$ que no se traslapan.

Para $\theta = \frac{\pi}{2}$ tenemos el grupo cíclico de orden 4 con los elementos

$$\phi_{\exp \frac{i\pi}{2}} : x \mapsto \frac{x+1}{-x+1}, \quad \phi_{\exp i\pi} : x \mapsto -\frac{1}{x}, \quad \phi_{\exp \frac{i3\pi}{2}} : x \mapsto \frac{x-1}{x+1}, \quad \mathrm{Id} : x \mapsto x$$

con información $[1, -1], [\infty, 0], [-1, 1]$ y $[0, \infty]$ respectivamente. La representación matricial de estos LFTs es respectivamente:

$$S^1 = \begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix} \quad S^2 = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} \quad S^3 = \begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} \quad S^0 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

Estas cuatro matrices forman un grupo cíclico de orden 4 (si los factores comunes no importan). Los superíndices indican que $S^i S^j = S^{(i+j) \mod 4}$. Tomamos estas como nuestras matrices signo.

4.4.2. Matrices Dígito

Fijando un entero $r \geq 0$. Todo número real en $[-1,1]$ tiene una representación $\sum_{n=1}^{\infty} k_n r^{-n}$ con dígitos enteros $|k_n| < r$ (Los dígitos pueden ser negativos) estos dígitos corresponden a las funciones afines $f_k : x \mapsto \frac{x+k}{r}$ que son LFT's o equivalentemente matrices $A_k^r = \begin{pmatrix} 1 & k \\ 0 & r \end{pmatrix}$, que envían el intervalo $[-1,1]$ en $[\frac{k-1}{r}, \frac{k+1}{r}]$. Estos intervalos imagen tienen longitud $\frac{2}{r}$ y cubren $[-1,1]$. Los intervalos imagen que vienen de sucesivos valores de k se traslapan en un intervalo común de longitud $\frac{1}{r}$. Esto provee la redundancia necesaria para aritmética real exacta.

Seleccionamos ahora un conjunto adecuado de matrices dígito desde $\mathbb{M}^+$. Dado que las matrices dígito se requieren para representar secuencias de intervalos contractantes, buscaremos matrices que contraen las distancias en $[0,\infty]$ con respecto a la métrica ρ. Las matrices dígito deben traslaparse y cubrir $[0,\infty]$.

Note que para $x, y \in [0,\infty]$, tenemos $\rho(x,y) = \left|\frac{x-1}{x+1} - \frac{y-1}{y+1}\right| = |S^3(x) - S^3(y)|$ y S^3 es un homeomorfismo desde $[0,\infty]$ a su imagen $[-1,1]$. Sea $\phi \in \mathbb{M}^+$ y considere la restricción $\phi : [0,\infty] \to [0,\infty]$. Entonces $S^3\phi(S^3)^{-1}$ es un homeomorfismo de $[-1,1]$ en sí mismo.

Proposición 4.4.2. *La función $\phi : [0,\infty] \to [0,\infty]$ es contracción con respecto a la ρ-métrica si y sólo si $S^3\phi(S^3)^{-1} : [-1,1] \to [-1,1]$ es una contracción con respecto a la métrica euclidiana.*

De esto se sigue que para cualquier base $r > 1$ la representación de dígitos signada en $[-1,1]$ en base r induce vía el homeomorfismo S^3 un conjunto adecuado de matrices dígito en $\mathbb{M}^+$.

Ahora definimos las matrices de dígitos en base r como el SFI en $[0,\infty]$ con funciones ρ-contractantes

$$D_k^r = (S^3)^{-1} A_k^r S^3 = \frac{c}{2} \begin{pmatrix} 1+r+k & -1+r+k \\ -1+r-k & 1+r-k \end{pmatrix},$$

donde c es una constante cualquiera, pues la representación es única salvo escalamiento y $k \in \mathrm{Dig}(r) = \{-r+n, r-n | n \in \mathbb{N}, 1 \leq n \leq \lfloor b \rfloor\}$. Además estas matrices dígito preservan la orientación del intervalo, pues su determinante es positivo.

