

Soluciones algorítmicas de sistemas semi-algebraicos

con una aplicación al análisis de redes químicas pseudoquirales

por

Daniel Dario Fula Arguello

Dirigido por

Juan Andrés Montoya Arguello

Documento requerido para obtener el título de

Científico de la Computación

en la

UNIVERSIDAD NACIONAL DE COLOMBIA

ABSTRACT

Se pretende estudiar la solución algorítmica de sistemas de ecuaciones y desigualdades polinómicas, que en la literatura se conocen como sistemas semi-algebraicos y en concreto se quiere mostrar una aplicación de eso al análisis de redes químicas pseudoquirales en donde se buscan condiciones para el rompimiento de simetría.

AGRADECIMIENTOS

Me gustaría agradecer al profesor Juan Andrés Montoya por su ayuda incondicional con respecto a los temas académicos y a mi familia por su apoyo moral y financiero durante toda mi estancia en la Universidad Nacional de Colombia. Estoy seguro de que sin la ayuda de alguno de ellos, este trabajo no sería posible.

ÍNDICE GENERAL

| | | |
|-------|--|----|
| 1 | INTRODUCCIÓN | 1 |
| 1.1 | Organización | 2 |
| 2 | ALGORITMO DE COLLINS | 3 |
| 2.1 | Algoritmo de Collins en pocas palabras | 3 |
| 2.2 | Descomposiciones cilíndricas | 4 |
| 2.3 | Geometría Algebraica Real | 6 |
| 2.3.1 | Proyecciones | 6 |
| 2.3.2 | Raíces reales de polinomios y álgebra computacional | 10 |
| 2.4 | Algoritmo de Collins | 11 |
| 2.4.1 | Fase de proyección | 11 |
| 2.4.2 | Fase Base | 11 |
| 2.4.3 | Fase de levantamiento | 12 |
| 2.4.4 | Complejidad | 12 |
| 2.4.5 | Implementaciones | 13 |
| 3 | APLICACIÓN: ANÁLISIS DE REDES QUÍMICAS PSEUDOQUIRALES | 15 |
| 3.1 | Modelos de mecanismos químicos | 16 |
| 3.1.1 | Redes de reacciones químicas | 16 |
| 3.2 | Modelos de síntesis enantioselectiva: Redes quirales | 19 |
| 3.2.1 | Redes quirales y pseudoquirales | 19 |
| 3.3 | Conexión entre la geometría algebraica real y el análisis de redes pseudoquirales | 23 |
| 4 | USO DE SACNA Y ALGUNOS EXPERIMENTOS COMPUTACIONALES | 27 |
| 4.1 | Un primer ejemplo - El modelo de Calvin | 27 |
| 4.2 | Una Versión modificada del Modelo de Calvin: <i>Enantioselectividad Limitada (LES)</i> | 28 |
| 4.3 | Modelo replicador de Hochberg y Ribó | 31 |
| 4.4 | El modelo APED | 33 |
| 4.5 | El modelo APED reversible | 35 |
| 5 | CONCLUSIONES | 37 |
| | SIGLAS | 39 |
| | BIBLIOGRAFÍA | 41 |

1 INTRODUCCIÓN

El único modo de hacer un gran trabajo es amar lo que haces. Steve Jobs

PLANTEAMIENTO DEL PROBLEMA

PROBLEMA GENERAL

Estudio de las soluciones algorítmicas de sistemas de ecuaciones y desigualdades polinomiales

PROBLEMA ESPECÍFICO

Desarrollo de algoritmos para calcular los valores de los parámetros de redes químicas que pueden dar lugar a rompimiento de simetría

ESTADO DEL ARTE

No es necesario ahondar en la importancia del estudio de sistemas de ecuaciones y desigualdades polinomiales dado que este es uno de los problemas centrales de las matemáticas desde sus orígenes y es el problema fundacional de la geometría algebraica real. Por otra parte, en este trabajo se pretende profundizar un poco más en la importancia de las aplicaciones en química. Específicamente, el tema de interés son los amplificadores quirales, que son unos estados que en principio corresponden a inestabilidades que pudieron haber estado presentes en la tierra prebiótica y que pudieron haber dado lugar al rompimiento de simetría quiral que precede al origen de la vida. Dichos procesos químicos pueden ser modelados parcialmente por estructuras combinatorias que dan lugar a sistemas de ecuaciones diferenciales no lineales que normalmente no se pueden resolver en forma cerrada, por lo cual es necesario realizar un análisis cualitativo o análisis de estabilidad para intentar obtener una solución. Tal análisis cualitativo consiste en determinar, de la mejor manera posible, cuáles son los equilibrios químicos del sistema y clasificarlos de acuerdo con diferentes criterios de estabilidad que pueden ser importantes en el análisis. Considerando que los equilibrios químicos están definidos por ecuaciones polinomiales que dependen de la estructura combinatoria de la red y que las condiciones de estabilidad que se estudian típicamente están definidas por desigualdades polinomiales, el problema se reduce a solucionar un sistema semi-algebraico. No obstante, dicho sistema puede ser muy complejo y es necesario tener algoritmos para hallar su solución computacionalmente. [9].

OBJETIVOS

GENERALES

- Estudiar los principios de la teoría de la geometría algebraica real
- Conocer algunos algoritmos para resolver sistemas semi-algebraicos

ESPECÍFICOS

- Mostrar una breve aplicación del algoritmo de Collins al análisis de modelos de redes químicas en las que se buscan condiciones de rompimiento de simetría
- Implementar una herramienta computacional que permita hacer el análisis de modelos de redes químicas

1.1. ORGANIZACIÓN

El documento está organizado de la siguiente forma. En la primera sección se dan a conocer los objetivos y el alcance de esta tesis. Para la sección 2, se introducen los conceptos de la geometría algebraica real necesarios para entender el algoritmo de Collins. Luego, en la sección 3 se explican algunos conceptos químicos, con el propósito de entender el problema de analizar cierto tipo de redes químicas que den lugar al rompimiento espontáneo de la simetría espejo. La sección 3 termina mostrando una conexión entre este problema y los conjuntos semialgebraicos, lo lleva a la formulación de un algoritmo para resolver dicho problema. Para terminar, en la sección 4 se presentan los resultados de algunos experimentos computacionales llevados a cabo con la herramienta computacional SACNA, desarrollada en Mathematica[®] usando como base lo expuesto en esta tesis.

2 ALGORITMO DE COLLINS

*Las matemáticas son el lugar donde puedes hacer cosas
que no puedes en el mundo real. Marcus du Sautoy*

El propósito de esta sección es presentar al lector el algoritmo de Collins y la teoría básica de la geometría algebraica real. No se pretende hacer un tratado extenso del tema, simplemente será lo suficientemente para entender el funcionamiento del algoritmo. La mayoría de las definiciones e imágenes de esta sección se tomaron de la referencia [19], la cual contiene ejemplos ilustrativos más detallados y las demostraciones de los teoremas, que por cuestiones de espacio no se presentarán aquí. La referencia [3] trata más profundamente estos temas, tanto de manera teórica como de manera algorítmica y es una referencia fundamental para entender el algoritmo de Collins. Otra referencia básica, aunque un poco más antigua, es el libro de Mishra [26].

2.1. ALGORITMO DE COLLINS EN POCAS PALABRAS

Antes de describir brevemente el algoritmo, es necesario saber qué es un conjunto semialgebraico.

Definición 2.1. *Un conjunto se dice semialgebraico si puede construirse por un número finito de aplicaciones de las operaciones de unión, intersección y complemento de conjuntos de la forma $\{x \in \mathbb{R}^n \mid f(x) \geq 0\}$, donde $f(x)$ es un polinomio de coeficientes reales.*

Observación 2.2. *Los conjuntos semialgebraicos son interesantes porque son cerrados bajo la proyección, es decir, la proyección de un conjunto semialgebraico a una dimensión menor es nuevamente un conjunto semialgebraico. Este resultado no es para nada trivial, pero su estudio está fuera del alcance de este documento. Su prueba puede encontrarse en [3].*

Conceptualmente, el algoritmo de Collins consiste de 3 fases: proyección, base y levantamiento. En la fase de proyección la idea es convertir de manera algorítmica el conjunto semialgebraico de entrada en un nuevo conjunto semialgebraico con propiedades especiales, de manera que si se comienza con un subconjunto de \mathbb{R}^n , se pueda llegar a uno en \mathbb{R}^{n-1} , luego a uno en \mathbb{R}^{n-2} y así hasta obtener uno en \mathbb{R} . El conjunto de ceros de los polinomios que conforman el sistema semialgebraico obtenido en \mathbb{R} definen una descomposición con signo invariante en \mathbb{R} . La fase de base tiene como propósito contar y aislar las raíces reales de dichos polinomios, así como muestrear un punto en cada partición de la descomposición y encontrar los signos correspondientes a cada una. Podría decirse que se realiza un procedimiento de cierta manera similar al de hacer el denominado *diagrama de cruces* o *diagrama del cementerio* para resolver sistemas de desigualdades en \mathbb{R} . En la fase de levantamiento o extensión se toman cada uno de los puntos muestreados y se reemplazan en el conjunto semialgebraico en \mathbb{R}^2 obtenido en la primera fase. Para cada sustitución, se obtiene

nuevamente un conjunto semialgebraico con una sola variable. La idea en este punto es ir aplicando la fase 2 y la fase 3 iterativamente hasta regresar al sistema semialgebraico original, el cual estará descrito en componentes de signo invariable en \mathbb{R}^n completamente identificadas y junto con su signo correspondiente (la llamada *descomposición cilíndrica algebraica*). La solución del sistema dependerá entonces de las componentes y sus signos correspondientes.

2.2. DESCOMPOSICIONES CILÍNDRICAS

El algoritmo de Collins es conocido de manera general como el algoritmo para calcular las descomposiciones algebraicas cilíndricas. De hecho, en Mathematica[®] se puede usar con el comando `CylindricalDecomposition`. Así, es de particular importancia entender lo que son las descomposiciones cilíndricas algebraicas.

Definición 2.3. Una región \mathcal{R} es un subconjunto conexo de \mathbb{R}^n .

Definición 2.4. Un cilindro sobre una región \mathcal{R} es un conjunto de la forma $\mathcal{Z}(\mathcal{R}) = \mathcal{R} \times \mathbb{R}$.

Definición 2.5. Sean f, f_1, f_2 funciones continuas de valor real en \mathcal{R} . Una f -sección de $\mathcal{Z}(\mathcal{R})$ es un conjunto de la forma $\{(\alpha, f(\alpha)) | \alpha \in \mathcal{R}\}$ y un (f_1, f_2) -sector de $\mathcal{Z}(\mathcal{R})$ es un conjunto de la forma $\{(\alpha, \beta) | \alpha \in \mathcal{R}, f_1(\alpha) < \beta < f_2(\alpha)\}$

Definición 2.6. Sea $\mathcal{X} \subseteq \mathbb{R}^n$. Una descomposición de \mathcal{X} es una colección finita de regiones disyuntas, también llamadas componentes, cuya unión es \mathcal{X} .

Definición 2.7. Una pila sobre \mathcal{R} es una descomposición que consiste de f_i -secciones y (f_i, f_{i+1}) -sectores, donde $f_0 < \dots < f_{k+1}$ para todo $x \in \mathcal{R}$, con $f_0 = -\infty$ y $f_{k+1} = +\infty$.

Los conceptos anteriores pueden visualizarse en la figura 2.1.

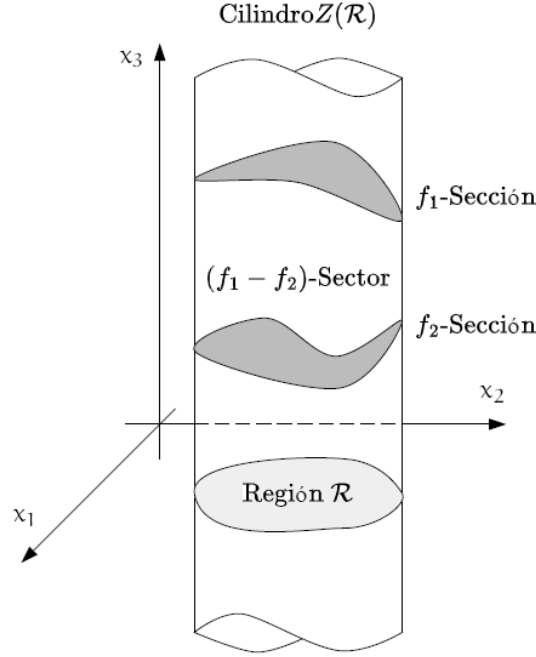


Figura 2.1: Visualización de los conceptos básicos de descomposición cilíndrica

Definición 2.8. Una descomposición \mathcal{D} de \mathbb{R}^n es cilíndrica si satisface la siguiente definición inductiva:

- $n = 1$) \mathcal{D} es una partición de \mathbb{R} en un conjunto finito de números y de los intervalos abiertos que definen dichos puntos.
- $n > 1$) $\mathcal{D}' = F_1 \cup \dots \cup F_m$ es una descomposición cilíndrica de \mathbb{R}^{n-1} y sobre cada F_i hay una pila que es subconjunto de \mathcal{D} .

Definición 2.9. Sean $\mathcal{X} \subseteq \mathbb{R}^n$ y $f \in \mathbb{R}[x_1, \dots, x_n]$. Entonces f es invariante en \mathcal{X} si se satisface una de las siguientes condiciones:

- (i) Para todo $x \in \mathcal{X}$, $f(x) > 0$.
- (ii) Para todo $x \in \mathcal{X}$, $f(x) = 0$.
- (iii) Para todo $x \in \mathcal{X}$, $f(x) < 0$.

El conjunto de polinomios $\mathcal{F} = \{f_1, \dots, f_r\} \in \mathbb{R}[x_1, \dots, x_n]$ es invariante en \mathcal{X} si cada f_i es invariante en \mathcal{X} . También se dice que \mathcal{X} es \mathcal{F} -invariante si \mathcal{F} es invariante en \mathcal{X} .

