

UCH-FC  
MA6-BMCN  
R 459  
C.1



# **ESTUDIO SOBRE CLAUSURA METABÓLICA**

Tesis entregada a la Universidad De Chile en cumplimiento parcial de los  
requisitos para optar al grado de Magíster en Ciencia Biológicas con  
Mención en Biología Molecular, Celular Y Neurociencias

Facultad De Ciencias

Por

**BRYAN REYNAERT GODEFROY**

Director de Tesis: Dr. Juan Carlos Letelier

Co-Director de Tesis: Dr. Jorge Mpodozis

Diciembre 2013

Santiago - Chile

FACULTAD DE CIENCIAS

UNIVERSIDAD DE CHILE

INFORME DE APROBACION

TESIS DE MAGISTER

Se informa a la Escuela de Postgrado de la Facultad de Ciencias que la Tesis de Magister presentada por el candidato

BRYAN REYNAERT GODEFROY

ha sido aprobada por la comisión de Evaluación de la tesis como requisito para optar al grado de Magister en Ciencias Biológicas con mención en Biología Molecular Celular y Neurociencias, en el examen de Defensa Privada de Tesis rendido el día 22 de Noviembre de 2013.

Director de Tesis:  
Dr. Juan Carlos Letelier Parga

Co-Director de Tesis:  
Dr. Jorge Mpodozis Marín

Comisión de Evaluación de la Tesis

Dr. Ricardo Cabrera Faucar

Dr. Jorge Soto Andrade

*Letelier*  
.....  
*J*  
.....  
*[Signature]*  
.....  
*[Signature]*  
.....





Perhaps the first lesson to be learned from biology is that there are lessons to be learned from biology.

- Robert Rosen

# Agradecimientos

Le agradezco de corazón a todos los que me ayudaron en el camino e hicieron este momento posible para mí. Honestamente, una enumeración sería posiblemente más extensa que esta tesis, así que pido perdón por la frustrante generalidad con la que deberé referirme a ustedes. Mis amorosos padres, gracias a quienes nunca me faltó nada. Mis abuelos, una fuente inagotable de inspiración y mi hermana, quien con su amistad y cómplice rivalidad me mantuvo motivado. También estoy profundamente agradecido con mis familias postizas (Mark, Barros, Vergara, Manriquez, Oñate y prontamente Romero) que me aguantan y cuidan como a uno más. Mis amigos me perdonarán por no nombrarlos uno a uno; estoy en deuda infinita con todos ellos: me apoyan y acompañan inamoviblemente en esta y otras aventuras. Espero que esto sea siempre así y que sientan que yo los acompaño a ustedes de vuelta. Aprovecho este segundo de fama y permanencia histórica para declararles que los considero mi familia. Además quiero mencionar aquí a los excelentes profesores que he tenido la dicha de conocer; su calidad humana y genio académico han enriquecido mi vida inmensamente durante estos años. Especialmente debo mencionar al Profesor Letelier, quien ha sabido ser Profesor, Líder, Guía, Maestro Zen y aunque a él posiblemente no le guste que lo diga, Amigo. Sin su ejemplar espíritu épico no estaría aquí hoy.



# Índice general

<b>Introducción</b>	<b>1</b>
Antecedentes . . . . .	1
Las cinco escuelas descriptivas . . . . .	3
Vida = Forma . . . . .	3
Vida = Sistema que consume Negentropía . . . . .	6
Vida = ADN o un Algoritmo . . . . .	8
Vida = Máquina sumamente complicada . . . . .	11
Vida = Modo especial de Organización . . . . .	12
Clausura: Un punto de encuentro . . . . .	13
Biología Relacional . . . . .	14
Sistemas $(M,R)$ . . . . .	15
Autopoiesis . . . . .	17
Sistemas $(M,R)$ y Autopoiesis . . . . .	23
Clausura metabólica . . . . .	24
Dinámica Autopoiética . . . . .	28
Teorías formales . . . . .	29
Conjuntos auto catalíticos . . . . .	29
Conjuntos RAF . . . . .	33
Teoría de la Organización Química . . . . .	34
Teoría de las Redes de Reacción Químicas . . . . .	35
Análisis del Control Metabólico . . . . .	40
Impacto relativo de estas teorías . . . . .	41
El Juego de las Citas . . . . .	42

<b>Antecedentes topológicos o relacionales</b>	<b>45</b>
Reacciones y redes de reacción . . . . .	45
Rosen in a (long and hard) nutshell . . . . .	49
Conjuntos RAF . . . . .	62
Teoría de la Organización Química . . . . .	66
<b>Antecedentes Dinámicos</b>	<b>69</b>
Ley de Acción de Masas . . . . .	69
Estado Estacionario . . . . .	72
Análisis del Control Metabólico . . . . .	73
Teoría de la Organización Química . . . . .	80
Dinámica y COT . . . . .	81
COT y Conjuntos RAF . . . . .	87
<b>Hipótesis y Objetivos</b>	<b>90</b>
<b>Resultados</b>	<b>92</b>
Organización de lo vivo: RAO . . . . .	92
Teoremas de Clausura . . . . .	94
Teorema de la Existencia del lazo . . . . .	98
Teorema de la Hegemonía del Lazo . . . . .	101
Teorema de la Estabilidad del Lazo . . . . .	102
Dinámica del Lazo . . . . .	108
Dinámica Sistémica . . . . .	120
Teoría de la Variedad Central . . . . .	124
Más allá del Lazo . . . . .	130

<b>Aplicaciones</b>	<b>133</b>
Cadena y Lazo de Catalizaciones . . . . .	133
Un caso pedagógico . . . . .	136
Un Lazo Biológico . . . . .	146
<b>Discusión, Perspectiva y Proyecciones</b>	<b>149</b>
Recapitulación de resultados . . . . .	149
Importancia del <i>cLST</i> . . . . .	154
Relación con resultados previos . . . . .	155
<b>Conclusión</b>	<b>163</b>
<b>Bibliografía</b>	<b>170</b>
<b>Apéndice</b>	<b>171</b>

# Índice de figuras

1.	Jardín Osmótico . . . . .	4
2.	Portada de “De máquinas y seres vivos” . . . . .	18
3.	Autocatálisis . . . . .	30
4.	Clausura auto catalítica . . . . .	31
5.	Número de trabajos centrados en Autopoiesis o Sistemas (M,R) . . . . .	42
6.	Metabolismo visto como grafo centrado en moléculas . . . . .	47
7.	Conjunto de reacciones interpretado como un grafo simple . . . . .	54
8.	Diagrama de Rosen (pre-clausura) . . . . .	59
9.	Diagrama de la Intuición de Rosen . . . . .	60
10.	Modelo de clausura eficiente de Rosen . . . . .	61
11.	Modelo de Rosen con Operadores . . . . .	61
12.	Representación gráfica de un CRS . . . . .	64
13.	Cinética de Michaelis-Menten . . . . .	75
14.	Descomposición de una organización . . . . .	87
15.	Jerarquía de las distintas organizaciones . . . . .	89
16.	Cadena y Lazo de Catalizaciones . . . . .	97
17.	Sistema constituido por sólo un Lazo . . . . .	103
18.	Plano de Argand con algunos valores propios reales y complejos . . . . .	106
19.	Teorema del corrimiento espectral . . . . .	107
20.	Distintos regímenes dinámicos para un lazo bidimensional . . . . .	111
21.	Campo vectorial para un sistema en tres dimensiones . . . . .	123
22.	Cadena y Lazo de Catalizaciones . . . . .	134
23.	Campo vectorial de un sistema que contiene un lazo . . . . .	135



24.	Campo vectorial de un sistema que no contiene un lazo . . . . .	136
25.	Ejemplo de sistema RAF y cRAO a la vez . . . . .	137
26.	Dinámica de un sistema simple con lazo . . . . .	139
27.	Dinámica de un sistema simple sin lazo . . . . .	140
28.	Simulaciones de un sistema RAF . . . . .	141
29.	Análisis de bifurcación para un sistema cRAO . . . . .	143
30.	Simulaciones de sistemas didácticos . . . . .	145
31.	Simulaciones del sistema SAC . . . . .	147
32.	Resumen conceptual de esta tesis . . . . .	151
33.	Reacciones cíclicas . . . . .	160
34.	Análisis por <i>SKM</i> . . . . .	161