Por ejemplo para la base 2 tenemos cuatro matrices signo S^0, S^1, S^2 y S^3 junto con las tres matrices dígito (para las cuales se utilizaron $c = 1, 2$ y 1 respectivamente).

$$D^2_{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \quad D^2_0 = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \quad D^2_1 = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}.$$

La aritmética exacta de punto flotante en base r esta definida como la representación de números reales por composición infinita de LFTs o equivalentemente producto infinito de matrices, tales que la primera matriz es una de las matrices signo y las matrices siguientes son matrices dígito.

Para cada composición finita $D^r_{k_1} D^r_{k_2} \cdots D^r_{k_n}$ de matrices dígito tenemos

$$S^3 D^r_{k_1} D^r_{k_2} \cdots D^r_{k_n}[0, \infty] = A^r_{k_1} A^r_{k_2} \cdots A^r_{k_n}[-1, 1]$$

Luego, para toda composición infinita de matrices dígito, tenemos

$$\bigcap_{n \geq 0} S^3 D^r_{k_1} D^r_{k_2} \cdots D^r_{k_n}[0, \infty] = \bigcap_{n \geq 0} A^r_{k_1} A^r_{k_2} \cdots A^r_{k_n}[-1, 1]$$

Esto nos da:

Proposición 4.4.3. *Un número real, con expansión digital signada $.k_1 k_2 k_3 \ldots$ (con $k_j \in Dig(r)$ para $j \geq 1$) es representado con punto flotante exacto por el producto infinito*

$$S^3 D_{k_1} D_{k_2} D_{k_3} \cdots$$

4.5. LFTs como operadores

Los LFTs no sólo pueden ser usados para representar números reales, sino también para llevar a cabo cómputos con números reales. Usando LFT's adecuados podemos expresar operaciones básicas tales como $x \mapsto x + 1$, $x \mapsto 2x$ y $x \mapsto \frac{1}{x}$.

Para ilustrar estos cálculos, recordemos como se utiliza el algoritmo euclidiano para construir una fracción continua. Tomando 1053/802, tenemos

1053		
802	1	1053/802=1+251/802=1+1/(802/251)
251	3	802/251=3+49/251 =3+1/(251/49)
49	5	251/49=5+6/49=5+1/(49/6)
6	8	49/6=8+1/6
1	6	6/1=6
0		

$$\frac{1053}{802} = 1 + \cfrac{1}{3 + \cfrac{1}{5 + \cfrac{1}{8 + \cfrac{1}{6}}}}$$

se dice que la fracción continua de 2.54 es 1 3 5 8 6.

El siguiente ejemplo y el ejemplo 4.7.1 de la Sección 4.7 son una generalización del algoritmo euclidiano para LFTs y Tensores. Para más detalles de este procedimiento ver [11].

Ejemplo 4.5.1. Ahora, suponga que se quiere aplicar el LFT $x \mapsto \frac{2}{3-x}$ a la entrada $\sqrt{2}$, es decir, se quiere construir la fracción continua de $\frac{2}{3-\sqrt{2}}$ a partir de la fracción continua de $\sqrt{2}$.

Sabiendo que la fracción continua de $\sqrt{2}$ es 1 2 2 2 ..., escribimos esta entrada de derecha a izquierda y el LFT que se quiere aplicar debajo de esta fila.

...	2	2	2	2	2	2	1	
							0	2
							-1	3

De izquierda a derecha, extendemos estas dos filas aplicando la siguiente recurrencia: multiplicar cada elemento por el término correspondiente de la primera fila, adicionarle el elemento a la derecha y escribir el resultado a la izquierda.