Definición 2.10. Una descomposición es algebraica si cada una de sus componentes es un conjunto semialgebraico.

Definición 2.11. Una descomposición algebraica cilíndrica (CAD, por sus siglas en inglés) de \mathbb{R}^n es una descomposición que es tanto algebraica como cilíndrica. Las componentes de una CAD son llamadas celdas.

Sea \mathcal{F} el conjunto semialgebraico definido por los polinomios $\{(x_1 - 2)^2 + (x_2 - 2)^2 - 1, (x_1 - 3)^2 + (x_2 - 2)^2 - 1\}$. La descomposición cilíndrica algebraica de \mathbb{R}^2 definida por el conjunto \mathcal{F} se muestra en la figura 2.2. Como se puede observar esta consiste gráficamente de puntos, arcos (f_i – secciones) y los espacios entre ellos ($(f_i - f_j)$ – Sectores). Además también se muestra la CAD de \mathbb{R} , que consiste de los puntos grises en el eje x_1 , así como de los intervalos abiertos que estos definen.

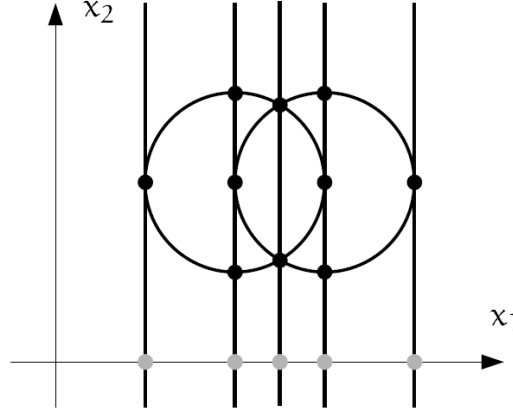


Figura 2.2: Descomposiciones cilíndricas del conjunto $\mathcal{F} = \{(x_1 - 2)^2 + (x_2 - 2)^2 - 1, (x_1 - 3)^2 + (x_2 - 2)^2 - 1\}$

2.3. GEOMETRÍA ALGEBRAICA REAL

Como se mencionó en la subsección 2.1, en la segunda fase del algoritmo es necesario contar y aislar las raíces de un polinomio. Para ello, es necesario conocer algunas herramientas que pertenecen al área de la geometría algebraica real. En particular estas herramientas son: las sucesiones de residuos polinomiales, las sub-resultantes, los coeficientes sub-resultantes principales, las cadenas de Sturm y la representación de números algebraicos.

2.3.1. PROYECCIONES

Dados dos polinomios con coeficientes reales en una variable, es posible formularse varias preguntas acerca de ellos. En particular uno podría preguntarse cuántos ceros comunes poseen dichos polinomios. Al igual que con los números enteros, la respuesta se encuentra en el máximo común divisor.

Lema 2.12. *El número de ceros comunes de $f, g \in \mathbb{R}[x]$ es $\deg(\gcd(f, g))$.*

Otra pregunta que se puede responder con el máximo común divisor, es el número de ceros distintos de un polinomio.

Lema 2.13. *El número de ceros distintos de $f \in \mathbb{R}[x]$ es $\deg(f) - \deg(\gcd(f, f'))$, donde f' denota la derivada de f .*

Así que es natural preguntarse acerca de la manera de calcular el máximo común divisor de dos polinomios de coeficientes reales. La respuesta a esto está dada por el algoritmo de Euclides para polinomios, el cual es una generalización del algoritmo de Euclides para enteros. No obstante, no siempre se tendrán polinomios en una variable, por lo que vale considerar la misma pregunta, pero esta vez para polinomios más complejos. Se introduce así, el concepto de pseudodivisión:

Lema 2.14. Sean f y g polinomios en $\mathbb{R}[x_1, \dots, x_{n-1}][x_n]$, de manera que

$$f = f_p x_n^p + \dots + f_0 y$$

$$g = g_m x_n^m + \dots + g_0,$$

es decir que f_i y g_i son polinomios en $\mathbb{R}[x_1, \dots, x_{n-1}]$ y $m \leq p$. Encones

$$g_m^s f = qg + r,$$

donde $q, r \in \mathbb{R}[x_1, \dots, x_{n-1}][x_n]$ son únicos, $s \geq 0$ y $\deg_{x_n}(r) < m$.

El residuo, r , de la pseudo-división se llama pseudo-residuo y se suele denotar como prem .

Definición 2.15. Sea \mathcal{R} un dominio de factorización única. Dos polinomios en $\mathcal{R}[x]$ se dicen similares si existen $a, b \in \mathcal{R}$ tales que $a f(x) = b g(x)$. Se denota por $f(x) \sim g(x)$.

Teniendo en cuenta la manera de calcular el máximo común divisor, surge el concepto de sucesión de residuos polinomiales.

Definición 2.16. Sean $\mathcal{R} = \mathbb{R}[x_1, \dots, x_{n-1}]$ y $f_1, f_2 \in \mathcal{R}[x_n]$ con $\deg_{x_n}(f_1) \geq \deg_{x_n}(f_2)$. La sucesión f_1, f_2, \dots, f_k es una sucesión de residuos polinomiales (PRS por sus siglas en inglés) para f_1 y f_2 si:

i. Para todo $i = 3, \dots, k$, $f_i \sim \text{prem}(f_{i-2}, f_{i-1})$.

ii. La sucesión termina con $\text{prem}(f_{k-1}, f_k) = 0$.

Existen varios tipos de sucesiones de residuos polinomiales, entre ellas EPRS (Euclidean Polynomial Remainder Sequences) y PPRS (Primitive Polynomial Remainder Sequences). Los dos tipos anteriores sufren de problemas de complejidad, por lo cual no se tratarán en esta tesis. Existe otro tipo de residuos polinomiales que ofrecen un tradeoff entre los problemas de los dos tipos de sucesiones mencionados anteriormente que se obtiene a partir de lo que se conoce como la subresultante de un polinomio. A partir de ahora y hasta que se diga lo contrario, suponga que \mathcal{R} es un dominio de factorización única.

Definición 2.17. Sea $f_i = \sum_{j=0}^{n_i} f_{ij} x^j \in \mathcal{R}[x]$, para $i = 1, \dots, k$. La matriz asociada con f_1, \dots, f_k es

$$\text{mat}(f_1, \dots, f_k) = [f_{i,l-j}],$$

donde $l = 1 + \max_{1 \leq i \leq k} (n_i)$.

Definición 2.18. Sea $\mathcal{M} \in \mathcal{R}^{k \times l}$. El polinomio determinante de \mathcal{M} se define como

$$\text{detpol}(\mathcal{M}) = \det(\mathcal{M}^{(k)})x^{l-k} + \cdots + \det(\mathcal{M}^{(l)}),$$

donde $\mathcal{M}^{(j)} = [\mathcal{M}_{\cdot,1} \cdots \mathcal{M}_{\cdot,k-1} \mathcal{M}_{\cdot,j}]$.

Definición 2.19. Sean $f, g \in \mathcal{R}[x]$ y $\deg(f) = m, \deg(g) = n, m \geq n$. El k -ésimo subresultante de f y g es

$$\text{subres}_k(f, g) = \text{detpol}(\mathcal{M}_k),$$

donde $\mathcal{M}_k = \text{mat}(x^{n-k-1}f, \dots, f, x^{m-k-1}g, \dots, g)$. Además la cadena subresultante de f y g es $(S_j)_{j=0}^{n+1}$, donde

$$\begin{aligned} S_{n+1} &= f, \\ S_n &= g, \\ S_{n-1} &= \text{subres}_{n-1}(f, g), \\ &\vdots \\ S_0 &= \text{subres}_0(f, g). \end{aligned}$$

Se debe comentar que \mathcal{M}_0 es la matriz de Sylvester de f y g y que $\text{detpol}(\mathcal{M}_0) = \det(\mathcal{M}_0)$ es el resultante de f y g . Además, las matrices \mathcal{M}_k se pueden obtener al eliminar filas y columnas de \mathcal{M}_0 .

Definición 2.20. Sean $f, g \in \mathcal{R}[x]$ y $\deg(f) = m, \deg(g) = n, m \geq n$. El k -ésimo coeficiente subresultante principal (psc por sus siglas en inglés) de f y g se define como

$$\text{psc}_k(f, g) = \det(\mathcal{M}_k^{(k)}), 0 \leq k \leq n.$$

En otras palabras, $\text{psc}_k(f, g)$ es el coeficiente de x^k en $\text{subres}_k(f, g)$.

Lema 2.21. Sean $f, g \in \mathcal{R}[x]$, donde $\deg(f) = m$ y $\deg(g) = n$. Entonces para todo $0 < i \leq \min(m, n)$ f y g tienen un factor común de grado igual a i , si y sólo si, $\text{psc}_j(f, g) = 0$, $j = 0, \dots, i-1$ y $\text{psc}_i(f, g) \neq 0$.

El concepto de delineabilidad permite formalizar la manera de encontrar regiones sobre las cuales un conjunto de polinomios tiene un número constante de raíces reales, lo cual es la idea principal de la fase de proyección. Antes de definir dicho concepto, es necesario conocer la siguiente terminología básica.

Observación 2.22. Sea $f_i \in \mathbb{R}[x_1, \dots, x_{n-1}][x_n]$ un polinomio real en n variables, de manera que

$$f_i(x_1, \dots, x_{n-1}, x_n) = f_i^{d_i}(x_1, \dots, x_{n-1})x_n^{d_i} + \cdots + f_i^0(x_1, \dots, x_{n-1})$$

y $\xi = (\xi_1, \dots, \xi_{n-1}) \in \mathbb{R}^{n-1}$. Entonces se denotará al polinomio univariado f evaluado en ξ como $f_{i,\xi}(x_n) = f(\xi_1, \dots, \xi_{n-1}, x_n)$.

Definición 2.23. Se define la reducción de un polinomio f_i de grado d_i , para $0 \leq k_i \leq d_i$ como $\hat{f}_i^{k_i}(x_1, \dots, x_{n-1}, x_n) = f_i^{k_i}(x_1, \dots, x_{n-1})x_n^{k_i} + \cdots + f_i^0(x_1, \dots, x_{n-1})$.

Definición 2.24. Sea $\mathcal{F} = \{f_1, f_2, \dots, f_r\} \subset \mathbb{R}[x_1, \dots, x_{n-1}][x_n]$ un conjunto de polinomios reales multivariados y $\mathcal{C} \subset \mathbb{R}^{n-1}$. Se dice que \mathcal{F} es *delineable* en \mathcal{C} si satisface las siguientes propiedades invariantes:

- Para todo $1 \leq i \leq r$, el número total de raíces complejas de $f_{i,\xi}$ (contando la multiplicidad) permanece invariante cuando ξ varía sobre \mathcal{C} .
- Para todo $1 \leq i \leq r$, el número de raíces complejas distintas de $f_{i,\xi}$ permanece invariante cuando ξ varía sobre \mathcal{C} .
- Para todo $1 \leq i \leq r$, el número total de raíces complejas comunes de $f_{i,\xi}$ y $f_{j,\xi}$ (contando multiplicidades) permanece invariante cuando ξ varía sobre \mathcal{C} .

Lema 2.25. Sea $\mathcal{F} \subset \mathbb{R}[x_1, \dots, x_{n-1}][x_n]$ un conjunto de polinomios y sea \mathcal{F} un conjunto *delineable* sobre $\mathcal{C} \subset \mathbb{R}^{n-1}$. Entonces el número de raíces reales distintas de \mathcal{F} es invariante sobre \mathcal{C} .

Teorema 2.26. Sea $\mathcal{F} \subset \mathbb{R}[x_1, \dots, x_{n-1}][x_n]$ un conjunto de polinomios y sea $\mathcal{C} \subset \mathbb{R}^{n-1}$. Entonces \mathcal{C} es *semialgebraico*.

Observación 2.27. Se denota por $\text{psc}_i^{x_n}$ al *iésimo* coeficiente subresultante principal con respecto a x_n y por D_{x_n} al operador derivada con respecto a x_n .

Definición 2.28. Sea $\mathcal{F} = \{f_1, f_2, \dots, f_r\}$, entonces $\text{proj}(\mathcal{F}) = \text{proj}_1(\mathcal{F}) \cup \text{proj}_2(\mathcal{F}) \cup \text{proj}_3(\mathcal{F})$, donde

$$\begin{aligned} \text{proj}_1(\mathcal{F}) &= \{f_i^k(x_1, \dots, x_{n-1}) \mid 1 \leq i \leq r, 0 \leq k \leq d_i\} \\ \text{proj}_2(\mathcal{F}) &= \{\text{psc}_l^{x_n}(\hat{f}_i^k(x_1, \dots, x_n), D_{x_n}(\hat{f}_i^k(x_1, \dots, x_n))) \mid \\ &\quad 1 \leq i \leq r, 0 \leq l < k \leq d_i - 1\} \\ \text{proj}_3(\mathcal{F}) &= \{\text{psc}_m^{x_n}(\hat{f}_i^{k_i}(x_1, \dots, x_n), \hat{f}_j^{k_j}(x_1, \dots, x_n)) \mid \\ &\quad 1 \leq i < j \leq r, 0 \leq m \leq k_i \leq d_i, 0 \leq m \leq k_j \leq d_j\} \end{aligned}$$

En resumen, dado un conjunto de polinomios $\mathcal{F} \subset \mathbb{R}[x_1, \dots, x_{n-1}][x_n]$ se puede calcular otro conjunto de polinomios de $(n-1)$ variables $\text{proj}(\mathcal{F}) \subset \mathbb{R}[x_1, \dots, x_{n-1}]$, el cual caracteriza los subconjuntos \mathcal{F} -delineables maximalmente conexos de \mathbb{R}^{n-1} . Para terminar esta subsección, se comentarán las implicaciones de que el conjunto $\text{proj}(\mathcal{F})$ sea invariante sobre $\mathcal{C} \subset \mathbb{R}^{n-1}$.