# Resumen

En esta tesis planteo un marco formal mixto que contempla tanto los aspectos estáticos como dinámicos del metabolismo y enfatiza las consecuencias de realizar la clausura metabólica, una propiedad reconocida como central en biología pero diversa y vagamente definida. Dentro de este formalismo concentro mi atención en un patrón de catalizaciones, el cual llamo lazo de catalizaciones, y derivó teoremas no sólo acerca de su existencia sino que además de su comportamiento dinámico, encontrando que este último da espacio para la estabilidad y adaptabilidad ampliamente reconocida en los sistemas biológicos. En particular formulé el teorema de la estabilidad del lazo (cLST), el cual se cumple para una amplia gama de sistemas y fundamenta la robustez biológica. Con esto traduzco los resultados teóricos y sobretodo las distintas nociones de clausura a un entendimiento más profundo de la conducta del metabolismo, lo que acerca la teoría al dominio experimental. Adicionalmente describo como se podría extender esta teoría. Los resultados obtenidos en esta tesis son profusamente contrastados con teorías alternativas y se concluye que hay un gran valor en la aproximación aquí tomada porque es simultáneamente general, fértil y se nutre y conecta con los resultados más importantes de la biología teórica de sistemas de los últimos 50 años.

# Abstract

In this thesis I pose a mixed formal framework which comprises both static and dynamical aspects of the metabolism and emphasises the consequences of realizing metabolic closure, a property known to be central to biology but at the same time diversely and vaguely defined. In this formalism I focus my attention on a particular catalization motif, which I call catalization loop, and I derive theorems not only about its existence, but on its dynamic behaviour, realizing that the catalization loop allows for biological stability and adaptability. Moreover, I formulate the loop stability theorem based on closure (cLST), which holds true for a wide range of systems and might be at the base of biological robustness. Thus, we translate theoretical results and more importantly, multiple definitions of closure into a deeper understanding of the behaviour of metabolisms, which narrows the gap between the theoretical and empirical domains. Additionally, I describe how this theory could be extended. The results of obtained in this work are vastly compared with alternative theories and it is concluded that the approximation followed here is of great value as it is general, fertile and it nourishes itself and connects to the most relevant results in theoretical systems biology of the last 50 years.

# Introducción

## Antecedentes

Los seres vivos, aunque tienen muchos menos componentes que una estrella (masa solar =  $M_{\odot} = 1.9891 \times 10^{30}$  kg), tienen muchos más *tipos* de componentes <sup>1</sup> y más procesos que una estrella. Entender como los sistemas vivos se originaron, funcionan y han resultado ser sistemas físicos increíblemente estables ha sido una de las más importantes preguntas de la ciencia occidental en los últimos 300 años. Esto no es sorprendente cuando se analiza a la luz de que los seres vivos se han mantenido, como clase de sistemas, por mas de 3.500.000.000 años y ocupando practicamente cada rincón del planeta. La búsqueda de este *secreto de los seres vivos*<sup>2</sup> ha sido intensa y no es nueva, pero justificada ya que los siguientes hechos apabullantes. En pleno periodo de los enciclopedistas franceses, en 1747, el médico Julien Offray de la Mettrie (1709-1751) publicó un pequeño panfleto (*L'Homme Machine*) donde afirmaba que el organismo humano era una máquina, sin secretos vitalistas, y que funcionaba esencialmente como un sistema (sofisticado) de engranajes transparentes (de La Mettrie y Thomson, 1996) <sup>3</sup>. Desde ese entonces, hace 263 años, muchas mentes se han dedicado a tratar de romper o descifrar el secreto de los seres vivos. Sorprendentemente, el secreto sobre la naturaleza de los seres vivos se ha aseverado muy duro de romper, a pesar de que se han usado aproximaciones muy distintas, desde la física, la química, la biología, la matemática y la filosofía. A continuación, y sólo para

---

<sup>1</sup>*E. coli* tiene aproximadamente 1200 metabolitos y 4300 proteínas = 5500 componentes distintos en sólo  $0.95 \times 10^{-12}$  g

<sup>2</sup>para ser preciso, con secreto me refiero a una definición de ser vivo que sea universal y operativa, y que permita derivar las propiedades generales más notorias de esta clase de sistemas

<sup>3</sup>Además publicó, en 1751 una obra pornográfica *L'Art de Jouir* que le obligó a una vida trashumante en varias cortes europeas hasta que murió de glotonería en la corte imperial Alemana.

fijar el contexto intelectual de esta tesis, sobrevolaré las principales líneas de ataque que ha sufrido este problema desde la obra de de la Mettrie. En todo caso, para los lectores que deseen saber más, aconsejo leer la siguiente revisión moderna (Letelier y cols., 2011).

Existen muchas teorías que pretenden entender lo que podría llamarse *el secreto de la vida*, pero muy poco consenso sobre su real utilidad en la biología contemporánea. Por ejemplo existe una enorme corriente de pensamiento que afirma que el secreto está en los ácidos nucleicos, especialmente en el ADN. Desde 1953 ese punto de vista se volvió mayoritario en los principales centros mínimamente sofisticados dedicados a la biología; al punto que la pregunta sobre la naturaleza especial de los seres vivos se transformó en una mala metáfora sobre autómatas y máquinas moleculares y la reflexión sobre el metabolismo se llenó de palabras como *código*, *programa*, *información*, *alfabetos* y hasta un innecesario *dogma central*<sup>1</sup>. Otra noción explicativa ampliamente invocada, pero muy rara vez explicada, ha sido asociar a los seres vivos propiedades como máquinas que se alimentan de *negentropía* (ver libro *What is Life?* de Erwin Schrödinger de 1944 (Schrödinger, 1944) así el maravilloso orden de los seres vivos se explicaría como una generación de orden (interno) a expensas de aumentar el desorden del resto del universo. Pero en paralelo a estas mega-ideas (Vida = ADN, Vida = *entropía negativa* o Vida = *Accionar de un autómata*) siempre ha existido una corriente de pensamiento que ha defendido que la propiedad esencial que caracteriza a los seres vivos es el *modo de organización* de su metabolismo. A continuación revisaré con mayor detenimiento cada una de estas corrientes con el fin de contextualizar el desarrollo de esta tesis y preparar las ideas preliminares que permitirán explicar porqué la clausura metabólica debiese estar al centro de la discusión sobre la organización biológica. En particular, para avanzar en los objetivos de esta tesis se hará una revisión de los puntos de encuentro y desacuerdo entre las teorías anteriormente mencionadas y se avanzará en el entendimiento del metabolismo utilizando la clausura

---

<sup>1</sup>dogma que, como todo en biología, cuenta con notables excepciones

metabólica como argumento central. Además se generarán resultados operativos para la biología experimental. Para lograr estos objetivos se establecerá un marco formal en el cual la clausura metabólica se exprese precisamente y simultáneamente se incorpore la dinámica temporal de los sistemas metabólicos. Así se entenderán las consecuencias de realizar la clausura metabólica a nivel de magnitudes empíricas (e.g. flujos metabólicos).