...	2	2	2	2	2	2	1	
					4	2	0	2
					3	2	-1	3

El resultado obtenido $\frac{4x+2}{3x+2}$ con $x \in [0, \infty]$, es el resto de la fracción continua y está entre 4/3 y 2/2, luego podemos llevar a cabo el paso euclidiano con cociente 1 como el primer termino de la respuesta:

...	2	2	2	2	2	2	1		
					4	2	0	2	
					3	2	-1	3	1
					1	0			

Aplicando la recurrencia a las últimas dos filas

...	2	2	2	2	2	2	1		
					4	2	0	2	
				8	3	2	-1	3	1
				2	1	0			

Cualquier número entre 8/2 y 3/1 tiene parte entera 3, luego 3 es el segundo término de la respuesta

...	2	2	2	2	2	2	1		
					4	2	0	2	
				8	3	2	-1	3	1
				2	1	0			3
				2	0				

Continuando:

...	2	2	2	2	2	2	1		
					4	2	0	2	
				8	3	2	-1	3	1
			5	2	1	0			3
		10	4	2	0				1
	5	2	1	0					4
10	4	2	0						1
2	1	0							4

Tenemos un ciclo desde la segunda ocurrencia del patron $\begin{pmatrix} 2 & 1 \\ 2 & 0 \end{pmatrix}$. Luego tenemos la fracción continua $\frac{2}{3-\sqrt{2}} = 1\ 3\ 1\ 4\ 1\ 4\ \ldots$

La acción de tomar un dígito del LFT del argumento se denominará absorción y la acción de devolver un dígito del LFT resultante se llamará emisión.

4.6. LFTs de refinamiento en [-1,1]

La definición de refinamiento utilizada en las secciones anteriores se hizo sobre el intervalo base $[0, \infty]$, las siguientes definiciones se hacen en el intervalo $[-1, 1]$:

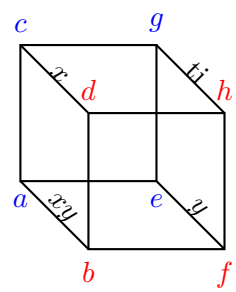
Un LFT ϕ es acotado, si $\phi([-1,1]^2) \subseteq \mathbb{R}$. Un LFT ϕ es refinado si $\phi([-1,1]) \subseteq [-1,1]$. Los LFT's acotados toman sus valores máximo y mínimo en los extremos del intervalo $[-1,1]$, luego un LFT $\phi(x) = \frac{ax+c}{bx+d}$ acotado es refinado si y sólo si $|\phi(-1)| = \left|\frac{c-a}{d-b}\right| \leq 1$ y $|\phi(1)| = \left|\frac{a+c}{b+d}\right| \leq 1$. Dado que la representación del LFT no es única, asumiendo $d \geq 0$ como una forma de normalización, tenemos $bx + d > 0,\ \forall x \in [-1,1]$ en LFT's acotados pues el denominador no puede ser cero, así que debe mantener el mismo signo en el intervalo $[-1,1]$. Por tanto si un LFT acotado es refinado $d-b > 0$ y $d+b > 0$, lo que es equivalente a

$$d - b \geq |a - c| \text{ y } d + b \geq |a + c|$$

(bajo la asunción general $d > 0$). La adición de estas dos condiciones nos lleva a $2d \geq |a-c|+|a+c| = 2\text{máx}\{|a|, |c|\}$, que implica $d \geq |a|$ y $d \geq |c|$ para matrices refinadas.

4.7. Tensores

Un tensor es una matriz de tres dimensiones $2 \times 2 \times 2$ compuesta por 8 enteros. La posición de cada entero en la matriz esta determinada por, si es o no coeficiente de x, si es o no coeficiente de y, y si está o no en el numerador (los coeficientes de xy son simultáneamente coeficientes de x y de y). En el siguiente cubo



los enteros están ubicados en los vértices, las aristas marcadas indican las variables por las que se multiplican los vértices en sus extremos, "ti" significa términos independientes, la cara $\begin{pmatrix} d & h \\ b & f \end{pmatrix}$ son los coeficientes del denominador y la cara $\begin{pmatrix} c & g \\ a & e \end{pmatrix}$ los coeficientes del numerador.

Luego la matriz

$$\begin{matrix} c & & g & \\ & d & & h \\ a & & e & \\ & b & & f \end{matrix}$$

representa la operación

$$\frac{axy + cx + ey + g}{bxy + dx + fy + h}$$

sobre las entradas x, y.