- La invarianza del conjunto $\text{proj}_1(\mathcal{F})$ implica que el grado con respecto a la variable x_n es constante sobre \mathcal{C} y por tanto el número de raíces de cada polinomio es constante sobre \mathcal{C} .
- La invarianza del conjunto $\text{proj}_2(\mathcal{F})$ implica que el máximo común divisor de cada polinomio y su derivada tiene un grado constante. Así, debido a la invarianza de $\text{proj}_1(\mathcal{F})$, se concluye que el número de ceros distintos de cada polinomio en \mathcal{F} es constante.
- La invarianza del conjunto $\text{proj}_3(\mathcal{F})$ junto con la invarianza del conjunto $\text{proj}_1(\mathcal{F})$ llevan a concluir que el número de ceros comunes de cada par de polinomios en \mathcal{F} es constante.

2.3.2. RAÍCES REALES DE POLINOMIOS Y ÁLGEBRA COMPUTACIONAL

Para la fase base, es necesario el conteo y el aislamiento de las raíces distintas de un polinomio en una sola variable. Para este fin, en esta subsección se presentan las cadenas de Sturm. Por último se discutirá acerca de la necesidad de tener implementada de antemano un sistema de álgebra computacional (CAS).

Definición 2.29. Sean $f_1, \dots, f_r \in \mathbb{R}[x]$ polinomios de variable real con las siguientes propiedades con respecto a un intervalo (a, b) :

- $f_k(x) = 0$ implica que $f_{k-1}(x)f_{k+1}(x) < 0, a < x < b$.
- $f_r(x) \neq 0, a < x < b$.

Entonces la sucesión f_1, \dots, f_r se llama una cadena de Sturm en el intervalo (a, b) . Si se multiplica cada polinomio de la cadena de Sturm por el mismo polinomio arbitrario $d(x)$, la cadena obtenida se llama cadena generalizada de Sturm.

Tómese un polinomio $f_1 \in \mathbb{R}[x]$ y $f_2 = f'_1$. Considérese la siguiente sucesión de residuos polinomiales:

$$\begin{aligned} f_1 &= q_1 f_2 - f_3 \\ f_2 &= q_2 f_3 - f_4 \\ &\vdots \\ f_{r-2} &= q_{r-2} f_{r-1} - f_r \\ f_{r-1} &= q_{r-1} f_r + 0, \end{aligned}$$

donde $f_r = \gcd(f_1, f_2)$. Esta es una sucesión de Sturm en (a, b) por construcción si $f_r \neq 0$ en (a, b) .

Definición 2.30. Sea $(a_i)_{i=1}^r$ una sucesión finita de números reales. Entonces se denota al número de variaciones de signo de la sucesión se denota por $\text{var}(a_1, \dots, a_r)$.

Teorema 2.31. Sea $f_1(x), \dots, f_r(x)$ la cadena obtenida anteriormente y $\mathcal{V}(x) = \text{var}(f_1(x), \dots, f_r(x))$. Entonces el número de ceros reales distintos de f_1 en (a, b) está dado por $\mathcal{V}(a) - \mathcal{V}(b)$.

El teorema anterior indica una manera para contar las raíces reales distintas de un polinomio en una variable. No obstante es necesario conocer un intervalo que contenga todas las raíces de dicho polinomio. En 1829, Cauchy dio una cota para todas las raíces de un polinomio. Este resultado dice que para un polinomio $p(x) = \sum_{i=0}^n a_i x^i$, si ρ es una raíz de $p(x)$, entonces

$$|\rho| < 1 + \max \left\{ \left| \frac{x_{n-1}}{x_n} \right|, \left| \frac{x_{n-2}}{x_n} \right|, \dots, \left| \frac{x_0}{x_n} \right| \right\}.$$

Resta entonces indicar un método para aislar las raíces. El tema en sí es bastante amplio y es fuente de investigación. Un método muy usado es el método de la bisección. Este consiste, de manera resumida, en la división recursiva del intervalo que contiene todas las raíces de un polinomio, de

manera tal que se obtengan subintervalos que contienen cero o una raíz del polinomio. De esta forma se consigue aislar cada raíz del polinomio en un intervalo.

Cabe resaltar la importancia de tener un sistema algebraico computacional para llevar a cabo todas las operaciones descritas anteriormente. Aunque se puedan aislar las raíces, estas deben evaluarse en un polinomio de varias variables al entrar en la fase de extensión o levantamiento. Pero no se pueden evaluar aproximaciones de la raíz, por más cercanas que puedan ser. Así, lo anterior lleva a la pregunta ¿cómo representar de manera exacta la raíz de un polinomio en un computador, de manera tal que se puedan realizar las operaciones aritméticas básicas para poder hacer la evaluación? Un número algebraico real es un número que es cero de un polinomio univariado con coeficientes reales. Una manera, quizá la más intuitiva (después de todo lo mencionado anteriormente), de representar un número algebraico en el computador sea por medio del intervalo en el que se encuentra, así como con un polinomio para el cual sea raíz. Dicho polinomio suele ser el llamado polinomio minimal. Y la herramienta principal para realizar las operaciones aritméticas con los números algebraicos son los resultantes. En cualquier caso, este tema está fuera del alcance de esta tesis, pero una respuesta más detallada a tal pregunta se puede encontrar en cualquier libro serio de álgebra computacional. Para el muestreo de los intervalos, se suele tomar un racional en cada uno de ellos y es de resaltar que la manera en la que se deberían escoger también es un tema de investigación.

2.4. ALGORITMO DE COLLINS

2.4.1. FASE DE PROYECCIÓN

En la literatura se pueden encontrar dos modelos para realizar esta fase. El primero de ellos es el que se discutirá aquí y fue dado por Collins y Hong. Su idea es tomar un conjunto \mathcal{F} de polinomios y convertirlo en otro conjunto \mathcal{P} de polinomios, tal que la descomposición algebraica natural del espacio euclidean en regiones conexas maximales para las cuales los elementos de \mathcal{P} son invariantes al signo sea una CAD (ver las referencias [9] y [17]). Este operador de proyección se aplica recursivamente $n - 1$ veces hasta obtener un polinomio en una sola variable. Sea $\mathcal{F} = \{f_1, \dots, f_r\}$, $\text{proj}^0 = \mathcal{F}$ y $\text{proj}^i(\mathcal{F}) = \text{proj}(\text{proj}^{i-1}(\mathcal{F}))$ para $1 \leq i \leq n - 1$.

El segundo modelo fue dado por Brown-McCallum y aquí la idea es producir una CAD para la cual las proyecciones sean invariantes al orden. Este modelo no se tratará en esta tesis, pero el lector interesado puede consultar información bien detallada en las referencias [23] y [24].

Se destaca que la implementación en Mathematica, por ejemplo, usa la proyección de Brown-McCallum para conjuntos bien orientados y la de Collins-Hong para cualquier otro tipo de conjunto.

2.4.2. FASE BASE

La fase base consiste básicamente de 2 pasos: El aislamiento o separación de raíces reales y la elección de los puntos de muestreo. Como entrada tiene el polinomio de una variable, $\text{proj}^{n-1}(\mathcal{F})$. Las raíces de cualquier polinomio univariado definen una descomposición invariante al signo en \mathbb{R} . Una vez se aíslan y se ordenan las raíces del polinomio, se debe obtener un punto de cada intervalo definido por dichas raíces. Para cada raíz, se debe guardar su representación exacta como

número algebraico real. Denótese por \mathcal{L} a la lista de puntos de cada intervalo y de las raíces del polinomio.

2.4.3. FASE DE LEVANTAMIENTO

En la fase de extensión o levantamiento, la idea es llevar una descomposición de signo invariante de \mathbb{R}^{i-1} , \mathcal{D}_{i-1} , a una descomposición de signo invariante de \mathbb{R}^i y usar la técnica de la fase base de manera repetitiva.

Supóngase que se está calculando la fase de levantamiento para $\text{proj}^i(\mathcal{F})$ y que su variable es x_i . La proyección inmediatamente anterior $\text{proj}^{i-1}(\mathcal{F})$, que tendrá como variables a x_i y x_{i-1} . Reemplace cada punto l en la lista, obtenida en la fase base, por la variable x_i . Se obtienen entonces polinomios en la variable x_{i-1} , que corresponden a los valores de $\text{proj}^{i-1}(\mathcal{F})$ sobre la línea "vertical" $x_i = l$. En este punto se aplica nuevamente la fase base a cada polinomio obtenido. Una vez se ha hecho lo anterior para todos los conjuntos proyección, se obtiene una lista de celdas de \mathbb{R}^n , junto con un punto que representa a cada celda (es decir, la descomposición cilíndrica). Estos puntos se deben evaluar en el sistema original para obtener los signos correspondientes a cada celda y se seleccionan únicamente las celdas que satisfagan las condiciones dadas para el sistema semialgebraico original. Probablemente la mejor manera de almacenar las celdas y los puntos sea por medio de una estructura de datos en forma de árbol, debido a que por construcción, una vez que se ha dado como entrada un polinomio en la fase base, de este salen, por decir, r puntos, que a su vez definirán r polinomios univariados diferentes usando la proyección inmediatamente anterior.

2.4.4. COMPLEJIDAD

La complejidad del algoritmo de Collins es doblemente exponencial en el número de variables del sistema semialgebraico. Este comportamiento puede explicarse teniendo en cuenta la siguiente observación: Para calcular el operador proyección (en particular para el segundo conjunto $\text{proj}_3(\mathcal{F})$ y el tercer conjunto $\text{proj}_3(\mathcal{F})$) de una familia de polinomios \mathcal{F} cuyo grado máximo es, digamos, d , se requiere el cálculo del psc de dos polinomios. Para el caso de $\text{proj}_3(\mathcal{F})$, según la definición, se pueden encontrar como un orden de $O(d^2)$ nuevos polinomios. Esto simplemente si se fijan 2 polinomios y se le calcula el psc de cada una de las posibles combinaciones de sus reducciones. Además el subresultante de dos polinomios de grado d será como máximo de d . De esta manera, una familia de polinomios de n variables dará lugar a una familia de polinomios de $n - 1$ variables de orden $O(d^2)$ cuyo grado máximo será d . Debido a que el procedimiento debe hacerse nuevamente para cada uno de los polinomios de dicha familia y que este proceso debe hacerse $n - 1$ veces, al final la etapa base se deberá realizar en el peor de los casos un orden de $O(d^{2^{n-1}})$ veces. De manera más formal, Collins probó el siguiente resultado:

Teorema 2.32. *Existe un algoritmo que calcula una descomposición cilíndrica de \mathbb{R}^n invariante a los signos de una familia \mathcal{F} de polinomios con coeficientes entero. Si la familia \mathcal{F} contiene m polinomios de grado menor o igual a d y con coeficientes acotados por 2^H (longitud acotada por H), el tiempo que toma finalizar este algoritmo está acotado por $(2d)^{2^{2n+8}} m^{2^{n+6}} H^3$.*

Esto implica que el rango de aplicaciones del algoritmo de Collins es bastante reducido.

2.4.5. IMPLEMENTACIONES

Para finalizar esta sección, se describen brevemente algunos programas que tienen implementado el algoritmo de Collins. A día de hoy se conocen 4 implementaciones razonablemente completas para calcular las descomposiciones cilíndricas de sistemas semialgebraicos.

- QEPCAD B: Este sistema es una mejora de QEPCAD, implementado por Hoon Hong, con la contribución de muchas personas, incluido George Collins. Está programado en C++ y está disponible en <https://www.usna.edu/CS/qepcadweb/B/QEPCAD.html> junto con su documentación. El sistema algebraico simbólico SageMath cuenta con una interfaz para el uso de QEPCAD.
- RLCAD: Esta es una implementación de CAD en Reduce como parte del sistema Redlog por Andreas Seild y Thomas Sturm. Reduce se puede descargar de la página <https://reduce-algebra.sourceforge.io> y la documentación del algoritmo está disponible en <https://www.redlog.eu/documentation/service.php?key=rlqe>.
- Mathematica: Mathematica[®] es un programa utilizado en áreas científicas, de ingeniería, matemática y áreas computacionales. La implementación es debida a Adam Strzbonski y se puede realizar con el comando `CylindricalDecomposition`. Su funcionalidad está descrita en la documentación de Mathematica[®].
- Maple: Al igual que Mathematica, Maple es un programa orientado a la resolución de problemas matemáticos, capaz de realizar cálculos simbólicos, algebraicos y de álgebra computacional. Se puede llevar a cabo una CAD usando el comando `CylindricalAlgebraicDecompose`. La implementación usada en este sistema es un algoritmo incremental basado en descomposiciones triangulares y fue propuesta por Changbo Chen y Marc Moreno Maza en el ASCM de 2012.

3

APLICACIÓN: ANÁLISIS DE REDES QUÍMICAS PSEUDOQUIRALES

Todas las fantasías deberían tener una base sólida en la realidad. Max Beerbohm

En esta sección se pretende mostrar una aplicación de la teoría expuesta en la sección anterior al análisis de redes químicas. Tal aplicación se formulará en forma de problema, el cuál irá cambiando a medida que se muestran todos los conceptos que lo involucran, de manera tal que al final se obtiene un problema matemático que se puede resolver usando el algoritmo de Collins. Por último se muestran algunos resultados (computacionalmente) experimentales obtenida al usar una implementación de todas las ideas de esta sección.

Antes de poder describir el problema que se pretende resolver, es necesario entender la terminología y el contexto en el cual éste se encuentra. A continuación se presenta la terminología química necesaria para entender el problema.

Sustratos homoquirales Supóngase que se tiene un sustrato químico que puede contener moléculas de las especies enantioméricas L y D . Se dice que este sustrato es homoquiral si la concentración de una de ellas es cero, mientras que la concentración de la otra no es cero.

Homoquiralidad biológica Se ha observado que 19 de los 20 aminoácidos naturales son homoquirales, siendo L-quirales (hacia la izquierda), mientras que los azúcares son D-quirales (hacia la derecha).

Sustratos racémicos Se dice que un sustrato es racémico, con respecto a la pareja (L, D) si las concentraciones de las dos especies mencionadas son iguales.