## Las cinco escuelas descriptivas

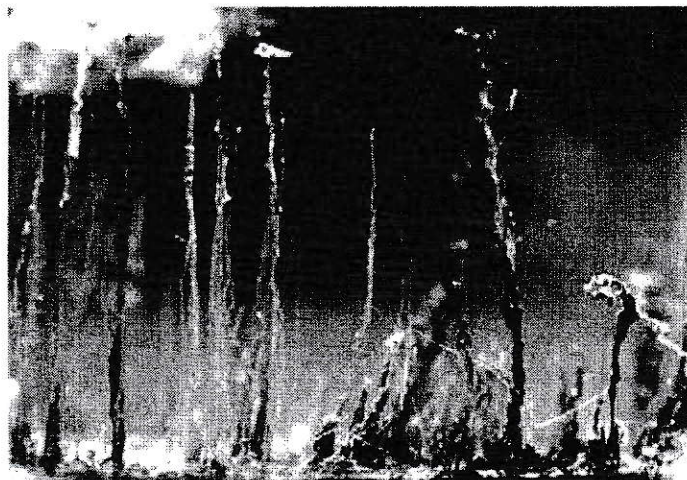
A continuación haré una revisión histórica y analítica de las distintas corrientes que han dominado la forma de entender la biología, muchas de las cuales se mantienen vigentes incluso hoy en día. Este proceso sirve para ilustrar la complejidad de la tarea y la diversidad de ideas característica del campo, contextualizar y plantear un vocabulario común, y crucialmente para resaltar porqué esta tesis se distingue de trabajos anteriores. He decidido clasificar a estas primeras cinco escuelas como descriptivas, pues por ahora no han logrado más<sup>1</sup> que describir el fenómeno de la vida. A pesar de contar con distintos énfasis, enfoques y léxicos, su aporte no ha permitido desentrañar o penetrar el secreto de la vida significativamente.

### Vida = Forma

Una tentativa notable y natural para entender a los seres vivos se debe al médico francés Stéphane Leduc(1853-1939) quien estudió los *bosques osmóticos* los cuales son crecimientos de cristales que asemejan formas biológicas como plantas, hongos o líquenes como los que se observan en la figura 1. Lo interesante de la visión de Leduc fue su foco hacia el metabolismo. Aunque no comprendió a cabalidad el mecanismo de crecimiento de sus formas osmóticas (sólo fueron entendidas hacia 1980 (Coatman y cols., 1980) y se

---

<sup>1</sup>sin intento de desconocer la utilidad y grandeza del impacto que han tenido



**Figura 1:** Jardín osmótico realizado en la Facultad de Ciencias utilizando una solución de silicato de sodio y distintas sales como sulfato de cobre. La morfología aparentemente biológica de estas estructuras inspiró a Stéphane Leduc a considerar la complejidad de los organismos vivos como resultado de una serie de procesos esencialmente químicos pero interconectados de forma particular.

han revisado recientemente por Cartwright y cols. (2002)) sí intuyó que la naturaleza de las reacciones de transformación eran importantes. De hecho Leduc fue una persona de pensamiento muy moderno como lo atestigua su libro (de 1912) *La biologie synthétique, étude de biophysique* (Leduc, 1912).

les êtres vivants sont formés des mêmes éléments et des mêmes énergies que le reste du monde. Les phénomènes de la vie suivent les lois physiques générales de la nature: la loi de la continuité, celle de la conservation de la matière, celle de la conservation de l'énergie. Les êtres vivants sont des transformateurs de forme, de matière, d'énergie. La vie est un phénomène physique (Capítulo III, (Leduc, 1912))

Claro que con 100 años de ventaja podemos darnos cuenta de lo moderno y errado que era el pensamiento de Leduc, sobretodo ya que su visión del metabolismo tenía como metáfora la química inorgánica y nunca consideró a las enzimas y el fenómeno de la catálisis (durante el resto de esta tesis explicaré porque esto es fundamental para incorporar

la noción central de clausura).

Otro aspecto importante de Leduc es su fijación en la morfología como criterio de lo viviente. Sin lugar a dudas, dado que en el desarrollo de esta tesis se han construido bosques osmóticos bastantes espectaculares, se afirma que la cautivación de Leduc por la riqueza de formas “orgánicas” que espontáneamente aparecen en estos bosques osmóticos es justificada. Esto lo llevo al convencimiento de que tenía entre ma un fenómeno fundamental para entender a los seres vivos ya que podía sintetizar su forma con reacciones inorgánicas. En esta tesis me muestro opuesto a esta corriente: la “forma” de un fenómeno no captura su “esencia”, así que el programa de Leduc aparece a nuestros ojos como irremediamente equivocado. En efecto, como el zorro de Antoine de Saint-Exupery creo que “lo esencial es invisible a los ojos”. Lo notable de esta pseudo-igualdad o falso isomorfismo es que aún después de 100 años de la obra de Leduc muchas personas del ambiente de las ciencias de la computación creen (al igual que Leduc) que si un programa produce una salida que se parezca a la forma de los seres vivos es que se ha capturado la esencia de estos sistemas. Ejemplos de estos programas son muchos sistemas de reacción-difusión (cuya belleza cautivó a Turing). Para ser justo, debo mencionar que esta corriente de pensamiento cuenta con la teoría celular entre sus filas! Si bien la teoría celular en su forma actual es reflejo de un pensamiento bastante más profundo e informado, el principio rector es el mismo. Nuevamente, en honor a la justicia, debo reconocer que los aportes de esta última teoría a la biología han sido mayúsculos.

Nuevamente, en un acto de consciencia, quiero recordar aquí al científico mexicano Alfonso Herrera (1868-1942) quién, en la década 1920/1930, expandía las ideas de Leduc e hizo experimentos más atrevidos para sintetizar sistemas vivos artificiales.

Por último me gustaria mencionar un pintoresco ejemplo de cómo esta corriente sigue vigente: El caso de las nanobacterias. Las nanobacterias (similarmemente a los nanobios) se postularon originalmente basadas en observaciones morfológicas pero posteriormente se

descartaron bajo fuerte evidencia. Resulta un entretenido ejercicio científico seguir esta literatura: Kajander y Çiftçioglu (1998); Urbano y Urbano (2007); Raoult y cols. (2008) o incluso este formativo artículo por García-Ruiz y cols. (2009)

## **Vida = Sistema que consume Negentropía**

La entropía, en su definición clásica por Rudolf Clausius (1822-1888) es una función de estado que sólo se puede computar para procesos reversibles de acuerdo a:

$$\Delta S = \int \frac{dQ_{rev}}{T} dx \quad (1)$$

La noción de entropía siempre ha ejercido un embrujo potente, sobre todo por leyes como que la entropía, en procesos espontáneos, sólo puede aumentar. Para ver un ejemplo del efecto que esta misteriosa noción puede tener incluso sobre autores del campo, cito textualmente a mi libro favorito de fisicoquímica (Castellan, 1987):

“... un extraño sentimiento que tiene la mayoría de las personas de que la entropía puede verse, sentirse, meterse en una botella, si pudieran ver el sistema desde un ángulo apropiado”

luego recomienda

“Podemos evitarnos noches de insomnio si, por el momento al menos, ignoramos la vaga pregunta ¿Qué es la entropía?”

y finalmente advierte

“Otros métodos de acción, como tratar de memorizar el máximo de ecuaciones posible, sólo produce pánico, parálisis y paranoia.”

Otro aspecto de atracción que tiene la entropía es que esta variable de estado es el resultado de la primera, y más efectiva, teoría de sistemas que se haya producido. No deja de ser notable que, al analizar la eficiencia de motores a vapor la ciencia del siglo



XIX se produjo una ciencia de sistemas que se terminó aplicando desde reacciones químicas, gracias a relaciones como  $dU = Tds - Pdv$ , hasta el universo, gracias al teorema  $H$ . Interesantemente para nuestros fines, la termodinámica usa muy pocas variables para describir sistemas (masa, temperatura, volumen, presión, energía interna, entropía, etc...). Por lo tanto, la termodinámica, en su versión clásica o en la versión estadística, no ve la micro-estructura interna (i.e. la red metabólica) de los sistemas que estudia, si no que observables macroscópicos. Como es de común acuerdo, uno de los genios más atormentados del siglo XIX (Ludwig Boltzmann (1844-1906) produjo un resultado fundamental que relaciona el mundo macroscópico con el microscópico a través de la entropía ( $S$ ) y el número de microestados correspondientes al macroestado observado  $\Omega$ :

$$S = k \cdot \ln(\Omega) \quad (2)$$

Esta ecuación, debido al concepto de número de microestados, ( $\Omega$ ) ha sido muy usada y ha llevado al abuso de establecer la errónea pero ampliamente invocada y enseñada equivalencia:

#### ENTROPÍA = DESORDEN

Esta relación, más el hecho obvio de que los seres vivos parecen ser la antítesis del desorden, ha hecho que muchas personas hayan tratado de hacer calzar (muchas veces a martillazos!) la noción de que los seres vivos –que son entidades organizadas– no sólo deben disminuir su entropía, si no que ésta es la actividad que los caracteriza como una clase especial de sistemas. Los principales propugnadores de esta metáfora son Erwin Schrödinger (1887-1961) en su libro *What is Life* (1944) y Walter Elsasser (1904-1991) con su idea de que existen leyes de la física, que se aplican a los seres vivos: leyes *biotónicas*. En esta tesis no exploraré ideas relacionadas con la entropía, ya sea en su versión clásica o en su versión estadística, pues creo que:

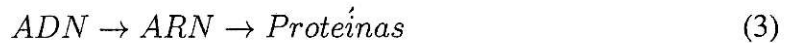
- Al ser un parámetro extensivo, no tiene la correcta “granularidad” para referirse a sistemas con tantas clases de componentes como los seres vivos.
- Es una variable que no se puede medir *usando las ideas estadísticas* al nivel de metabolismo celular (¿Cuál es el  $\Omega$  de *E. coli*?). De hecho, desde hace 20 años hay una muy interesante literatura que trata de medir la entropía del crecimiento bacteriano usando ideas de la más clásica fisicoquímica (Battley, 1993; Trevors, 2011).

Esto no significa que la escuela este muerta, al contrario, se ha extendido y reformulado con grandes expositores como Ilya Prigogine apoyándola. También me gustaría aclarar que de ninguna forma pretendo ignorar que el metabolismo debe seguir principios fisicoquímicos, pero creo que el enfoque no es el adecuado para comprender a los organismos vivos.

### **Vida = ADN o un Algoritmo**

Otra escuela de pensamiento, la predominante en la biología actual, identifica todo lo interesante de los seres vivos con la dinámica molecular de los ácidos nucleicos (ADN y ARN). Desde que se insinuó su función, en 1953 (Watson y Crick, 1953), y que se pudo describir el sorprendente mecanismo de la biosíntesis de proteínas hubo un verdadero *tsunami* que identificó (y sigue identificando) a “lo vivo” = “poseedor de ADN”. Hay que reconocer que los avances en conocimiento de la *biología molecular* (i.e. código genético, control de la expresión génica, el ribosoma como una máquina proteica) dan la apariencia que el ADN es absolutamente central. Poco ayuda a la hora de rebatir esta idea que este tipo molecular se encuentre en todos los sistemas vivos e incluso de alguna forma en aquellos que exhiben propiedades de lo vivo y ocupan una zona gris (e.g virus). Nada describe esta infatuación mejor que el llamado *dogma central* donde la parte efectora del

metabolismo (enzimas, proteínas estructurales) es sólo el *output* de la dinámica del ADN y ARN. Este dogma reza así:



Aunque se hayan escrito montañas sobre esta infatuación como noción central de lo vivo, obviamente no puedo más que estar en completo desacuerdo ya que este enfoque indica que la fenomenología de lo vivo es consecuencia de un componente y no de una red de relaciones. Justamente un punto de vista alternativo es que “lo vivo” es una consecuencia de cómo se organizan procesos, no de componentes en particular o de variables físicas y esta tesis es un avance microscópico en esta dirección. Además es siempre instructivo darse cuenta de que el “dogma central” acepta una lectura dual, donde el ADN/ARN son sólo cofactores de un proceso de replicación de proteínas.



En forma paralela al “culto al ADN” la segunda mitad del siglo XX inauguró el “culto al algoritmo”. En efecto el operar de todo el mecanismo de biosíntesis de proteínas se puede interpretar (“sobre-interpretar” dirían algunos) como una especie de máquina de Turing<sup>1</sup> operando sobre una cinta que contiene un programa codificado. Esta noción acoplada al auge infinito de la capacidad de cómputo, de almacenamiento y de visualización de los computadores modernos hizo que algunos en la comunidad de los físicos y de los informáticos generaran su propio “dogma” que se puede encapsular así:

Sistema Vivos = Manifestación física del operar de una cierta clase particular de algoritmos.

Esta noción ha sido muy aceptada desde muy temprano. Uno de los grandes íconos

---

<sup>1</sup>Alan Turing (1901-1954) el verdadero inventor de las ciencias de la computación

de la matemática del siglo XX (John von Neumann, 1903-1957) <sup>1</sup> de hecho hasta hizo un trabajo mostrando como una máquina, siguiendo un programa, podía auto-construirse (Von Neumann y Burks, 1966). Una interesante conjunción entre esta corriente y el enfoque morfológico son los sistemas-L (sistemas de Lindenmayer), utilizados para modelar el desarrollo y crecimiento de organismos vegetales (Rozenberg, 1980).

Desde entonces hasta ahora mucha gente se ha dedicado a generar teorías computacionales de los seres vivos. La muestra más palpable de esto es la creación de un nuevo campo de investigación: *Artificial Life (AL)* (Langton y Santa Fe Institute, Santa Fe(1989) <sup>2</sup>. El campo de ALIFE, sobre todo en EEUU, se centró demasiado en la metáfora de los computadores y hasta se generó una escuela alternativa con la *European Conference on Artificial Life* en 1991 (Varela y Bourguine, 1992). Pero esta metáfora no se ha demostrado fructífera. De hecho en 2012, en la conferencia de *ECAL* en París se pidió a los asistentes votar por cual había sido EL resultado del campo de *AL* en 25 años de vida y se seleccionó a un programa de computación (*TERRA* de Thomas Ray) que simula un proceso ecológico-evolutivo. De más esta decir que nadie, fuera de la estrecha comunidad de *Al* o *ECAL* ha escuchado hablar de este programa como de una “gran contribución”.

Mi opinión resumida es que esta corriente no es más que aplicar el razonamiento de la física tradicional a los irreducibles fenómenos biológicos. La obsesión por encontrar los mecanismos causales entre componentes logran un embrujo poderoso pues parecen muy explicativos, sin embargo creo que esto es una ilusión que puede mantenerse sólo mientras uno observe el sistema como partes de existencia y sentido independientes, y no como un todo. Cuando se da un paso hacia atrás aparece lo inabarcable por esta cosmovisión, los llamados comportamientos emergentes que son *inmanentes* a la biología.

Adicionalmente a esta búsqueda de algoritmos existe una “meta-discusión” acerca de

---

<sup>1</sup>La arquitectura de los computadores modernos se llama “Arquitectura de von Neumann”

<sup>2</sup>No deja de ser curioso que la conferencia inicial de este campo se haya efectuado en el laboratorio militar de *Los Alamos, Nuevo México*

si los sistemas vivos son simulables o no. Este es otro tópico que ha consumido a grandes mentes <sup>1</sup> pero yo sólo deseo dejar estampado aquí que: la simulabilidad de los seres vivos es un interesante problema con múltiples aristas, pero que no estará en el foco de mi tesis. En efecto, mi tesis busca no seguir “avivando polémicas” sino que obtener un resultado técnico que sea incontrovertible, y más o menos entendible por la gran mayoría, y que use las nociones de clausura metabólica.

### **Vida = Máquina sumamente complicada**

Sin lugar a dudas tratar de entender cuál es “EL” secreto de los seres vivos es un problema científico trascendente de primera magnitud y aunque estuvo mucho tiempo relegado a la periferia del *establishment*, se ha puesto en el centro mismo del pensar biológico debido al auge en la cantidad de datos que se están obteniendo. Por ejemplo, hacia 1995 cuando salió la primera versión del genoma humano algunos (optimistas) creyeron que al saber la secuencia de los 3200 millones de pares de bases de nuestra ADN se sabría TODO sobre el organismo. Pero gran fue su desilusión cuando se descubrió que saber la secuencia iluminaba muy poco los problemas fundamentales. Al poco andar algunos biólogos envueltos en genética, metabolismo y análisis cuantitativos redescubrieron algo esencial: la biología es una ciencia que trata de sistemas y no de partes. Así, esta colección de biólogos reduccionistas hizo suya las banderas del análisis sistémico y bautizó un nuevo campo: *Systems Biology*, nombre dado por Hiroaki Kitano (1961-) en el año 2000 (Kitano, 2001).