La información contenida en un tensor T se define como $\text{info}(T) = T[(0,\infty),(0,\infty)]$. Es fácil ver que $\text{info}(T) = \left[\text{mín}\{\frac{a}{b}, \frac{c}{d}, \frac{e}{f}, \frac{g}{h}\}, \text{máx}\{\frac{a}{b}, \frac{c}{d}, \frac{e}{f}, \frac{g}{h}\}\right]$. El conjunto de todos los tensores es denotado por $\mathbb{T}$. Sea $\mathbb{T}^+$ el conjunto de todos los tensores T tal que $\text{info}(T) \subseteq [0,\infty]$. Note que $\mathbb{T}^+$ es el conjunto de tensores cuyas entradas son todas no positivas o todas no negativas.

Es inmediato que las operaciones aritméticas básicas se pueden expresar con tensores

$$x + y = \begin{matrix} 1 & & 0 & \\ & 0 & & 1 \\ 0 & & 1 & \\ & 0 & & 0 \end{matrix} \qquad xy = \begin{matrix} 0 & & 0 & \\ & 0 & & 1 \\ 1 & & 0 & \\ & 0 & & 0 \end{matrix} \qquad x/y = \begin{matrix} 1 & & 0 & \\ & 0 & & 0 \\ 0 & & 0 & \\ & 0 & & 1 \end{matrix}$$

Ahora para mostrar como operan los tensores, se muestra el siguiente ejemplo un poco más elaborado que el correspondiente a las operaciones básicas.

Ejemplo 4.7.1. Si se quiere computar $\frac{2xy+x}{xy+y}$ con las entradas $x = \coth 1 = 1\ 3\ 5\ 7\ 9\ \ldots$, e $y = \sqrt{6} = 2\ 2\ 4\ 2\ 4\ 2\ 4\ \ldots$, escribimos el tensor, la entrada x en la parte superior de derecha a izquierda y la entrada y al lado derecho del tensor de arriba hacia abajo, como se muestra a continuación

$$
\begin{array}{ccccc}
\ldots 5 & 3 & 1 & & \\
 & & 1 & 0 & \\
 & & 0 & 0 & \\
 & & 2 & 0 & 2 \\
 & & 1 & 1 & \\
 & & & & 2
\end{array}
$$

Ahora como x, y varían entre $[0, \infty]$ los límites de la cantidad

$$\frac{axy + cx + ey + g}{bxy + dx + fy + h}$$

varían entre $\min\{\frac{a}{b}, \frac{c}{d}, \frac{e}{f}, \frac{g}{h}\}$ y $\max\{\frac{a}{b}, \frac{c}{d}, \frac{e}{f}, \frac{g}{h}\}$. Para la matriz en cuestión dos de estos denominadores son cero y ellos deben ser desplazados en la gráfica por un término de la entrada y

$$
\begin{array}{ccccc}
\ldots 5 & 3 & 1 & & \\
 & & 1 & 0 & \\
 & & 0 & 0 & \\
 & & 2 & 0 & 2 \\
 & & 1 & 1 & \\
 & & 5 & 0 & 2 \\
 & & 2 & 2 & \\
 & & & & 4
\end{array}
$$

Aquí la matriz 2×2 $\begin{array}{cc} 2 & 0 \\ 1 & 1 \end{array}$ fue multiplicada por 2 y se le adicionó $\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}$ para obtener $\begin{array}{cc} 5 & 0 \\ 2 & 2 \end{array}$. Ahora observe que en la parte izquierda 2/2 y 5/4, tienen la misma parte entera, mientras que en la parte inferior 5/2 y 0/2 no. Como nuestro objetivo es que las partes enteras sean iguales para poder aplicar el paso euclidiano, procedemos a la izquierda con una entrada de x.

$$
\begin{array}{ccccc}
\ldots 5 & 3 & 1 & & \\
 & & 1 & 0 & \\
 & & 0 & 0 & \\
 & 2 & 2 & 0 & 2 \\
 & 2 & 1 & 1 & \\
 & 5 & 5 & 0 & 2 \\
 & 4 & 2 & 2 & \\
 & & & & 4
\end{array}
$$