Diferencia enantiomérica Si un sustrato no es homoquiral, usualmente exhibe una diferencia enantiomérica, la cual se puede calcular al tomar la diferencia entre las concentraciones de las especies L y D .

Rompimiento espontáneo de la simetría espejo Este término designa la dinámica que se obtiene al perturbar un equilibrio químico racémico y la cual lleva a su respectivo sistema químico hacia estados homoquirales, o hacia estados con diferencia enantiomérica. Se usaran las siglas en inglés SMSB por simplicidad.

Síntesis enantioselectiva Este término es usado para designar el proceso químico que transforma sustratos racémicos en sustratos homoquirales, el cual funciona bajo condiciones completamente simétricas (aquirales). Observe que la síntesis enantioselectiva es una consecuencia del rompimiento espontáneo de la simetría espejo.

Se ha argumentado que uno de los pasos fundamentales de la evolución química fue el surgimiento de sustratos homoquirales en la vida prebiótica (ver la referencia [25]). También se ha debatido que el surgimiento de dichos sustratos homoquirales fue un prerequisite para el origen de la vida,

así, si se está interesado en el origen de las biomoléculas, se debe estar también interesado en explicar cómo esos sustratos homoquirales pudieron emerger a partir de los racémicos. La química de los enantiómeros nos dice que las dos posibles quiralidades son igualmente probables y esto implica que los sustratos originarios tuvieron que ser racémicos.

La ocurrencia de la síntesis enantioselectiva en la tierra prebiótica ha sido explicada como una inestabilidad química (ver la referencia [14]). Esto significa que los sustratos homoquirales que precedieron el origen de la vida fueron alcanzados a través de perturbaciones en estados racémicos inestables. En la literatura hay varios modelos de mecanismos químicos que se supone que exhiben SMSB (ver la referencia [22] y las referencias que esta contiene). Sin embargo, se ha observado que esta literatura contiene muchos modelos incorrectos, es decir, modelos que no son consistentes con los principios básicos de la termodinámica y la cinemática; mientras que otros modelos son incapaces de producir SMSB.

El problema que se pretende resolver está especialmente relacionado con el último tipo de modelos mencionados en el párrafo anterior. Específicamente, el problema es el siguiente:

Problema: *Dado un modelo \mathcal{M} que pretende ser usado para explicar el origen de la homoquiralidad biológica, identificar si \mathcal{M} es capaz de producir SMSB*

Es muy común poder enunciar un problema de manera sencilla y aún así estar lejos de poder entenderlo y es justamente lo que ocurre aquí. Para entender el problema (matemáticamente) es necesario describir de manera formal los modelos de mecanismos químicos y en particular los modelos de homoquiralidad biológica capaces de producir SMSB, además de saber cuál es la relación de estos con las matemáticas.

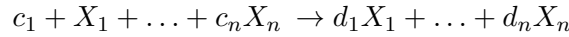
3.1. MODELOS DE MECANISMOS QUÍMICOS

Para resolver el problema, debemos entender bien qué es un modelo de mecanismos químicos. Así, el enfoque en esta sub-sección está en los mecanismos de reacción química y se afirma que cualquier mecanismo de reacción química puede describirse adecuadamente por una red de reacción química (ver [18], [13] y sus referencias). Así, el enfoque ahora es determinar los conjuntos de redes de reacciones químicas que pueden considerarse como modelos de la síntesis enantioselectiva.

3.1.1. REDES DE REACCIONES QUÍMICAS

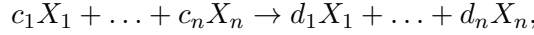
Una reacción química tiene lugar cuando se colocan algunas especies químicas en un recipiente. El recipiente tiene extensión espacial y allí ocurren tanto las reacciones químicas como los fenómenos de difusión que afectan las concentraciones de las especies químicas. Para simplificar el problema, se ignorará el fenómeno de difusión, se supondrá que los mecanismos de reacción actúan homogéneamente en el recipiente y que además el recipiente se agita continuamente haciendo que su contenido esté distribuido homogéneamente en el espacio. Se pueden representar estos mecanismos de reacción como procesos temporales que tienen lugar en un único punto en el espacio. Así se reduce esta dinámica a la evolución temporal de las concentraciones de las especies químicas.

Definición 3.1. Una reacción química sobre las especies químicas X_1, \dots, X_n se representa como una expresión de la siguiente forma



, donde c_1, \dots, c_n y d_1, \dots, d_n son enteros no negativos.

La expresión



indica que la mistura de c_1 unidades de X_1, \dots y c_n unidades de X_n da como resultado d_1 unidades de X_1, \dots y d_n unidades de X_n . Si se tienen las igualdades $c_i = 0$ para todo $i \leq n$, entonces la reacción representa una aparición mientras que si se tienen las igualdades $d_i = 0$ para todo $i \leq n$, entonces la reacción representa una fuga.

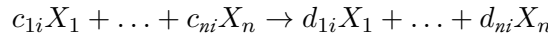
Definición 3.2. Una red de reacciones químicas es una pareja

$$\Omega = (\{X_1, \dots, X_n\}, \{R_1, \dots, R_r\}),$$

donde $\{R_1, \dots, R_r\}$ es un conjunto de reacciones químicas que actúan sobre las especies $\{X_1, \dots, X_n\}$.

Las redes de reacciones químicas son descripciones simbólicas de procesos químicos que están enteramente constituidos por mecanismos de reacciones químicas, es decir que en ellas no ocurren procesos fisicoquímicos diferentes a la reactividad química, como lo son la difusión y la advección.

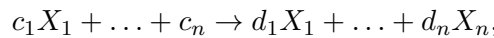
Observación 3.3. Sea $\Omega = (\{X_1, \dots, X_n\}, \{R_1, \dots, R_r\})$ una red de reacciones químicas. Se usa la expresión



para denotar la reacción R_i . Se usan las variables $[X_1], \dots, [X_n]$ para denotar las concentraciones de las n especies químicas.

DINÁMICA DE LAS REDES DE REACCIONES QUÍMICAS

Suponga que se conocen las concentraciones iniciales de las especies químicas en Ω . ¿Puede describirse la evolución temporal de tales concentraciones? Es posible, si se supone que la dinámica es gobernada por un mecanismo determinista y que conocemos dicho mecanismo. La ley de acción de masas enuncia que la velocidad de ocurrencia de una reacción química es proporcional al producto de las concentraciones de sus reactivos. Esto es, dada una reacción



hay una ley química k (una *constante de reacción*) tal que la velocidad de ocurrencia de esta reacción, en el estado representado por la tupla $([X_1], \dots, [X_n])$, es igual a $k[X_1]^{c_1} \dots [X_n]^{c_n}$. La ley de acción de masas es la expresión macroscópica de la dinámica microscópica que se supone que ocurre en el interior de los reactores.

Observación 3.4. Cada reacción química tiene asociada una constante de reacción. Por lo tanto, dada una red Ω , se asocia a la reacción R_i una constante de reacción k_i . La velocidad instantánea de R_i , en el estado $([X_1], \dots, [X_n])$ se considera igual a $k_i[X_1]^{c_{1i}} \dots [X_n]^{c_{ni}}$.

3 Aplicación: Análisis de redes químicas pseudoquirales

Sea

$$\Omega = (\{X_1, \dots, X_n\}, \{R_1, \dots, R_r\})$$

una red de reacciones químicas. Se quiere calcular las funciones

$$[X_1], \dots, [X_n] : \mathbb{R}_+ \rightarrow \mathbb{R}_+,$$

que representan las concentraciones de las especies químicas en cada instante de tiempo. Las matemáticas variacionales indican que describir la variación instantánea de funciones suaves es por lo general más fácil que describir las funciones mismas. Por lo tanto, se pondrá más enfoque en las derivadas

$$\frac{d[X_1]}{dt}, \dots, \frac{d[X_n]}{dt}.$$

Supóngase que se inicializa el sistema en el estado $[X_1]_0, \dots, [X_n]_0$. Sean $i \leq n$ y $j \leq r$. Note que la reacción R_j contribuye a la variación instantánea de $[X_i]$. Esta última contribución es proporcional al producto de su *velocidad instantánea* $k_j[X_1]^{c_{1j}} \dots [X_n]^{c_{nj}}$ y su *coeficiente estequiométrico* v_{ij} , el cual es igual a $d_{ij} - c_{ij}$. También se tiene que la variación instantánea de $[X_i]$ se reduce a la suma de las contribuciones de todas las reacciones químicas en la red. Se obtiene entonces que la igualdad

$$\frac{d[X_i]}{dt} = \sum_{j=1}^r k_j v_{ij} ([X_1]^{c_{1j}} \dots [X_n]^{c_{nj}})$$

es válida para todo $i \leq n$. Además, se obtiene que el sistema de ecuaciones diferenciales ordinarias

$$\begin{aligned} \frac{d[X_1]}{dt} &= \sum_{j=1}^r k_j v_{1j} ([X_1]^{c_{1j}} \dots [X_n]^{c_{nj}}) \\ &\vdots \\ \frac{d[X_n]}{dt} &= \sum_{j=1}^r k_j v_{nj} ([X_1]^{c_{1j}} \dots [X_n]^{c_{nj}}), \end{aligned}$$

es el *sistema autónomo de ecuaciones diferenciales ordinarias de primer orden* que, de acuerdo a la ley de acción de masas [18], conduce a la dinámica de la red Ω . Note que todas las dinámicas son determinísticas y por tanto la evolución temporal de aquellos mecanismos químicos está completamente determinada por sus estados iniciales. También debe notarse que un estado inicial de Ω se puede describir totalmente por una $(n + r)$ -tupla $([X_1], \dots, [X_n], k_1, \dots, k_r)$ de números reales no negativos.

Considérese el siguiente ejemplo de red de reacciones químicas. Se usará el símbolo Ω_F para denotar la red $(\{L, D\}, \{R_1, R_2, R_3\})$, donde:

- La reacción R_1 es igual a $L \rightarrow 2L$,
- la reacción R_2 es igual a $D \rightarrow 2D$ y

- la reacción R_3 es igual a $L + D \rightarrow \square$

donde el símbolo \square indica que esta reacción no tiene ningún producto. Entonces, se supone que Ω_F describe un proceso químico que involucra únicamente dos especies que reaccionan de tres maneras diferentes. La dinámica de Ω_F está dada por el sistema no lineal

$$\frac{d[L]}{dt} = k_1[L] - k_3[L][D]$$

$$\frac{d[D]}{dt} = k_2[D] - k_3[L][D].$$

¿Qué se puede decir acerca de la dinámica de estos mecanismos químicos abstractos? Se tiene que considerar un acercamiento cualitativo de la última pregunta, puesto que la dinámica no es lineal. De todo lo anterior, se concluye que el problema que se quiere resolver se ha vuelto el siguiente:

Problema: *Dada una red de reacciones químicas \mathcal{M} que pretende ser usado para explicar el origen de la homoquiralidad biológica, identificar si \mathcal{M} es capaz de producir SMSB*

3.2. MODELOS DE SÍNTESIS ENANTIOSELECTIVA: REDES QUIRALES

Las redes de reacciones químicas modelan una gran variedad de mecanismos químicos y claramente no todas ellas sirven para modelar la homoquiralidad biológica. Es natural preguntarse cuáles de ellas son las que se deben considerar.

3.2.1. REDES QUIRALES Y PSEUDOQUIRALES

Cualquier modelo realista de homoquiralidad biológica debe satisfacer algunas simetrías relacionadas a la *indistinguibilidad* de los enantiómeros: las dos especies que constituyen un par enantiomérico no pueden distinguirse por especies aquirales. Entonces, la primera restricción de simetría está descrita por la siguiente definición.

Definición 3.5. *Una red quiral es una red de reacciones químicas*

$$\Omega = (\{L_1, \dots, L_k, D_1, \dots, D_k, X_{2k+1}, \dots, X_n\}, \{R_1, \dots, R_r\}),$$

la cual contiene un subconjunto no vacío de tamaño k pares enantioméricos dentro de su conjunto de especies. Además dicho subconjunto es indistinguible en la red, es decir que si R_i es igual a

$$\begin{aligned} & c_1 L_1 + \dots + c_k L_k + e_1 D_1 + \dots + e_k D_k + c_{2k+1} X_{2k+1} + \dots + c_n X_n \\ \rightarrow & d_1 L_1 + \dots + d_k L_k + f_1 D_1 + \dots + f_k D_k + d_{2k+1} X_{2k+1} + \dots + d_n X_n, \end{aligned}$$

entonces existe $j \leq r$ tal que R_j es igual a

$$\begin{aligned} & e_1 L_1 + \dots + e_k L_k + c_1 D_1 + \dots + c_k D_k + c_{2k+1} X_{2k+1} + \dots + c_n X_n \\ \rightarrow & f_1 L_1 + \dots + f_k L_k + d_1 D_1 + \dots + d_k D_k + d_{2k+1} X_{2k+1} + \dots + d_n X_n. \end{aligned}$$

En la situación anterior, se dice que la pareja (R_i, R_j) es una pareja dual de reacciones. Además también se dice que la red Ω es una red **pseudoquiral** o quiral de orden k . Si la reacción R_i es dual consigo misma, se dice que es autodual y será listada dos veces en el conjunto de reacciones $\{R_1, \dots, R_r\}$.

La red Ω_F mostrada en un ejemplo anterior es una red quiral de orden 1. Las reacciones $L \rightarrow 2L$ y $D \rightarrow 2D$ son reacciones duales, mientras que la reacción $L + D \rightarrow \square$ es autodual. Esta red es conocida como la red de Frank y fue el primer modelo de síntesis enantioselectiva en la literatura. [14] Hasta el día de hoy, se considera que todos los modelos de homoquiralidad biológica propuestos en la literatura son redes pseudoquirales. Esto es debido a que la restricción de simetría que define la clase de redes quirales es la restricción de simetría más débil que debe ser satisfecha por cualquier modelo factible de homoquiralidad biológica. Si una red Ω no es pseudoquiral, entonces hay asimetría en la estructura de Ω que favorece la síntesis de especies en uno de los dos grupos y esto implica que el mecanismo de síntesis modelado por Ω no puede ser considerado un modelo de síntesis enantioselectiva.