En la práctica todo el campo de *Systems Biology* es sólo una vuelta de tuerca más al problema visto por biólogos reduccionistas armados con una infinidad de datos, poder computacional y matemáticas de lo complejo, que sin embargo usan puntos de vista que ya llevan varias décadas de existencia, pero que aún rehúyen las ideas teóricas más avanzadas o abstractas.

---

<sup>1</sup>Mi director de tesis (JCL) ha tenido alguna participación en esta *trifurca* (Letelier y cols., 2003)

## Vida = Modo especial de Organización

Por razones que tienen un origen multifacético, y que no voy a analizar aquí, la segunda mitad del siglo XX fue muy fructífera respecto la consideración de los sistemas vivos como realizaciones de un modo particular de organización y aparecieron varias teorías o modelos como: Sistemas ( $M,R$ ) (Rosen, 1958a,b), La *Autopoiesis* (Maturana y Varela, 1973), el *Chemoton* (Ganti, 2003) y los *Conjuntos Auto catalíticos* (Kauffman, 1986) por mencionar lagunas. A pesar de que estas teorías poseen lenguajes distintos y enfoques teóricos diferentes es posible rescatar que todas tienen un crucial punto en común ya que consideran que la quintaesencia de los sistemas vivos radica en el hecho de que son redes de procesos que, como red, tienen la propiedad de fabricar los propios componentes que sustentan esta red de transformaciones. A esta propiedad se le denominará en esta tesis como *clausura metabólica*. Tal vez no sea una sorpresa que en más de medio siglo estas teorías hayan generado una abundante literatura pero –muy a pesar nuestro– ningún resultado relevante. La falta de relevancia no puede quedar mejor expresada por la casi completa obscuridad en la que viven estas nociones: los biólogos experimentales, *biólogos de mesón*, no sólo no las mencionan sino que desconocen su mera existencia. Este hecho encuentra su explicación probablemente en el lenguaje poco *operativo* en el cual se describe la clausura metabólica, alejándola así del mundo experimental. En este trabajo retomaremos el espíritu de esta escuela pero, aprendiendo del pasado, seguiremos un camino más técnico. En particular revisaremos un par de teorías formales que rescatan o reinterpretan los puntos esenciales de esta escuela pero esta vez basándose en matemáticas para tener herramientas de derivación de principios y rigor indiscutible. Prestaré particular atención a la noción de clausura en sus distintas formas ya que me parece una idea sumamente poderosa.

## Clausura: Un punto de encuentro

Antes de sumergirme en un mar de definiciones y teoremas considero adecuado revisar a profundidad algunas teorías que siguen la corriente de la organización. Haré esto con la intención de contraponerlas y mostrar como dentro de la misma escuela pueden realizarse exposiciones tan distintas sobre el mismo problema. Esta experiencia formativa además nos dará un vocabulario base y las nociones más relevantes de este punto de vista. Por otro lado, creo que un aporte de este trabajo, quizás no científico pero si histórico, es la interpretación, yuxtaposición y síntesis de distintas teorías desde un enfoque moderno. A continuación realizare una revisión en realidad somera, pues los detalles técnicos de los sistemas  $(M,R)$  serán visitados *in extenso* en el capítulo 2. Si me gustaria, a modo de guiar la lectura, indicar que la clausura es lo que deseo rescatar de estas teorías principalmente.

La noción de que los sistemas vivos se autoproducen, el núcleo de la idea de al autopoiesis, es un concepto que se encuentra en distintas formas en muchas teorías nacidas desde 1950s a la fecha. Resulta extraño que si bien todas hacen referencia a la clausura metabólica, sus definiciones y enfoques difieren bastante. Por lo tanto, ante esta diversidad de enfoques, es importante estudiar estas teorías y extraer cuales son los puntos de concordancia y en cuales difieren. Para dar un ejemplo de esta situación de acuerdo/desacuerdo, mencionaré que la autopoiesis trata sobre la clausura metabólica, los sistemas  $(M,R)$  se preocupan de la clausura a causas eficientes, y los formalismos más técnicos utilizan el término clausura o clausura catalítica. Si bien todos los conceptos apuntan a lo mismo, los lenguajes son distintos. Las similitudes y diferencias serán apreciables una vez revisemos con detalle estas exposiciones.



## Biología Relacional

Aunque Nicolas Rashevsky (1899-1972) tuvo una vida excepcional <sup>1</sup> y grandes intuiciones, la biología del siglo XX lo terminó olvidando casi totalmente. Rashevsky fue un físico ruso que, escapando del ejército rojo, llegó a EEUU hacia 1924. Trabajó inicialmente como físico, pero fue atraído a la biología ya que consideraba que el enfoque físico-matemático era lo que faltaba en la biología y que los biólogos no sabían matemáticas <sup>2</sup>. Con increíble ahínco, desde 1935 hasta 1965 trabajó en la University of Chicago donde: creó en 1938 la primera revista científica en biología teórica (*The Bulletin of Mathematical Biophysics*), y en 1947 el primer programa de doctorado en esta área. Su trabajo, en biología teórica, puede dividirse en dos partes. Entre 1938-1955 publicó muchos modelos sobre fenómenos tales como: división celular, conducción nerviosa, redes de neuronas, epidemias, redes de personas. Todos estos modelos, con nuestros ojos modernos, pueden parecer un poco delirantes *Mathematical Biophysics: Physico-Mathematical Foundations of Biology* (Rashevsky, 1960) ya que usaba un punto de vista *fisicista o ingenieril* un tanto extremo: división celular modelada como la separación de cargas eléctricas del mismo signo (!!!). No es de extrañar que muy poco de ese esfuerzo haya sobrevivido en la historia del pensamiento biológico. De hecho Norbert Wiener (1894–1964) en su influyente libro *Cybernetics: Or Control and Communication in the Animal and the Machine* (Wiener, 1948) le dedica un párrafo no muy diplomático, pero esencialmente ajustado a la realidad, al programa de investigación de Rashevsky.

Pero, después de 1953, se produce un cambio drástico en el punto de vista de Rashevsky, ya que abandona la idea de hacer modelos “ingenieriles” y pasa a considerar a la organización biológica como un sujeto de estudio *per se*. No se deben explicar los fe-

---

<sup>1</sup>oficial de la marina Zarista, físico, biólogo teórico, gran profesor de una gran universidad norteamericana

<sup>2</sup>un síndrome bastante común entre los físicos/matemáticos/informáticos que se dedican a la biología, pero no del todo errado



nómenos en forma disgregada *hay que buscar principios generales y buscar leyes de la organización* (Rashevsky, 1954). A estas ideas, que a pesar de no ser muy claras eran atrayentes y llenas de potencialidades, les puso el nombre de *Relational Biology* o *Topological Biology*. Desde ese año (1955) no cesó de escribir sobre este punto de vista, insistiendo que los sistemas biológicos se caracterizaban por *relaciones* especiales entre sus componentes. Una lectura actual de esos trabajos muestra ideas difíciles de entender, enseñar y aplicar, seguramente no sobrevivirían a los revisores modernos del *Journal Of Theoretical Biology*, pero lo importante es que la emoción de la línea de investigación cambió irreversiblemente. Desde entonces (i.e. 1955) Rashevsky y sus alumnos trataron de generar una teoría de la organización biológica. Como veremos este camino tuvo su exponente más ilustre en Robert Rosen (1934-1998). En la pluma de Rashevsky la teoría de la organización sólo queda como una petición de principio; nada concreto, el grito de alguien convencido. Una exposición sobre la real contribución de este “ruso blanco” puede encontrarse en: (<http://www.uvm.edu/pdodds/files/papers/others/2004/abraham2004a.pdf>)

### **Sistemas $(M,R)$ <sup>1</sup>**

En 1954, Robert Rosen ingresa como alumno de doctorado al programa de biología teórica creado por Rashevsky en la Universidad de Chicago (*Ph.D Program on Mathematical Biology: 1947-1965*). Rosen, proveniente de una familia judía de New York, entró al programa de doctorado entusiasmado con la idea de aplicar nociones de la mecánica cuántica a la biología. Pero justamente llega en el momento de la transformación del foco de investigación del indiscutido líder del grupo, del programa y editor en jefe del único Journal de biología teórica existente entonces. Seguramente deben haber sido meses y años fascinantes en Chicago. Un aspecto bastante único de esta situación debe haber sido, además de la personalidad de Rashevsky, la posibilidad de Rosen de interactuar con profesores

---

<sup>1</sup>y una compleja epistemología hyper-postmoderna

de la talla de Saunders Mac-Lane (1909–2005) uno de los creadores de la *Teoría de Categorías*; para leer una autobiografía se aconseja el siguiente trabajo: *Autobiographical Reminiscences* (Rosen, 2006).