Infortunadamente esta entrada no da suficiente información para determinar una salida. Sin embargo las dos nuevas razones 2/2 y 5/4 tiene partes enteras iguales, así que continuamos hacia la izquierda

... 5	3	1			
		1	0		
		0	0		
8	2	2	0	2	
7	2	1	1		
20	5	5	0	2	
14	4	2	2		
				4	

Esto determina que el primer término de la respuesta es 1, y substraemos una vez el denominador $\begin{pmatrix} 7 & 2 \\ 14 & 4 \end{pmatrix}$ a la matriz $\begin{pmatrix} 8 & 2 \\ 20 & 5 \end{pmatrix}$ para obtener el nuevo denominador $\begin{pmatrix} 1 & 0 \\ 6 & 1 \end{pmatrix}$

... 5	3	1			
		1	0		
		0	0		
8	2	2	0	2	
7	2	1	1		salidas
1	0				1
20	5	5	0	2	
14	4	2	2		
6	1				
				4	

De nuevo un denominador cero frusta las esperanzas de una salida, pero este esta en la esquina que no le corresponde ni a x ni a y, así que nos podemos librar de el. En efecto, dado que las partes enteras de las otras tres razones son todas diferentes, necesitaremos al menos dos términos de entrada más para librarnos de ellas.

... 5	3	1			
		1	0		
		0	0		
8	2	2	0	2	
7	2	1	1		
1	0				salidas
20	5	5	0	2	1
14	4	2	2		
6	1				
35	10			4	
13	2				

Ahora 14/6 y 35/13 tienen partes enteras iguales, así que nos movemos en el sentido de x

...	5	3	1			
			1	0		
			0	0		
	8	2	2	0	2	
	7	2	1	1		salidas
	1	0				1
	20	5	5	0	2	
74	14	4	2	2		
31	6	1				
185	35	10			4	
67	13	2				

Lo cual nos da la salida 2.

...	5	3	1			
			1	0		
			0	0		
	8	2	2	0	2	
	7	2	1	1		
	1	0				salidas
						1
	20	5	5	0	2	2
74	14	4	2	2		
31	6	1				
12	2					
185	35	10			4	
67	13	2				
51	9					

Ahora debemos entrar el 4 de la secuencia y, con la cual obtendremos la salida 1, dejándonos con la matriz

51	9
16	4
216	38
83	20.

...	5	3	1			
			1	0		
			0	0		
	8	2	2	0	2	
	7	2	1	1		
	1	0				
	20	5	5	0	2	salidas
74	14	4	2	2		1
31	6	1				2
12	2					1
185	35	10			4	
67	13	2				
51	9					
16	4					
299	58				2	
216	38					
83	20					

Así sucesivamente continua el proceso.

En adelante utilizaremos la siguiente convención de escritura

$$\begin{pmatrix} a & c & e & g \\ b & d & f & h \end{pmatrix}$$

para denotar el tensor

c	g
d	h
a	e
b	f

Luego las operaciones aritméticas básicas se denotarán

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} (x, y) = x + y$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} (x, y) = x \cdot y$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} (x, y) = x/y.$$

4.8. Tensores de refinamiento en [-1,1]

Decimos que un tensor esta compuestos por dos LFTs $T = (T_0\ T_1)$ y que un LFT M esta compuesto por dos vectores $M = (M_0\ M_1)\cdot$ Un tensor $T = \begin{pmatrix} a & c & e & g \\ b & d & f & h \end{pmatrix}$ denota la función $T : \mathbb{R}^* \times \mathbb{R}^* \to \mathbb{R}^*$ dada por $T(x, y) = \frac{axy+cx+ey+g}{bxy+dx+fy+h}$. De nuevo los factores comunes no importan. Asumiremos $h > 0$ como una forma de normalización.