Definición 3.6. Sea

$$(\mathbf{a}, \mathbf{b}) = ([L_1], \dots, [L_k], [D_1], \dots, [D_k], X_{2k+1}, \dots, [X_n], k_1, \dots, k_r)$$

un estado de Ω . Se dice que este estado es racémico si se satisfacen las siguientes condiciones: Para todo $i \leq k$ se tiene la igualdad $[L_i] = [D_i]$. Si R_i y R_j son reacciones duales, entonces $k_i = k_j$. Se define la diferencia enantiomérica total de (\mathbf{a}, \mathbf{b}) como

$$ee(\mathbf{a}, \mathbf{b}) = |[L_1] - [D_1]| + \dots + |[L_k] - [D_k]|,$$

y se define el rompimiento espontáneo de la simetría espejo como la producción de diferencias enantioméricas grandes a partir de perturbaciones arbitrariamente pequeñas de los estados estacionarios racémicos.

Con todo lo anterior, el problema se puede definir ahora como

Problema: Dada una red pseudoquiral \mathcal{M} , identificar si \mathcal{M} es capaz de producir SMSB

CARACTERIZACIÓN ALGEBRAICA DE UN SMSB: LA CONDICIÓN MM

Se ha identificado el tipo de redes que se desea analizar, describiendo de manera precisa el problema que se quiere resolver. En esta parte, se verá como analizar matemáticamente el problema. Desde ahora, se supondrá que las concentraciones de especies quirales permanecen constantes. Esta última suposición permite enfocarse en la dinámica de las especies quirales. Además, se puede prescindir de las concentraciones constantes de las especies quirales como parámetros cinéticos, puesto que el efecto de estas se puede incluir en las constantes de las reacciones. Así, dada una red quiral de orden k , se supondrá que esta red contiene únicamente las correspondientes $2k$ especies quirales.

Sea Ω la red quiral $(\{L_1, \dots, L_k, D_1, \dots, D_k\}, \{R_1, \dots, R_r\})$. El análisis de estabilidad lineal provee herramientas para calcular el conjunto de estados inestables de Ω . Sin embargo, estas herramientas obtienen cualquier estado inestable de Ω y el problema indica explícitamente que dichas estabilidades deben producir SMSB. Entonces, es necesario hacer un análisis de estabilidad enfocado en este fenómeno. A continuación se enuncia la definición de estado estacionario:

Definición 3.7. Sea

$$\Theta = (\{X_1, \dots, X_n\}, \{R_1, \dots, R_r\})$$

una red de reacciones químicas. Se dice que $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^n \times \mathbb{R}^r$ es un estado estacionario de Θ si se satisfacen las igualdades $\sum_{j=1}^r \mathbf{b}[j] v_{ij} (\mathbf{a}[1]^{c_{ij}} \cdots \mathbf{a}[n]^{c_{nj}}) = 0$; para $i = 1, \dots, n$. El símbolo \mathbf{a} se usa para denotar la sub tupla de (\mathbf{a}, \mathbf{b}) conformada por las concentraciones de n especies químicas y el símbolo \mathbf{b} para denotar la sub tupla conformada por las r constantes de reacción.

Nótese que los estados estacionarios de Θ son un equilibrio químico, es decir, que si se inicializa el proceso químico en un estado estacionario, no ocurrirá nada y el sistema permanecerá en un punto fijo de su dinámica.

Observación 3.8. Sea Ω la red quiral anterior. El símbolo $\mathcal{J}_\Omega(\mathbf{a}, \mathbf{b})$ se usa para denotar al jacobiano de Ω en el estado (\mathbf{a}, \mathbf{b}) .

Definición 3.9. Sea (\mathbf{a}, \mathbf{b}) un estado estacionario racémico de Ω . Se dice que (\mathbf{a}, \mathbf{b}) es un amplificador quiral si existe un valor propio de $\mathcal{J}_\Omega(\mathbf{a}, \mathbf{b})$ tal que $\text{Re}(\lambda) \geq 0$ y existe $v \in \mathbb{R}^{2k}$ tal que v es un valor propio de $\mathcal{J}_\Omega(\mathbf{a}, \mathbf{b})$ asociado a λ y tal que

$$(v[1], \dots, v[k]) \neq (v[k+1], \dots, v[2k]).$$

En la referencia [2] se afirma que los amplificadores quirales de Ω son los estados de Ω que producen SMSB. En lo que resta de esta subsección se estudiará más detalladamente dicho resultado.

Definición 3.10. El subespacio racémico de \mathbb{R}^{2k} es el subespacio de dimensión k $\mathcal{R}^k = \{(u, u) : u \in \mathbb{R}^k\}$. Se dice que $b \in \mathbb{R}^{2k}$ es un vector racémico si $v \in \mathbb{R}^k$.

Definición 3.11. Sea C una matriz de dimensión $2k \times 2k$. Se dice que C es una matriz quiral si existen dos matrices de dimensión $k \times k$, A y B , tales que

$$C = \begin{pmatrix} A & B \\ B & A \end{pmatrix}.$$

Se usa el símbolo $[A, B]$ para denotar a la matriz C y se dice que $[A, B]$ es la pareja de k -bloques de la matriz C .

De las definiciones se puede concluir fácilmente el siguiente lema.

Lema 3.12. Sea (\mathbf{a}, \mathbf{b}) un estado estacionario racémico de Ω . Se tiene que $\mathcal{J}_{\Omega(\mathbf{a}, \mathbf{b})}$ es una matriz quiral.

Las matrices quirales tienen propiedades interesantes, como la siguiente:

Lema 3.13. Sea $C = [A, B]$, se tiene que $\det(C) = \det(A + B) \cdot \det(A - B)$.

El lema anterior implica que el espectro de $[A, B]$ está completamente determinado por el espectro de las matrices $A + B$ y $A - B$.

Teorema 3.14. Sea $C = [A, B]$. Se tiene que $p_C(\lambda) = p_{A+B}(\lambda) \cdot p_{A-B}(\lambda)$, donde dada una matriz cuadrada M , el símbolo $p_M(\lambda)$ indica su polinomio característico.

Demostración. La matriz simbólica $[A, B] - \lambda \cdot I_{2k}$ también es quiral puesto que se tiene que

$$[A, B] - \lambda \cdot I_{2k} = [A - \lambda \cdot I_k, B - \lambda \cdot I_k].$$

Entonces, $\det([A, B] - \lambda \cdot I_{2k}) = \det(A - \lambda \cdot I_k) \cdot \det(B - \lambda \cdot I_k)$. \square

Observación 3.15. Se usa el símbolo $\text{Spec}(M)$ para denotar el espectro de la matriz cuadrada M .

Como corolario del teorema anterior se tiene la igualdad

$$\text{Spec}([A, B]) = \text{Spec}(A + B) \cup \text{Spec}(A - B).$$

También se tiene el siguiente resultado:

Teorema 3.16. Supóngase que C es igual a la matriz quiral $[A, B]$. Las dos afirmaciones siguientes se cumplen:

1. Sean $\lambda \in \text{Spec}(A - B)$ y v un vector propio de λ . Se tiene que

$$C \cdot \begin{pmatrix} v \\ -v \end{pmatrix} = \lambda \cdot \begin{pmatrix} v \\ -v \end{pmatrix}.$$

2. Supóngase que $\lambda \in \text{Spec}(A + B)$ y que $\lambda \notin \text{Spec}(A - B)$, entonces todos los vectores propios generalizados de C que están asociados con λ son vectores racémicos.

Definición 3.17. Se dice que una matriz cuadrada M es inestable si existe $\lambda \in \text{Spec}(M)$ tal que $\text{Re}(\lambda) \geq 0$.

Observación 3.18. Sea (\mathbf{a}, \mathbf{b}) un estado racémico de Ω . La matriz $\mathcal{J}_\Omega(\mathbf{a}, \mathbf{b})$ es una matriz quiral. Se usan los símbolos $A_\Omega(\mathbf{a}, \mathbf{b})$ y $B_\Omega(\mathbf{a}, \mathbf{b})$ para denotar a la pareja de matrices bloque de la matriz $\mathcal{J}_\Omega(\mathbf{a}, \mathbf{b})$. Así $\mathcal{J}_\Omega(\mathbf{a}, \mathbf{b}) = [A_\Omega(\mathbf{a}, \mathbf{b}), B_\Omega(\mathbf{a}, \mathbf{b})]$.

Teorema 3.19. Sea (\mathbf{a}, \mathbf{b}) un estado racémico de Ω . Se tiene lo siguiente: El estado racémico (\mathbf{a}, \mathbf{b}) es un amplificador quiral, si y sólo si, la matriz $A_\Omega(\mathbf{a}, \mathbf{b}) - B_\Omega(\mathbf{a}, \mathbf{b})$ de dimensión $k \times k$ es inestable. Si el estado racémico (\mathbf{a}, \mathbf{b}) no es un amplificador quiral, entonces no puede producir SMSB.

Se tiene que la inestabilidad de $A_\Omega(\mathbf{a}, \mathbf{b}) - B_\Omega(\mathbf{a}, \mathbf{b})$ es una condición necesaria para la amplificación quiral. Se usará el término *condición MM* para denotar esta última condición. En este punto cabe considerar la siguiente pregunta: ¿La inestabilidad de la matriz $A_\Omega(\mathbf{a}, \mathbf{b}) - B_\Omega(\mathbf{a}, \mathbf{b})$ también es una condición suficiente para la amplificación quiral? Si $A_\Omega(\mathbf{a}, \mathbf{b}) - B_\Omega(\mathbf{a}, \mathbf{b})$ es inestable, existe un vector propio completamente asimétrico de $[A_\Omega(\mathbf{a}, \mathbf{b}), B_\Omega(\mathbf{a}, \mathbf{b})]$ que es tangente

a la variedad inestable en (\mathbf{a}, \mathbf{b}) . Lo último implica que esta variedad inestable contiene una órbita que es tangente a ella, pero la cual se mueve por una trayectoria que es ortogonal a \mathcal{R}^k . Sin embargo, esto solo se puede asegurar a corto plazo y no es suficiente para asegurar la amplificación quiral a nivel macroscópico. Parece difícil encontrar condiciones suficientes para el rompimiento de simetría dado que es una propiedad a largo plazo de la dinámica.

Con todo lo anterior, se puede redefinir el problema inicial de la siguiente manera:

Problema: Dada una red pseudoquiral \mathcal{M} , identificar si \mathcal{M} satisface la condición MM y en dicho caso calcule un amplificador quiral para \mathcal{M} .

3.3. CONEXIÓN ENTRE LA GEOMETRÍA ALGEBRAICA REAL Y EL ANÁLISIS DE REDES PSEUDOQUIRALES

Se ha podido reescribir el problema inicial de manera matemática, por lo que ahora es pertinente preguntarse si existe alguna manera algorítmica de resolver el problema. En lo que queda de esta sección se indicará de qué forma, dada una Ω red quiral de orden k , se puede calcular una definición semialgebraica de los amplificadores quirales de Ω usando la condición MM.

Defínase L_Ω^1 como la lista de las condiciones mostrada a continuación:

- Condiciones de estado estacionario:

$$\begin{aligned} 0 &= \sum_{j=1}^r k_j v_{1j} ([X_1]^{c_{1j}} \cdots [X_n]^{c_{nj}}) \\ &\vdots \\ 0 &= \sum_{j=1}^r k_j v_{nj} ([X_1]^{c_{1j}} \cdots [X_n]^{c_{nj}}). \end{aligned}$$

- Condiciones de no negatividad:

$$L_1, \dots, L_k, D_1, \dots, D_k, X_{2k+1}, \dots, X_n \geq 0, L_1 + \dots + L_k > 0, k_1, \dots, k_r > 0.$$

- Las condiciones racémicas:

$$L_1 = D_1, \dots, L_k = D_k.$$

Como se puede observar, la lista L_Ω^1 conforma un sistema semialgebraico, pues está conformada por un sistemas de ecuaciones y desigualdades polinomiales. No obstante, para resolver el problema describir las condiciones que hacen que en una red pseudoquiral se produzca SMSB. Para obtenerlas se hace uso del *Criterio de estabilidad de Routh-Hurwitz* para analizar la estabilidad de la matriz simbólica $A_\Omega(\mathbf{s}) - B_\Omega(\mathbf{s})$ (ver la referencia [30]).

Definición 3.20. Sea $p(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ un polinomio de grado n con coeficientes reales y supóngase que $a_{n-1} \neq 0$. La tabla de Hurwitz asociada $p(\lambda)$ se define de la siguiente manera:

$$H(p(\lambda)) = \begin{bmatrix} a_n & a_{n-2} & a_{n-4} & \cdots & 0 \\ a_{n-1} & a_{n-3} & a_{n-5} & \cdots & 0 \\ H_{3,1} & H_{3,2} & H_{3,3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{n+1,1} & H_{n+1,2} & H_{n+1,3} & \cdots & 0 \end{bmatrix},$$

donde $H_{i,j}$ para $3 \leq i \leq n+1$ y $j \geq 1$ se define recursivamente como

$$H_{i,j} = H_{i-2,j+1} - \frac{H_{i-2,1}}{H_{i-1,1}} \cdot H_{i-1,j+1}.$$

El teorema de Routh-Hurwitz enuncia que el número de cambios de signo en la primera columna de la tabla de Hurwitz es igual al número de raíces del polinomio asociado con parte real no negativa. Es decir que si $p(\lambda)$ es el polinomio característico de una matriz, se tendrá que dicha matriz es inestable cuando el número de cambios de signo en la primera columna de la tabla de Hurwitz es mayor que cero. Así, para verificar la condición MM, se debe calcular la tabla de Hurwitz asociada al polinomio característico de la matriz $A_\Omega(\mathbf{s}) - B_\Omega(\mathbf{s})$ y contar el número de cambios de signo en la primera columna; si dicho número es mayor que cero, entonces se debe concluir la red quiral da lugar a SMSB. Note que si $p(\lambda)$ proviene de una matriz simbólica, los elementos de la primer columna de H no serán polinomios, sino funciones racionales. Sin embargo, se debe recordar que el signo de $\frac{p(X_1, X_2, \dots, X_m)}{q(X_1, X_2, \dots, X_m)}$ es el mismo de $p(X_1, X_2, \dots, X_m) \cdot q(X_1, X_2, \dots, X_m)$.