A continuación veremos que la contribución de Rosen es absolutamente fundamental para el campo del estudio de la organización biológica, pero desgraciadamente la actividad académica alrededor de Rashevsky (y del programa en biología matemática) fue sólo una aventura personal de Rashevsky. Una vez que Rashevsky se alejó de la Universidad de Chicago dejaron de existir; el programa de doctorado, el *Bulletin*, y las ideas propiamente tales. En el 2010 nadie, en la Universidad de Chicago, se acordaba ya sea de Rashevsky, Rosen o el *Ph.D Program on Mathematical Biology* y hasta se había extraviado la tesis de doctorado de Rosen.

Rosen, hacia 1956-1957, durante la ejecución de su tesis de doctorado generó una enigmática idea que ocupará el centro de toda su vida intelectual. A continuación daré un resumen de lo que yo he entendido de lo que se puede denominar *Rosen's intuition*. Es importante dejar en claro que la lectura de Rosen es un ejercicio que requiere determinación y constancia. En efecto, Rosen escribió, absolutamente todos sus escritos, en un lenguaje difícil, oscuro y sin dar ejemplo alguno de sus ideas. No es de extrañar que, afuera de un muy pequeño grupo de iniciados (ver [www.panmere.com](http://www.panmere.com)), sus trabajos sean totalmente desconocidos, a pesar de haber escrito un libro con el ambicioso título de *Life Itself* (Rosen, 1991)<sup>1</sup>. Como dije anteriormente, revisaré sus ideas con bastante más detalle en el siguiente capítulo.

---

<sup>1</sup>En este punto recomiendo leer el trabajo de Letelier y cols. que trata de explicar que se debe entender cuando se lee a Rosen (Letelier y cols., 2006)

## Autopoiesis

Hace 40 años y bajo las difíciles circunstancias que imperaban en Chile en 1970-1973 <sup>1</sup>, los reconocidos biólogos chilenos Humberto Maturana (1928-) y Francisco Varela (1946-2001) escribieron un pequeño pero inmensamente influyente libro: *De Máquinas y Seres Vivos* (Maturana y Varela, 1973).

En ese libro se plasmó un punto de vista para entender a los seres vivos. Uno de los aspectos centrales de este punto de vista fue la noción de *autopoiesis* (auto-fabricación), es decir, los sistemas vivos se caracterizan por el hecho de que producen la mayoría de sus componentes a través de una red de procesos. Por esto se denomina que su organización metabólica es *circular*, o *cerrada sobre sí misma*. Es instructivo leer, con 40 años de perspectiva, el párrafo que define a las *máquinas autopoieticas* <sup>2</sup>.

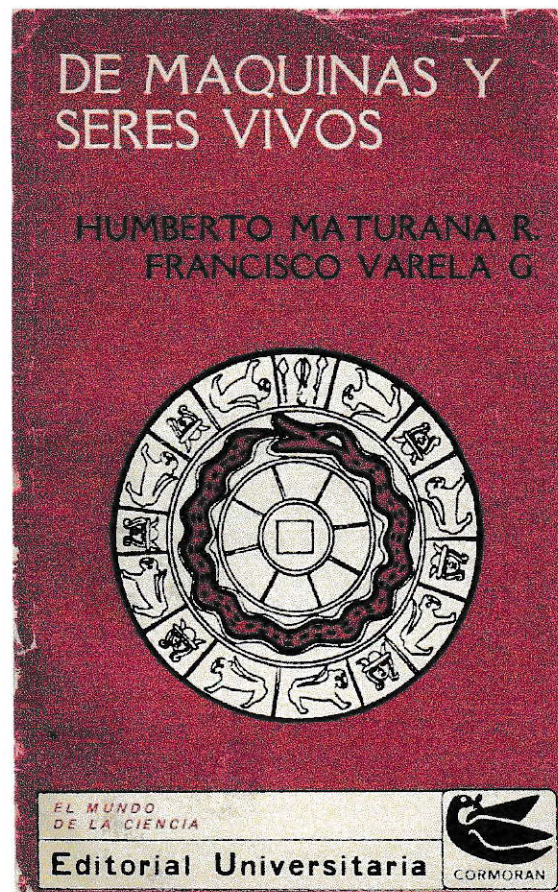
An autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components which: (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in space in which they (the components) exist by specifying the topological domain of its realization as such a network.

Esta definición consta de 2 cláusulas. La primera (i) es forma bastante compleja de definir clausura focalizándose en los procesos y no en los componentes. Es además una definición moderna ya que se refiere a una red de procesos. Esta cláusula es el núcleo conceptual de la autopoiesis. La segunda cláusula (ii) afirma que las máquinas autopoieticas son entidades en el espacio físico que, a causa de su propio metabolismo, se auto-encapsulan.

---

<sup>1</sup>claro que después de 1973 las circunstancias se harían bastante peores...

<sup>2</sup>La definición fue inicialmente escrita en inglés. El libro en español (Maturana y Varela, 1973) es una traducción de un documento escrito para el *Biological Computing Laboratory* de la Universidad de Illinois en 1972



**Figura 2:** Un pequeño libro rojo con muchas consecuencias. En 1973 apareció **De Máquinas y Seres Vivos** editado por la Editorial Universitaria. Su formato es muy pequeño, menor que un *Livre de poche*, debido a las restricciones existentes en el uso del papel prensa. Inicialmente pasó desapercibido, pero como dijo Stafford Beer “Este es un pequeño libro que contiene un universo”. En 2013 se cumplen 40 años de la publicación, mucho se ha escrito sobre las (muchas) ideas contenidas en este libro que van desde el entendimiento del metabolismo, el operar del sistema nervioso y una nueva manera de entender el fenómeno de la evolución biológica. Sin embargo, la influencia de estas ideas en el día a día de los biólogos ha sido mucho menor; se está aún en las etapas iniciales de la revolución augurada y propulsada por Maturana y Varela en 1973.

La influencia del concepto de autopoiesis ha sido enorme, contando con más de 15 publicaciones ISI en el 2010 y 10 en el 2011, cerca de medio millón (!!) de resultados en Google y media docena de patentes en Europa y Estados Unidos. Estas publicaciones son de todo tipo, desde sistemas sociales, pasando por economía y política organizacional, hasta su campo de origen, la biología. Sin lugar a dudas que la noción de clausura meta-

bólica, vista desde el punto de vista de la Autopoiesis, es una de las contribuciones más importancia la Facultad de Ciencias de la Universidad de Chile en sus 50 años de vida. Además de haber generado miles de referencias, se han producido media docenas de patentes <sup>1</sup>. No se puede evitar mencionar que la cláusula (i) –como fue escrita originalmente por Maturana y Varela– ha sido parafraseada muchas veces como una manera *interna* por la cual un sistema puede definir su propia realidad. De hecho en *De Máquinas Y Seres Vivos* es posible encontrar la siguiente propiedad de las maquinas autopoieticas:

the space defined by an autopoietic system is self-contained and cannot be described by using dimensions that define another space. When we refer to our interactions with a concrete autopoietic system, however, we project this system on the space of our manipulations and make a description of this projection.