Para un x fijo, la función $T|_x : y \to T(x, y)$ es un LFT con

$$T|_x = \begin{pmatrix} ax + e & cx + g \\ bx + f & dx + h \end{pmatrix}.$$

De forma similar para un y fijado $T|^y : x \to T(x, y)$ donde

$$T|^y = \begin{pmatrix} ay + c & ey + g \\ by + d & fy + h \end{pmatrix}.$$

En la sección anterior se trabajo con el intervalo base $[0, \infty]$, por otro lado en el intervalo $[-1, 1]$ tenemos las siguientes definiciones:

Un tensor T es acotado, si este evita los valores ∞ y $\perp$ en sus argumentos de $[-1, 1]$, es decir, si $T([-1, 1]^2) \subseteq \mathbb{R}$. Estudiemos la expresión $bxy + dx + fy + h > 0$ para todo $x, y \in [-1, 1]$. Para x fijo y y variable en $[-1, 1]$, la expresión $bxy + dx + fy + h = (bx + f)y + (dx + h)$ alcanza su mínimo en $y = 1$ o $y = -1$. De manera similar las expresiones resultantes $(d+b)x+(f+h)$ y $(d-b)x+(f-h)$ alcanza su mínimo en $x = 1$ o $x = -1$. La suposición $h \geq 0$ es equivalente a $bxy + dx + fy + h > 0,\ \forall x, y \in [-1, 1]$ en tensores acotados. Luego de la expresión acotada $bxy + dx + fy + h > 0$ evaluada en los puntos $\{(-1, -1), (-1, 1), (1, -1), (1, 1)\}$ resultan cuatro condiciones que pueden resumirse en $h + b > |f + d|$ y $h - b > |f - d|$. Para $b = 0$ estas dos condiciones son equivalentes a la condición $h > |d| + |f|$.

Un tensor T es refinado si $T([-1, 1]) \subseteq [-1, 1]$. Para T acotado y x fijo en $[-1, 1]$, $T|_x : [-1, 1] \to \mathbb{R}$ es un LFT acotado y monótono, por tanto obtiene sus valores extremos en 1 y -1. Las funciones resultantes $T|_1 : [-1, 1] \to \mathbb{R}$ y $T|^1 : [-1, 1] \to \mathbb{R}$ son monótonas y obtienen sus valores extremos en 1 y -1. Esto muestra que un tensor T es refinado si y sólo si $|T(x, y)| \leq 1$, para todo $(x, y) \in \{(-1, -1), (-1, 1), (1, -1), (1, 1)\}$.

4.9. Escogiendo el intervalo base

Dado que $S^3 : [0, \infty] \to [-1, 1]$ es un homeomorfismo, todo lo que puede hacerse en $[0, \infty]$ puede hacerse en $[-1, 1]$.

Hay dos ventajas de usar $[0, \infty]$: Primero no es necesario hacer ninguna computación para encontrar un LFT para el que la imagen de $[0, \infty]$ sea un intervalo cualquiera como se concluyó después de la proposición 4.2.1. En contraste este mismo calculo requiere algunas

computaciones para el intervalo $[-1,1]$. Segundo un LFT o un tensor es refinado en $[0,\infty]$ si y sólo si todas sus entradas son no negativas o no positivas. Esta condición es mucho mas simple que la derivada para el refinamiento de $[-1,1]$ en las secciones 4.6 y 4.8.

Por otro lado el intervalo base $[-1,1]$ tiene matrices dígito $A_k^r = \begin{pmatrix} 1 & k \\ 0 & r \end{pmatrix}$, mientras que para $[0,\infty]$ las matrices dígito son un poco más complejas $D_k^r = (S^3)^{-1} A_k^r S^3 = \begin{pmatrix} 1+r+k & -1+r+k \\ -1+r-k & 1+r-k \end{pmatrix}$ (sección 4.4.2), principalmente porque el cero en la entrada de A_k lleva a cabo emisiones y absorciones de matrices dígito de manera más eficiente. El intervalo base $[-1,1]$ evita algunos problemas que tiene el valor ∞ en transformaciones algebraicas, operaciones importantes tales como sustracción, multiplicación, exponenciales son totales en $\mathbb{R}$, pero indefinidas para ciertos argumentos que involucran ∞. Aún más, se puede trabajar con la métrica estándar y las derivadas estándar en el intervalo $[-1,1]$, mientras que trabajar con $[0,\infty]$ excluye la métrica estándar.