De esta manera la primer columna de H se puede reescribir de manera tal que sus elementos sean polinomios en función de los parámetros de la matriz simbólica. Además, el primer elemento de esta columna será el coeficiente principal del polinomio característico, el cual está dado por $(-1)^n$. Claramente, multiplicar por este valor a todos los elementos de la primer columna no alterará el número de cambios de signo. Hacer este procedimiento hace que se evite revisar el primer elemento de H , así como establecer de manera única el sistema semialgebraico. Sea H' el vector columna dado al hacer las dos alteraciones anteriores a la primer columna de H y al descartar su primer elemento. Por lo tanto, si Ω es una red quiral de orden k , la condición MM se puede traducir al lenguaje semialgebraico como

$$\left(\bigvee_{1 \leq i \leq k} H'_i \right) < 0.$$

Se denotará por RHC_Ω^i a la i -ésima condición de Routh-Hurwitz.

Con lo anterior, y teniendo en cuenta se se cuenta un algoritmo para resolver sistemas semialgebraicos, se propone el siguiente algoritmo para solucionar el problema de encontrar las condiciones que hacen que una red quiral Ω de orden k pueda exhibir SMSB:

1. Calcule la lista L_Ω^1 , así como las condiciones RHC_Ω^i , para $1 \leq i \leq k$.

2. Defina sol_{Ω} como una lista vacía.
3. Para $i \leq k$, defínase $SA_{\Omega}^i = L_{\Omega}^i \cup RHC^i$. Use el algoritmo de Collins para revisar la factibilidad del sistema SA_{Ω}^i . Si encuentra soluciones, defina sol_{Ω}^i como el conjunto de soluciones y agréguelo a sol_{Ω} .
4. Si no se encuentra la solución para ninguno de tales sistemas, devuelva el mensaje *La red Ω no produce SMSB*. En caso contrario, imprima sol_{Ω} .

Para finalizar esta sección, se mencionan algunos aspectos acerca de la eficiencia del algoritmo anterior:

- El algoritmo de Collins tiene una complejidad doblemente exponencial en función al número de parámetros de entrada.
- El paso 3 debe llevarse a cabo de esa manera, dado que al tomar la conjunción de todas las condiciones de Routh-Hurwitz, se obtiene una expresión demasiado compleja para que el algoritmo de Collins pueda resolverla.
- Calcular la tabla de Hurwitz no es una tarea fácil, pues requiere del cálculo de los coeficientes del polinomio característico de una matriz simbólica. Usualmente esta tarea es llevada a cabo calculando los menores principales de la matriz, es decir, que se debe trabajar con *determinantes* cuyo tamaño depende del tamaño de la red.

4

USO DE SACNA Y ALGUNOS EXPERIMENTOS COMPUTACIONALES

Para conquistar la frustración, uno debe permanecer intensamente enfocado en el resultado, no en los obstáculos. T.F. Hodge

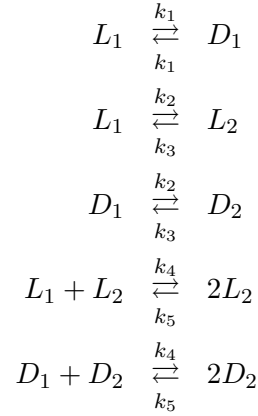
Esta sección describe algunos resultados (computacionalmente) experimentales del algoritmo expuesto en la sección anterior, por medio de la herramienta computacional SACNA (Semialgebraic Chemical Network Analyzer). Las dos secciones anteriores sirven como fundamento teórico para el desarrollo de la herramienta computacional SACNA. Esto implica que las dos secciones anteriores describen las limitaciones y el contexto en el cuál se debe usar esta herramienta. SACNA se desarrolló como un paquete en Mathematica[®], el cual contiene dos maneras de hacer el análisis y un simulador. El uso de SACNA se muestra por medio del análisis de algunos modelos extraídos de la literatura relacionada. Los cinco modelos escogidos representan los posibles resultados al usar SACNA para analizar redes quíricas:

- Cuando el algoritmo termina sin encontrar una solución, determina que la red no tiene ningún amplificador quíric. (Modelo de Calvin)
- Cuando el algoritmo encuentra soluciones para todas las condiciones de Routh-Hurwitz. (Modelo de Calvin-LES)
- Cuando el algoritmo encuentra soluciones para una de las condiciones de Routh-Hurwitz. (Modelo replicador de Hochberg-Ribó)
- Cuando la red ingresada es demasiado compleja, haciendo que la mayoría de las condiciones Routh-Hurwitz sean imposibles de analizar o no tengan soluciones en el corto plazo. (Modelo APED)
- Cuando la red ingresada es demasiado compleja, haciendo que todas de las condiciones Routh-Hurwitz sean imposibles de analizar o no tengan soluciones en el corto plazo. (Modelo APED reversible)

4.1. UN PRIMER EJEMPLO - EL MODELO DE CALVIN

Melvin Calvin, uno de los pioneros de la evolución química, introdujo un modelo de red de homociclicidad biológica [6], la cual está descrita de manera detallada en el esquema 4.1. Esta red corresponde, como entrada a SACNA, como la primera de las cinco situaciones posibles listadas

anteriormente: el algoritmo no puede encontrar soluciones para los dos sistemas semialgebraicos obtenidos de la red, lo que permite establecer que este modelo no tiene amplificadores quirales. Así, SACNA permite asegurar de manera formal que el modelo de Calvin es incapaz de exhibir SMSB.

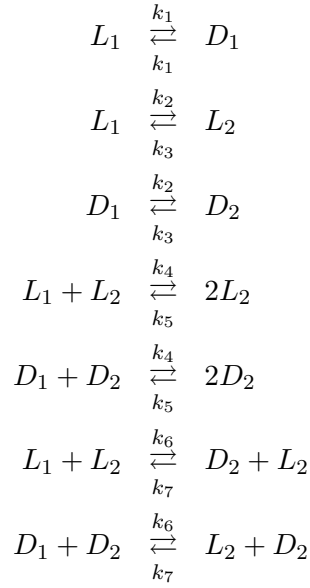


Esquema 4.1: Modelo de Calvin

4.2. UNA VERSIÓN MODIFICADA DEL MODELO DE CALVIN: *ENANTIOSELECTIVIDAD LIMITADA (LES)*

Se puede modificar ligeramente el modelo de Calvin al agregar cuatro nuevas reacciones y serán llamadas reacciones de enantioselectividad limitada (LES, por sus siglas en inglés). Al hacer esta modificación, se obtiene la red descrita por el esquema 4.2 y el cual se denotará por el símbolo Ω_C . Incluir estas cuatro reacciones hace que el modelo sea capaz de exhibir SMSB.

4.2 Una Versión modificada del Modelo de Calvin: Enantioselectividad Limitada (LES)



Esquema 4.2: Modelo de Calvin LES

Se debe notar que el modelo de Calvin es una red quiral de orden 2, por lo que al calcular las desigualdades de Routh-Hurwitz, se obtienen 2 sistemas semialgebraicos diferentes: el sistema $SA_{\Omega_C}^1$ y el sistema $SA_{\Omega_C}^2$. SACNA pudo encontrar soluciones para ambos. Se simuló la dinámica de Ω_C usando tales soluciones como condiciones iniciales y se observó que se produce SMSB incluso en el caso donde no se perturbaban ligeramente las condiciones iniciales. Esto es causado por los errores numéricos introducidos por el algoritmo de simulación, el cual no acepta números irracionales como parámetros.

Las soluciones del sistema $SA_{\Omega_C}^1$ llevan, casi siempre, a una dinámica oscilatoria. Las soluciones de este sistema están parametrizadas de la siguiente forma:

$$\begin{aligned}
 D_1 &> 0 \wedge 0 < D_2 \wedge k_1 > 0 \wedge k_2 > 0 \wedge k_3 > 0 \wedge k_4 > \frac{2k_1+k_2+k_3}{D_1-D_2} \\
 &\wedge 0 < k_5 < \frac{D_1k_4-D_2k_4-2k_1-k_2-k_3}{2D_2} \\
 &\wedge 0 < k_6 < \frac{D_1k_4-D_2k_4-2D_2k_5-2k_1-k_2-k_3}{D_1+D_2} \\
 \wedge k_7 &= \frac{D_1D_2k_4+D_1D_2k_6+D_1k_2+D_2^2(-k_5)-D_2k_3}{D_2^2} \\
 &\wedge L_1 = D_1 \wedge L_2 = D_2
 \end{aligned}$$

Si se toman los valores

| Parámetro | D_1 | D_2 | L_1 | L_2 | k_1 | k_2 | k_3 | k_4 | k_5 | k_6 | k_7 |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------------|-------------------------|
| Valor | 769 | 62 | 769 | 62 | 24 | 74 | 63 | 11 | 29 | $\frac{5}{11}$ | $\frac{1341048}{10571}$ |

se obtiene el comportamiento mostrado en la figura 4.3. Se puede verificar que para estos valores de parámetros el jacobiano \mathcal{J}_{Ω_C} tiene dos valores propios con partes imaginarias no nulas y partes

reales positivas. Así, se tiene que tales amplificadores quirales no son bifurcaciones Hopf pero producen una dinámica oscilatoria.

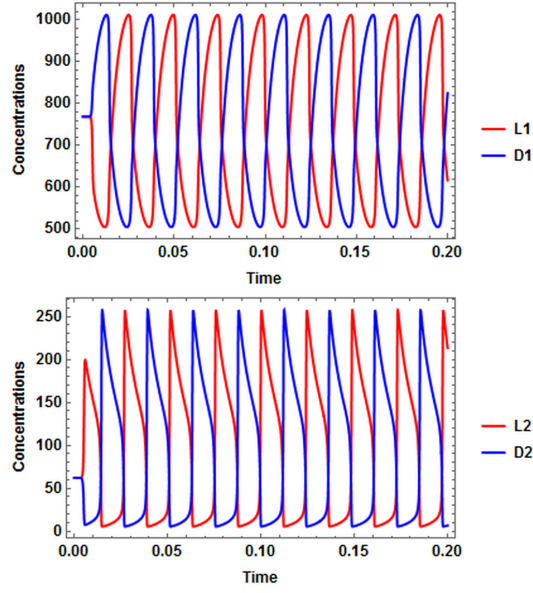


Figura 4.3: Simulación para algunos valores del sistema $SA_{\Omega_C}^1$

Las soluciones de $SA_{\Omega_C}^2$ están dadas por las condiciones

$$\begin{aligned}
 &D_1 > 0 \wedge 0 < D_2 \wedge k_1 > 0 \wedge k_2 > 0 \wedge k_3 > 0 \wedge k_4 > \frac{k_3}{D_1} \\
 &\wedge 0 < k_5 < \frac{-k_3 + D_1 k_4}{2D_2} \\
 &\wedge 0 < k_6 < \frac{-k_1 k_3 + D_1 k_1 k_4 - 2D_2 k_1 k_5}{D_1 k_1 + D_1 k_2 + D_2 k_3 + 2D_2^2 k_5} \\
 &\wedge k_7 = \frac{D_1 k_2 - D_2 k_3 + D_1 D_2 k_4 - D_2^2 k_5 + D_1 D_2 k_6}{D_2^2} \\
 &\wedge L_1 = D_1 \wedge L_2 = D_2
 \end{aligned}$$

El comportamiento en este segundo caso es el comportamiento típico de bifurcación que suelen mostrar los amplificadores quirales. Por ejemplo, al escoger los valores

| Parámetro | D_1 | D_2 | L_1 | L_2 | k_1 | k_2 | k_3 | k_4 | k_5 | k_6 | k_7 |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------------|---------------------------|
| Valor | 769 | 62 | 769 | 62 | 24 | 74 | 63 | 11 | 29 | $\frac{2}{53}$ | $\frac{12396201}{101866}$ |

Se obtiene el comportamiento mostrado en la figura 4.4

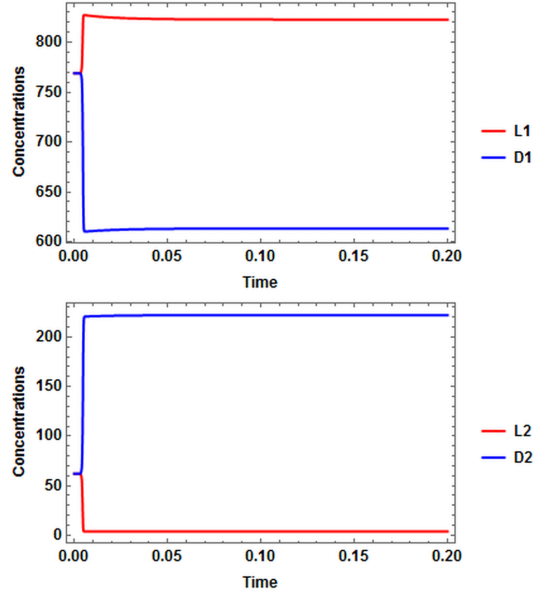
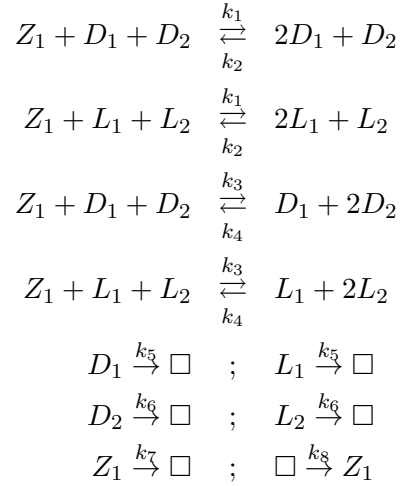


Figura 4.4: Simulación para algunos valores del sistema $SA^2_{\Omega_C}$ donde se muestra el típico comportamiento de *bifurcación*

4.3. MODELO REPLICADOR DE HOCHBERG Y RIBÓ

Hochberg y Ribó propusieron el modelo representado por el siguiente esquema (ver la referencia [22]):



Esquema 4.5: Modelo replicador de Hochberg y Ribó

Esta es una red quiral de orden 2, la cual se denotará con el símbolo Ω_{HR} . SACNA puede calcular fácilmente las dos condiciones de Routh-Hurwitz correspondientes. Al hacerlo, se obtienen dos sistemas semialgebraicos: el sistema $\text{SA}_{\Omega_{\text{HR}}}^1$ y el sistema $\text{SA}_{\Omega_{\text{HR}}}^2$.