Este párrafo parece indicar que “el espacio” (en su definición más amplia) donde existe un sistema autopoietico es definido por la dinámica interna de este sistema y además no es conmensurable con el espacio desde donde se describe el sistema. No es de extrañar que, con afirmaciones de este calibre, la autopoiesis se haya transformado en una idea usada por muchos campos. Un uso, extra-biológico, se puede encontrar en *The Autopoiesis of Architecture: A New Framework for Architecture (2010)* (Schumacher, 2011). Pero no es el único, de hecho, una de las aplicaciones con más seguidores se refiere a una interpretación “autopoietica” de los sistemas legales por Niklas Luhmann (1927-1998) y hasta existe una “teoría autopoietica” de la guerra y de la “Actor-Network Theory” (Actor-Network Theory and Autopoiesis: A New Perspective on Knowledge Management. Lars Steiner (University of Gävle, Sweden). Claro que tampoco sería justo olvidar teorías psicoanalíticas (Is

---

<sup>1</sup>ver patente: US 6553355, Autopoietic network system endowed with distributed artificial intelligence for the supply of high volume high-speed multimedia telesthesia telemetry, telekinesis, telepresence, telemanagement, telecommunications, and data processing services

Freud's model of the mind autopoietic?) o incluso de la resolución de conflictos. Como se puede ver, la utilización de la autopoiesis como metáfora ha sido (y continúa siendo) un campo muy fértil en *áreas alejadas de la biología*. No deseo hacer un análisis fino, pero yo estimo que esta popularidad se basa en el hecho de que, con la noción de autopoiesis se puede comenzar a hablar de auto-referencia y modos de relacionarse de unos sistemas con otros.

Pero mi interés no es hablar de autoreferencia, arquitectura, sistemas legales o del arte de la guerra. Mi interés está centrado en la biología y por lo tanto aparece la pregunta: ¿Cómo usar la autopoiesis en biología, el campo donde fue concebida?

Han existido varias tentativas para: a) matematizar la autopoiesis y b) aplicarla concretamente al metabolismo celular <sup>1</sup>. Debo dejar en claro que mi interés no es en “matematizar” algo por si mismo, si no que *entender* el metabolismo y ambas acciones no son sinónimos, más la matematización es un paso o herramienta poderosa para explorar y comunicar las consecuencias de distintos axiomas. Un problema para poder aplicar *verbatim* la definición de máquinas autopoieticas al metabolismo es la dificultad en definir *procesos* y *componentes* ambas nociones son claves y es siempre complicado definir las.

Un ejemplo de intento de formalizar la autopoiesis fue bastante temprano, hacia 1976, un alumno de la Facultad de Ingeniería de la Universidad de Chile (Aquiiles Limones) hizo su memoria de título sobre una visión de la autopoiesis como un sistema retro-alimentado. Una tentativa más interesante fue llevada a cabo por Francisco Varela quien, en 1979, publicó un libro (Principles of Biological Autonomy) donde usa el *cálculo indicacional* de Spencer Brown (1923-) (Spencer-Brown, 1972) para definir la idea de auto-referencia (Varela, 1979). Estas dos tentativas no atacan el problema de cómo hablar de metabolismo dentro del contexto de la autopoiesis, en cambio se centran en entender la idea de “auto-referencia”.

---

<sup>1</sup>se recuerda que ese fue el campo donde nació

Pero la simulación computacional de sistemas autopoieticos sí ha generado escuela y mucha. Hay que recordar que el *paper* original de Varela, Uribe y Maturana (Varela y cols., 1974) tenía una sección (*Protobio*) donde se simulaba un metabolismo muy simple consistente en tres reacciones. El operar de este metabolismo producía una encapsulación y tenía cierto grado de clausura. Esta simulación, que algunos han visto como iniciando el campo de *Artificial Life o Automatas Celulares* ha sido replicado por muchas personas desde entonces al punto de que se creó el campo de *Computational autopoiesis* (McMullin, 2004; McMullin y Program, 1997). Estas simulaciones son interesantes ya que en la práctica son simulaciones de redes metabólicas muy simples y capturan el hecho de la auto-encapsulación. Pero todas estas simulaciones son, para nuestro conocimiento moderno, bastante simples ya que sólo contienen pocas reacciones (del orden de 10) y usan químicas artificiales muy simples.

Una tentativa más interesante se puede encontrar en el trabajo de Pier Luigi Luisi (1938-) que ha tratado de hacer sistemas celulares artificiales basados en la autopoiesis y ha logrado interesantes resultados en la autocatálisis de la formación de lípidos (ver por ejemplo la conferencia: <http://www.pt-ai.org/smlc/2013>) y el libro *The Emergence of Life: From Chemical Origins to Synthetic Biology* (Luisi, 2006). Lo interesante es que la noción de autopoiesis, aún con todas sus imprecisiones, ha servido como principio heurístico para guiar los experimentos. Por ello no es de extrañar que se haya creado otro pequeño campo de investigación denominado: *Chemical Autopoiesis* (Luisi, 2012).

Lo que ha pasado con *Computational Autopoiesis* y *Chemical Autopoiesis* es un indicador del campo en su conjunto: Existen ideas interesantes (y atrapantes) que mucha gente cree que van a tener un gran impacto futuro. Pero ese impacto no se ha podido coagular en verdaderos resultados ya que (aparentemente) falta una idea fundamental para poder hablar con propiedad de términos como procesos, componentes, redes de procesos o autonomía. Pero, en términos conceptuales, existen cientos (sino miles) de papers

y documentos web que tratan de usar las ideas de la autopoiesis en una gran variedad de situaciones. Mi opinión es que muchas de estas tentativas son meras *metáforas* y, aunque tengan profundidad conceptual, no tienen una profundidad técnica.

Pero, dado que se podría pensar que estos comentarios son demasiado negativos *vis-a-vis* de la autopoiesis quiero dejar en claro que realmente considero muy importante la obra de Maturana y Varela por los siguientes motivos:

- Clausura como elemento definitorio de lo vivo
- Énfasis en el concepto de red de procesos
- Énfasis en la noción de encapsulamiento
- La idea de la autonomía como *derivada* de la noción de clausura
- La organización como característica diagnóstico de clase
- Su relación histórica con la cibernética y la importancia del *Biological Computing Laboratory* bajo la dirección de Heinz von Foerster (1911-2002)
- Su íntima relación con una teoría sobre el sistema nervioso. Poca gente sabe que el libro **De Máquinas y Seres vivos** es la expansión de una idea de Humberto Maturana sobre entender al sistema nervioso como una máquina que continuamente está generando correlaciones senso-motoras. Por ello, el libro que introduce la noción de autopoiesis al mundo anglosajón lleva el enigmático título de **Autopoiesis and Cognition** (Maturana y Varela, 1980)
- Su afán totalizador y unificador de la Biología; muchos creen que funda una nueva epistemología, y una nueva teoría del lenguaje. En todo caso, a pesar de los casi nulos avances (en el dominio de la acción científica predictiva) es normal esperar



que cada 10 años se publiquen trabajos con el sugestivo título de “autopoiesis: Una Re-Evaluación” (McMullin, 2004; Luisi, 2003; Razeto-Barry, 2012).

## **Sistemas (*M,R*) y Autopoiesis**

Una pregunta que se puede hacer es: ¿Que tanto conocían Maturana y Varela, hacia 1970-1973, de los sistemas (*M,R*)? Hay que enfatizar que tanto Rosen como Maturana/Varela hacían énfasis en la idea de clausura... y que en ese énfasis eran pioneros. ¿Leyeron Maturana/Varela los papers de Rosen?