Conclusiones y recomendaciones

Este trabajo empezó enfatizando que para hablar de teoría de la computabilidad en los números reales era necesario un modelo de aproximación de los números reales en el cual se pudiera generar los dígitos de la expansión decimal de un número real computable y en el que las operaciones básicas entre números reales computables, fuesen computables. Luego vimos que los números reales pueden ser aproximados por una sucesión contractante de intervalos encajados, señalando que esta aproximación corresponde a la teoría de dominos. Finalmente se materializó la idea de aproximación de los reales en la representación por intervalos centrados diádicos y LFT's, definiendo también las operaciones básicas en cada contexto.

Para dar un panorama más amplio de los alcances de las representaciones tratadas en este trabajo, mencionamos que otros procedimientos en los reales pueden computarse: calculo de funciones transcendentales, calculo de ceros de funciones (método de trisección), técnicas numéricas de integración (Trapecio, Simpson, formula de Gauss-Legendre), y los métodos de Euler y Runge-Kutta para resolver problemas de valor inicial de ecuaciones diferenciales ordinarias. También se han hecho análisis de la complejidad temporal y espacial de la aproximación de estos procedimientos.

Para futuros trabajos de investigación que extiendan la teoría de la computabilidad en los reales, es interesante construir procedimientos computables y eficientes de otros métodos numéricos. También se pueden plantear preguntas de computabilidad de subconjuntos compactos del espacio euclídeo. Por ejemplo se sabe que para el polinomio $f : \mathbb{C} \to \mathbb{C}$ de grado ≥ 2, el conjunto de Julia $J(f)$ es recursivo si f pertenece a una clase de funciones conocidas como *hiperbólicas*. Una pregunta abierta es si el conjunto de Mandelbrot $M := \{c \in \mathbb{C} : (\forall n)|f_c^n(0)| < 2\}$ es un subconjunto cerrado recursivo de $\mathbb{R}^2$ (aquí $f_c : \mathbb{C} \to \mathbb{C}, z \mapsto z^2 + c$).

Bibliografía

[1] S. Abramsky & A. Jung. Domain theory Handbook of Logic in Computer Science. Clarendo Press, 1994

[2] J. Blanck, Exact real arithmetic using centred intervals and bounded error terms, The Journal of Logic and Algebraic Programming, vol. 66 2006, pg. 50 - 67.

[3] J. Blanck, Eficient Exact Computation of iterated maps, The Journal of Logic and Algebraic Programming, vol. 64 2005, pg. 41 - 59.

[4] J. Blanck, Computability on topological spaces by effective domain representations, Ph.D. thesis, Uppsala University, 1997, Uppsala Dissertations in Mathematics 7.

[5] X. Caicedo Elementos de Lógica y Computabilidad, Universidad de los Andes, Segunda edición 1990.

[6] N. Cutland, Computability, Cambridge University Press, First edition 1980.

[7] B. A. Davey & H. A. Priestley, Introduction to Lattices and Order, Cambridge University Press, Second edition 2002.

[8] A. Edalat, domains for computation in mathematics, physics and exact real arithmetic, Bulletin of Symbolic Logic, Volume 3, Number 4, 1997, pg 401-452.

[9] A. Edalat & Philipp Sünderhauf, A domain-theoretic approach to computability on the real line

[10] A. Edalat and P. J. Potts, A new representation for exact real numbers, Proceedings of mathematical foundations of programming semantics 13, Electronic Notes in Theoretical Computer Science, vol. 6, Elsevier Science B. V., 1997.

[11] W. Gosper. Continued fraction arithmetic. Technical Report HACKMEM Item 101,B MIT Artifical Intelligence Memo 239, MIT 1972.

[12] R. Heckmann, How Many Arguments Digits Are Needed to produce n Result Digits?, Electronic Notes in Theoretical Computer Science, 1999.

[13] V. Stoltenberg-Hansen and J. V. Tucker, Computability on topological spaces via domain representations, New Computational Paradigms Changing Conceptions of What is Computable, Springer, 2008, pg. 153-194.

[14] V. Stoltenberg-Hansen, I. Lindstrom and E. R. Griffor, Mathematical Theory of Domains, Cambridge University Press, 1994.