El sistema $\text{SA}_{\Omega_{\text{HR}}}^1$ no tiene soluciones, mientras que para el segundo SACNA encontró las siguientes soluciones:

$$\begin{aligned}
D_1 &> 0 \wedge 0 < D_2 \wedge k_1 > 0 \wedge k_2 > 0 \wedge k_3 > 0 \wedge k_4 > 0 \\
\wedge k_5 &> \frac{D_2^2 k_1 k_4}{k_3} \wedge k_6 = \frac{D_1^2 D_2 k_2 k_3 - D_1 D_2^2 k_1 k_4 + D_1 k_3 k_5}{D_2 k_1} \\
\wedge k_7 &> \frac{2D_1^2 D_2 k_2 k_3 - 2D_1 D_2^2 k_1 k_4 + 2D_1 k_1 k_5 + 2D_1 k_3 k_5}{k_1} \\
\wedge k_8 &= \frac{D_2 k_1}{-k_5 - D_1 D_2 k_2} \left(-\frac{2k_3 (D_1 (-k_5) - D_1^2 D_2 k_2)}{k_1} \right) \\
&+ \frac{D_2 k_1}{-k_5 - D_1 D_2 k_2} (-2 (D_1 (-k_5) - D_1^2 D_2 k_2) - 2D_1^2 D_2 k_2 - 2D_1 D_2^2 k_4 - k_7) \\
\wedge Z_1 &= \frac{D_1 D_2 k_2 + k_5}{D_2 k_1} \wedge L_1 = D_1 \wedge L_2 = D_2
\end{aligned}$$

Si se escogen algunos valores que satisfagan las condiciones de arriba, se obtiene siempre un amplificador quiral que produce SMSB. Por ejemplo se escogieron los valores:

| | | | | | | | | | | | |
|-----------|-------|-------|-------|--------|-------|-------|-------|-------|-------|-------|-------|
| Parámetro | D_1 | D_2 | L_1 | L_2 | Z_1 | k_1 | k_2 | k_3 | k_4 | k_5 | k_6 |
| Valor | 0,25 | 0,5 | 0,25 | 0,5 | 1 | 1 | 1 | 1 | 1 | 0,375 | 0,125 |
| | | | | k_7 | | k_8 | | | | | |
| | | | | 1,3125 | | 1 | | | | | |

Los resultados de la simulación se muestran en la figura 4.6

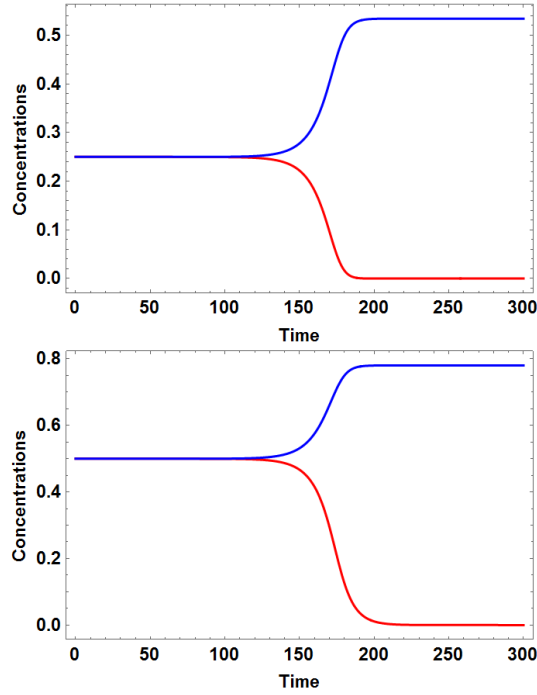
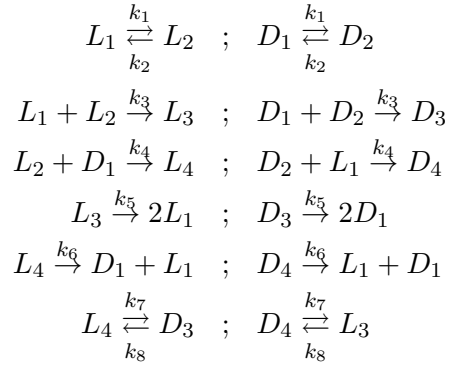


Figura 4.6: Simulación para el modelo replicador de Hochberg y Ribó

4.4. EL MODELO APED

SACNA puede ser utilizado para analizar modelos más grandes. Se considerará el modelo APED propuesto en [4], cuyo esquema se muestra a continuación.



Esquema 4.7: Modelo APED

Se han analizado cuatro sistemas semialgebraicos. SACNA hizo este trabajo y concluyó que el primero de ellos no tiene soluciones. El segundo y el tercero de dichos sistemas son demasia-

do complejos para que el algoritmo pueda procesarlos. Sin embargo, el cuarto de ellos pudo ser analizado, obteniendo una respuesta que es demasiado compleja para incluirla aquí. Dicha respuesta se puede encontrar en el siguiente enlace: <https://github.com/ddfulaa/SACNA/blob/main/Tutorial/SACNATutorial4-TheAPEDmodel.pdf>. Se tiene, por ejemplo, el amplificador quiral usando los siguientes valores de los parámetros:

| Parámetro | D_1 | D_2 | D_3 | D_4 | k_1 | k_2 | k_3 | k_4 | k_5 | k_6 |
|-----------|-------|-------|-------|-------|-------------------|---------------------|-------------------|-----------------------|-------|-------------------|
| Valor | 532 | 49 | 10 | 9 | 90 | 47 | $\frac{111}{116}$ | $\frac{42739}{53998}$ | 2523 | $\frac{20347}{9}$ |
| | | | | | k_7 | k_8 | | | | |
| | | | | | $\frac{2080}{63}$ | $\frac{2339}{2030}$ | | | | |

Los resultados de las simulaciones de este amplificador pueden verse en las siguientes gráficas:

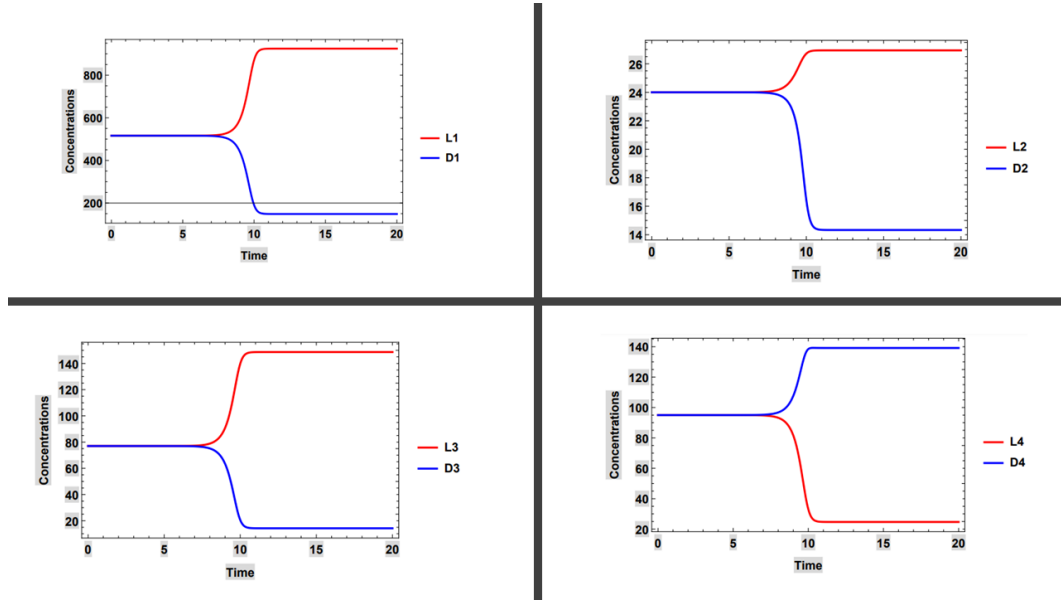


Figura 4.8: Simulación para el modelo APED

Además, los autores de [2] afirman que las condiciones $k_3 \neq k_4$ y $k_5 \neq k_6$ son necesarias, en esta red específica, para que se produzca el rompimiento de simetría. Esto no es del todo correcto y SACNA permite encontrar amplificadores quirales que satisfacen la igualdad $k_5 = k_6$. Se tiene por ejemplo la solución dada por:

| Parámetro | D_1 | D_2 | D_3 | D_4 | k_1 | k_2 | k_3 | k_4 |
|-----------|-------|-------|-------|-------|---------|---------|----------|-----------|
| Valor | 52 | 62 | 29 | 1,25 | 30 | 13,751 | 0,158616 | 0,0608131 |
| | | | | | k_5 | k_6 | k_7 | k_8 |
| | | | | | 23,3864 | 23,3864 | 227 | 4,03178 |

La dinámica producida por este amplificador quiral se muestra en la figura 4.9.

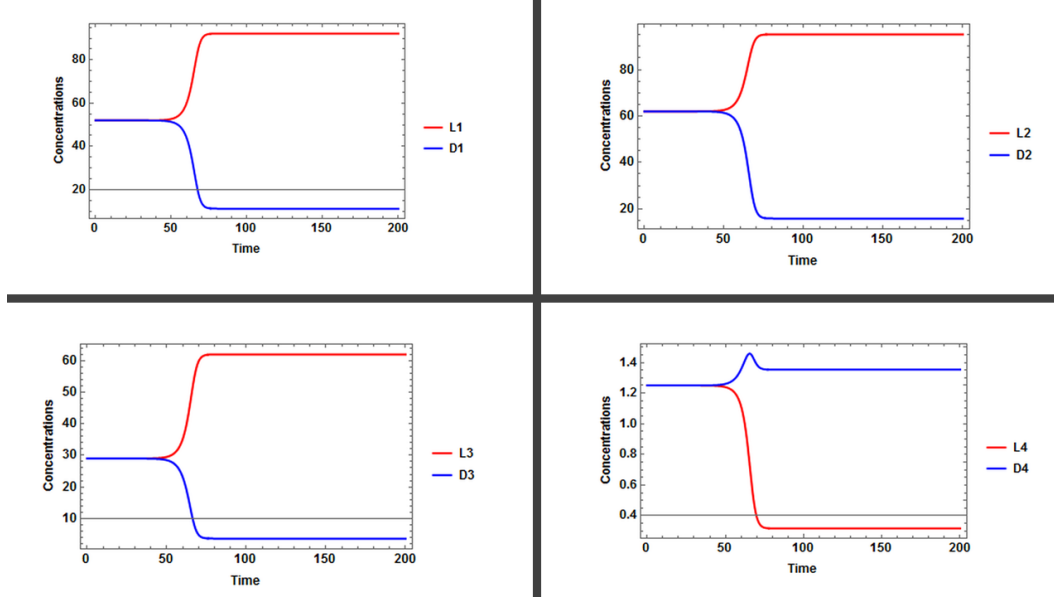
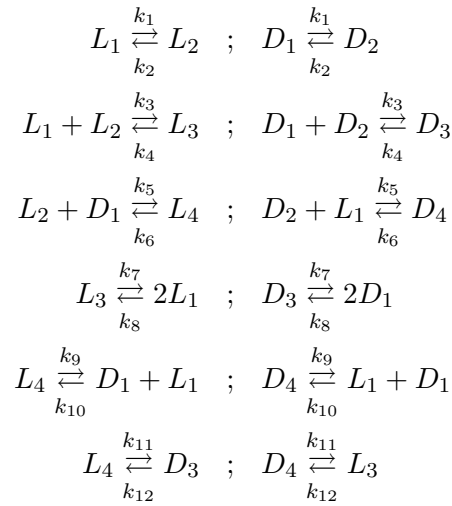


Figura 4.9: Simulación para el modelo APED con $k_5 = k_6$

De esta forma, SACNA permite arreglar algunos errores que ocurren en la literatura relacionada.

4.5. EL MODELO APED REVERSIBLE

El modelo APED contiene reacciones irreversibles. Si se agregan las acciones reversibles, se puede obtener el modelo APED reversible. El modelo está descrito por el esquema 4.10.



Esquema 4.10: Modelo APED reversible

Uno se puede preguntar si la irreversibilidad es necesaria para la ocurrencia de SMSB en el modelo APED. Este nuevo modelo es más complejo y SACNA tiene problemas para procesar esta red. Esta red tiene 4 condiciones de Routh-Hurwitz y el algoritmo se estanca al hacer el análisis de tanto la segunda como la tercera. No obstante, es posible enfocarse únicamente en la primera y la cuarta condición. Al analizar el sistema con la primera condición, se obtiene una respuesta negativa y al analizar la cuarta, nuevamente el algoritmo se estanca. Se ha observado experimentalmente que la primera o la última condición de Routh-Hurwitz suelen dar resultados afirmativos, por lo que vale la pena hacer una simplificación de la condición 4 al introducir valores para algunos parámetros. Por ejemplo, considérense los siguientes valores:

| Parámetro | D_1 | D_2 | k_1 | k_2 | k_3 | k_4 | k_9 | k_{10} | k_{11} | k_{12} |
|-----------|-------|-------|-------|-------|-------|----------------------|-------|----------|----------|---------------|
| Valor | 1 | 1 | 1 | 1 | 1 | $\frac{30761}{1401}$ | 1 | 1 | 1 | $\frac{3}{2}$ |

Dichos valores hacen que el algoritmo encuentren una solución particular para este modelo. La solución es demasiado compleja para ponerla aquí, pero en el siguiente enlace es posible consultarla: <https://github.com/ddfulaa/SACNA/blob/main/Tutorial/SACNATutorial5-ThereverseAPEDmodel.pdf>. Por ejemplo, se pueden obtener los siguientes valores para los parámetros restantes:

| D_3 | D_4 | k_5 | k_6 | k_7 | k_8 |
|-----------|----------|----------|---------|-------|----------|
| 0,0243732 | 0,285855 | 0,740741 | 4,21749 | 53 | 0,577636 |

El comportamiento de este sistema se muestra en la figura 4.11

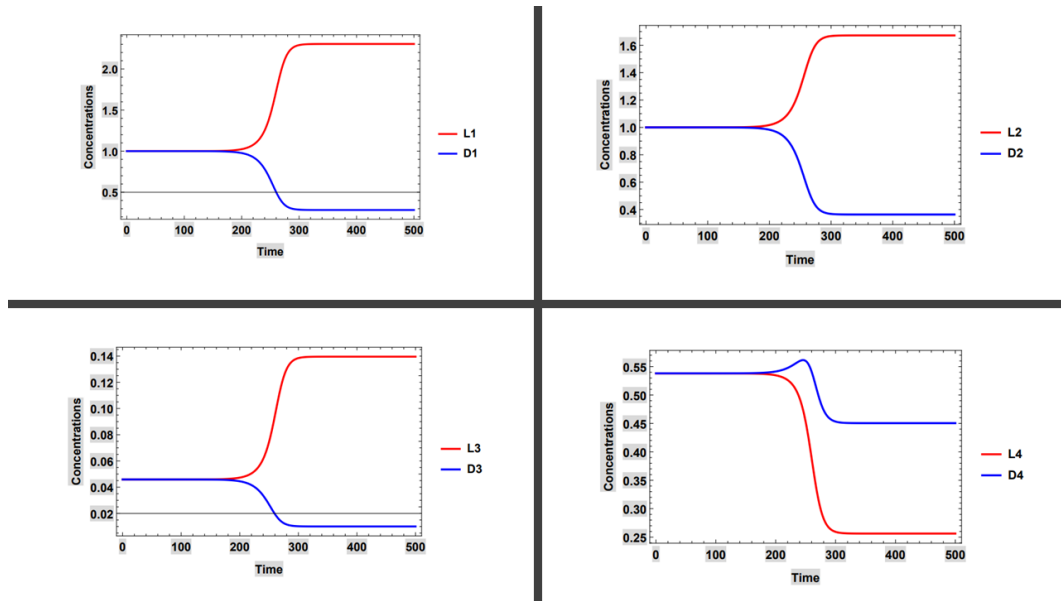


Figura 4.11: SMSB del modelo APED reversible con los valores dados

5 CONCLUSIONES

Esta tesis tiene 2 objetivos principales, el primero de ellos es mostrar una aplicación del algoritmo de Collins al análisis de modelos de redes químicas en las que se buscan condiciones del rompimiento de simetría, lo cual fue expuesto en la sección 3. Este tema ya había sido tratado en artículos anteriores por otras personas [ver la referencia [27] y [2]]. En particular lo expuesto en esta tesis puede verse como un resumen de la referencia [27]. El segundo objetivo consiste en la implementación de una herramienta que llevara a la práctica el algoritmo descrito al final de la sección 3. Este objetivo se cumple al realizar y mostrar la implementación del paquete SACNA, el cual puede ser descargado de <https://github.com/ddfulaa/SACNA>. Para el análisis de redes pseudoquirales, existía anteriormente una implementación (ver la referencia [1]) de las ideas expuestas en [2], pero tal implementación es un poco débil en comparación con la herramienta SACNA, pues no hace uso del algoritmo de Collins para resolver el sistema semialgebraico. Una ventaja de usar el algoritmo de Collins es que es un algoritmo exacto y formal, en el sentido que el resultado obtenido por medio de este no es una aproximación. La implementación del software se realizó usando Mathematica[®] debido a la solidez de su paquete simbólico. Cabe resaltar que es difícil encontrar buenas implementaciones del algoritmo de Collins y debido a que ya está implementado en Mathematica[®] como el comando `CylindricalDecomposition`, hizo que la implementación del programa de análisis de redes químicas SACNA fuese más simple. Sin embargo, considero que, antes de iniciar la elaboración de un sistema a partir de herramientas ya implementadas, es necesario conocer al menos de manera superficial el funcionamiento de dichas herramientas. Esto permite saber cuáles son los límites del nuevo sistema que se creará. Tal necesidad fue suplida en la exposición de la sección 2. Como toda herramienta, SACNA tiene limitaciones, las cuales se listan a continuación:

- El usuario debe escoger una de las sub-condiciones de Routh-Hurwitz del sistema que SACNA analizará, ya que intentar analizarlas todas a la misma vez, hace que la herramienta sea prácticamente inutilizable. No obstante, debido a que la condición de Routh-Hurwitz es una conjunción de las sub-condiciones, se podría paralelizar este proceso y detener el la búsqueda cuando se encuentre una solución en alguna de las sub-condiciones.
- Existen problemas a la hora de analizar redes químicas de tamaño moderado. Esto es debido a que tiene como base el algoritmo de Collins, el cual tiene una complejidad doblemente exponencial con respecto al número de variables. Aún así, se implementó una versión que permite al usuario reducir el número de variables al introducir valores para algunos de ellos.
- Se sabe que las descomposiciones obtenidas dependen del orden de las variables de entrada, lo cual puede influir a la hora de obtener una respuesta. En SACNA se usa un orden específico de las variables, por lo que es muy probable que dicho orden no sea óptimo para todas las redes que se intenten analizar. Hay trabajos interesantes con respecto a la manera

de escoger el orden de las variables, por ejemplo está en la referencia [8], en donde usan redes neuronales artificiales.

- Mathematica[®] requiere de una licencia para su uso, lo cual hace que el uso de SACNA se vea limitado.
Hay buenas implementaciones libres del algoritmo de Collins, como la que puede ser encontrada en <https://www.usna.edu/CS/qepcadweb/B/QEPCAD.html> o usando el software REDUCE. De hecho, la primera de ellas se puede llamar desde otro potente software, alternativo a Mathematica[®], llamado SAGE, haciendo que la idea que se muestra aquí pueda ser implementada en una versión libre. Hay otros paquetes de cálculo simbólico libres, pero en ninguno de ellos está implementado el algoritmo Collins, como por ejemplo Sympy(Python). Aún así, este tiene implementadas muchas herramientas que se mencionaron en esta tesis, como por ejemplo el cálculo de los coeficientes subresultantes principales o la representación de números algebraicos, haciendo viable una posible implementación del algoritmo y de una posible herramienta similar a la que se muestra aquí.

A pesar de lo anterior, SACNA permite confirmar de manera práctica lo expuesto en la sección 3, lo que lleva a pensar que la condición MM podría ser también una condición suficiente para la amplificación quiral. En la tesis también se mostró el análisis de los modelos de Calvin, Calvin LES (con enantioselectividad limitada), el modelo de Hochberg-Ribó, el modelo APED y su versión reversible. Por medio del simulador que tiene implementado, se corroboró que las soluciones encontradas por el algoritmo corresponden con el comportamiento esperado. Además, se observó curiosamente que aunque el modelo Calvin-LES no es una bifurcación de Hopf, produce dinámicas oscilatorias.

Por último, como se menciona en la referencia [29], para resolver sistemas semialgebraicos también existen los llamados métodos de punto crítico. Estos permiten calcular al menos un punto en cada componente semialgebraicamente conexa y es mucho más eficiente que el algoritmo de Collins al momento de determinar si un determinado sistema tiene solución (siendo simplemente de complejidad exponencial, lo que comparado con la complejidad doble exponencial del algoritmo de Collins es una mejora importante). Podrían usarse estos métodos de 2 maneras diferentes: 1) Como algoritmo inicial antes de buscar la solución exacta con el algoritmo de Collins, es decir como un verificador más eficiente para saber si el sistema tiene o no solución; o 2) Aplicarlo directamente al análisis de redes químicas de manera similar a la que se explicó aquí e intentar explorar los vecindarios de cada punto obtenido para buscar nuevas soluciones. Una breve descripción de dichos métodos puede ser consultada en las referencias [29] y [3].

SIGLAS

| | |
|------|--|
| CAD | Descomposición algebraica cilíndrica |
| CAS | Sistema de álgebra computacional |
| EPRS | Sucesión de residuos polinomiales euclidianos |
| PPRS | Sucesión de residuos polinomiales primitivos |
| PRS | Sucesión de residuos polinomiales |
| psc | Coficiente subresultante principal |
| SMSB | Rompimiento espontáneo de la simetría espejo |
| SPRS | Sucesión de residuos polinomiales subresultantes |

BIBLIOGRAFÍA

1. J. Agreda y E. Cruz. *Listanalchem*. URL: <https://gitlab.com/homochirality/listanalchem> (visitado 29-12-2021).
2. J. Agreda, E. Cruz y J. A. Montoya. «Computing the Parameter Values for the Emergence of Homochirality». *Life* 9(3) 3, 2019.
3. S. Basu, R. Pollack y M.-F. Roy. *Algorithms in Real Algebraic Geometry*. Springer, 2016.
4. H. Bersini, A. Commeyras y R. Plasson. «Recycling Frank: Spontaneous emergence of homochirality in noncatalytic systems». *Proc. Nat. Acad. Sci. USA* 101:48, 2004, págs. 16733-16738.
5. C. W. Brown. «Companion to the Tutorial Cylindrical Algebraic Decomposition», 2004.
6. M. Calvin. «Chemical Evolution: Molecular Evolution towards the Origin of Living Systems on the Earth and Elsewhere», 1969.
7. B. Caviness y J. Johnson. «Quantifier Elimination and Cylindrical Algebraic Decomposition». *Texts and Monographs in Symbolic Computation*, 1998.
8. C. Chen, Z. Zhu y H. Chi. *Variable Ordering Selection for Cylindrical Algebraic Decomposition with Artificial Neural Networks*. DOI: [10.1007/978-3-030-52200-1_28](https://doi.org/10.1007/978-3-030-52200-1_28).
9. G. Collins. «Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition». *Proceedings of Second GI Conf. Automata Theory and Formal Languages*, 1975.
10. M. Coste. *An Introduction to Semialgebraic Geometry*. 2002.
11. J. H. Davenport. *Computer Algebra (Notas de Clase)*. URL: <https://people.bath.ac.uk/masjhd/JHD-CA.pdf> (visitado 02-01-2022).
12. J. H. Davenport, Y. Siret y E. Tournier. *Computer Algebra: Systems and Algorithms for Algebraic Computation*. URL: <https://people.bath.ac.uk/masjhd/masternew.pdf> (visitado 02-01-2022).
13. M. Feinberg. «Complex balancing in general kinetic systems». *Arch. Rational Mech. Anal* 49, 1972, págs. 187-194.
14. C. Frank. «On spontaneous asymmetric synthesis». *Biochim. Biophys. Acta* 11, 1953, págs. 459-463.
15. J. Gal. «Molecular Chirality: Language, History, and Significance». *Topics in Current Chemistry* 340, 2013, págs. 1-20.
16. P. Gruber. *Convex and discrete geometry*. Springer-Verlag, New York, 2007.
17. H. Hong. «An improvement of the projection operator in cylindrical algebraic decomposition». *Proc. International Symposium on Symbolic and Algebraic Computation*, 1990, págs. 261-264.

18. F. Horn y R. Jackson. «General Mass Action Kinetics». *Archive Rational Mech.* 47, 1972, págs. 81-103.
19. M. Jirstrand. «Cylindrical Algebraic Decomposition: An introduction», 1995.
20. M. Kauers. «How to Use Cylindrical Algebraic Decomposition». *Séminaire Lotharingien de Combinatoire* 65, 2011.
21. W. Koepf. *Computer Algebra: An Algorithm-Oriented Introduction*. Springer, 2021. DOI: <https://doi.org/10.1007/978-3-030-78017-3>.
22. R. J. M., H. David, C. Joaquim, E.-H. Zoubir y M. Albert. «Spontaneous mirror symmetry breaking and origin of biological homochirality». *J. R. Soc. Interface* 14, 2017. DOI: <http://doi.org/10.1098/rsif.2017.0699>.
23. S. McCallum. «An improved projection operation for cylindrical algebraic decomposition of three-dimensional space». *Journal of Symbolic Computation* 5, 1988, págs. 141-161.
24. S. McCallum. «An improved projection operator for cylindrical algebraic decomposition. In Quantifier Elimination and Cylindrical Algebraic Decomposition». *Texts and monographs on Symbolic Computation*, 1998.
25. U. Meierhenrich. *Amino acids and the asymmetry of life caught in the act of formation*. Springer, Berlín, 2008.
26. B. Mishra. *Algorithmic Algebra*. Springer., 1993.
27. J. Montoya y C. Mejía. «On the Algebraic Analysis of Chiral Amplification in Chemical Reaction Networks». *MATCH Commun. Math. Comput. Chem.* 82, 2019, págs. 5-28.
28. N. S. Nise. *Control Systems Engineering*. 7.^a ed. John Wiley & Sons, Inc., 2015. ISBN: 0-471-44577-0.
29. A. Quadrat y E. Zerz. *Algebraic and Symbolic Computation Methods in Dynamical Systems*. Springer, 2020. DOI: <https://doi.org/10.1007/978-3-030-38356-5>.
30. E. Routh. *A Treatise on the Stability of a Given State of Motion: Particularly Steady Motion*. Macmillan, NY, 1877.
31. S. Strogatz. *Nonlinear Dynamics and Chaos*. 2.^a ed. Westview Press, NY, 2011.