Aunque parezca increíble Maturana y Varela desconocían completamente el trabajo de Rosen. Hay varias evidencias –y razones– de este desconocimiento:

- Como se dijo la idea central de la autopoiesis no proviene de una idea sobre el metabolismo, sino que sobre una manera de entender el sistema nervioso.
- este interés en el sistema nervioso de remonta a la propia carrera académica de Maturana y a su interacción con EL centro más duro de la nueva ciencia de la cibernética (el BCL)
- El ostracismo que el mundo de la cibernética tenía respecto de Rashevsky. Aunque BCL (1958-1974) co-existió temporalmente con Rashevsky y su grupo... y además sólo lo separaban 200 kms (los dos estaban en el mismo estado, Illinois), seguramente no interactuaron mucho. De hecho, en 1961, para anunciar su creación BCL organizó una conferencia con el ambicioso título de *Principles of Self-Organization* (Zopf y cols., 1962), entre los asistentes está la aristocracia más rancia de la cibernética (Gordon Pask, Stafford Beer, Ludwig von Bertalanffy, Ross Ashby, Warren McCulloch), pero Robert Rosen y Rashevsky, que están tratando de hacer una ciencia de la organización biológica y que viven muy cerca, están –notablemente–

ausentes.<sup>1</sup>

- Conceptualmente están en lados opuestos de la luna. La preocupación primaria de Maturana y Varela es epistemológica y tienen como modelo de pensamiento de libros como “La Lógica de lo Viviente” (de Francois Jacob (1920-2013)). La prueba de esto es el siguiente trabajo en conjunto (Varela y Maturana, 1972). Este es un manuscrito que habla de mecanismos, organización, estructura y explicaciones en biología. Es un trabajo poco conocido pero que muestra hacia donde se dirigían en 1972/1973 Maturana y Varela. Este manuscrito no se puede entender fuera del contexto de la cibernética.
- En 1982, cuando JC Letelier, le presentó a Francisco Varela fotocopias de los papers de Rosen de 1958/1959 Varela estuvo absolutamente sorprendido. Además, a pesar de que F. Varela y JC. Letelier tenían buena formación matemática fueron incapaces de avanzar en el entendimiento de estos trabajos (Letelier, comunicación personal).

## Clausura metabólica

En este punto creo que es prudente sacar al lector de un innecesario y aburrido suspenso e indicar cuales son las ideas con las que me alinee, a modo de que mi crítica a los puntos de vista mencionados anteriormente tenga un contexto adecuado. Mi postura, la cual menciono no por interés político si no porque permea y guía el desarrollo completo de esta tesis, se fundamenta en las siguientes nociones:

- existe un “secreto”<sup>2</sup> de los sistemas vivos

---

<sup>1</sup>Seguramente von-Foester sabía de la existencia de Rashevsky y su grupo, mal que mal este era el editor de la única revista dedicada a la teoría de la biología. Además, McCulloch, en 1956, había escrito una *Letter to the Editor* de *Science* defendiendo a Rashevsky durante la persecución liderada por Joseph McCarthy (1908-1957) en las universidades de EEUU

<sup>2</sup>aquí con secreto me refiero a una cualidad que identifica unívocamente a los seres vivos y de la cual se pueden derivar el resto de las propiedades estos sistemas

- el cual no tiene que ver directamente con ideas sacadas de la física (entropía), la computación (algoritmos especiales) o por propiedades únicas de ciertas moléculas (ADN)
- un paso importante para comprender el secreto es construir una verdadera ciencia de la **organización**
- el secreto consiste en la forma en la cual está organizada la red de reacciones metabólicas que necesariamente subyace en cualquier ser vivo
- para entender el secreto debemos poder entender con precisión las organizaciones
- en particular, y por ahora en términos vagos, el aspecto definitorio, es que el metabolismo se autoproduce. Es decir la enorme mayoría de los componentes del metabolismo (más del 95 %) son generados por el propio metabolismo a través de procesos sustentados por el metabolismo. De aquí en adelante me referiré a esta propiedad como **clausura metabólica**.

Sin lugar a dudas la noción de clausura ha tenido una acogida muy buena en el área teórica, prueba de ello es la continua aceptación de trabajos en las conferencias de *Artificial Life* y *European Conference on Artificial Life* desde que fueron creadas en 1987 (en *Los Alamos, NM*) y 1990 (en *Paris*) respectivamente. Desgraciadamente, las teorías que explotan la noción de la clausura metabólica no se han plasmado en resultados concretos para la biología experimental. Esto ha hecho que esta importante noción sea virtualmente ignorada por los biólogos experimentales. Además, cuando algunos biólogos entusiastas tratan de entender por ellos mismos la teoría encuentran que el campo es una maraña inescrutable de conceptos difícilmente entendibles y más innacesibles aún a la hora de aterrizarlos en manipulaciones experimentalmente relevantes. Uno de los principales hechos que entorpecen la penetración de esta idea en biología es que no se cuenta con una

gran teoría unificada, si no que el estado del arte consiste de un conjunto inconexo de teorías autárticas que muy raramente hacen referencias cruzadas. Si se contase con una visión más uniforme y amplia, podría tenerse una comprensión más profunda del metabolismo y por extensión, de los seres vivos. Es por eso que en esta tesis intentaré avanzar, utilizando la clausura metabólica como eje central, en su clarificación y matematización, teniendo como meta final obtener resultados operativos de esta noción.

Inicialmente puede parecer una tarea titánica poder hablar en general, y con un cierto grado de precisión técnica, de un sistema con clausura cuando este modelo puede realizarse a través de una variedad infinita de configuraciones de conectividad y de posibles cinéticas. No hay que olvidar que el metabolismo de *E. coli* (sin contar las múltiples reacciones de señalización) es una red que contiene del orden de 1200 metabolitos y 1600 reacciones. Buscar un resultado general, que se aplique a todos los metabolismos conocidos –y por conocer– que pueden tener conectividades muy distintas, parece una meta imposible.

Ambiciosamente, y perfectamente consciente del párrafo anterior, trataré de obtener un (1) resultado técnico que a la vez sea general y que tenga *algún* grado de aplicación práctica. Confío en que si se encuentra dicho resultado se habrá abierto un camino de avance (para entender los sistemas metabólicos) alejado de la ya trillada discusión conceptual que tiene ya 40 años; y valiosamente más cercano a la biología experimental.

Mi proyecto de tesis se enmarca en un esfuerzo continuo, de muchas décadas, que trata de entender y generar herramientas para usar esta noción de autoproducción en forma cuantitativa. Si bien es cierto que desde hace 10-15 años existe una escuela de pensamiento que trata de entender el metabolismo desde un punto de vista sistémico; esta escuela no ha logrado cristalizar la idea de clausura, si no que se ha concentrado en generar simulaciones detalladas siguiendo la tradición del *Análisis de Control Metabólico*; además de buscar principios esencialmente empíricos. Uno de los mayores representantes de este modo de

pensar es Bernhard Palsson y su trabajo principal trata de predecir los flujos de todas las reacciones metabólicas, pero en este enfoque se produce una dicotomía muy especial: las enzimas no se consideran en lo absoluto, sino que se les asigna el rol de factores externos constantes. Así, aunque se cuenta con esta visión cuantitativa que ha llegado a ser muy importante para la biotecnología, no utiliza el principio de la clausura metabólica ya que los catalizadores no son considerados como partes de la red.

Así, antes de empezar, me gustaría insistir en que la idea de que la clausura metabólica es lo central ya tiene más de medio siglo. Esto queda en evidencia por la cantidad de teorías que orbitan este concepto. Es más, en este tiempo se han realizado una multitud de análisis y (sobre todo) meta-análisis y por eso mismo creo que el avance ahora depende de lograr un resultado que convenza a todos. Esta tesis es justamente un paso en esa dirección, no es otro sesudo análisis estilo de Razeto-Barry (2012) o Letelier y cols. (2011). Las teorías que se deben conocer para contextualizar correctamente lo que viene en esta tesis son (cronológicamente):

- Biología Relacional (1954)
- Sistemas  $(M, R)$  (1958)
- Teoría de las Redes de Reacciones Químicas (1972)
- Metabolic Control Analysis (1973)
- Autopoiesis (1973)
- Redes Auto catalíticas (1986)
- Teoría de la Organización Química (2001)
- Conjuntos RAF (2004)

Habiendo ya revisado algunas de estas teorías, me dedicaré a continuación a exponer los representantes más técnicos o matemáticos de la corriente de pensamiento que se preocupa por la organización de lo vivo.

## **Dinámica Autopoiética**

Como he dicho anteriormente, en los últimos 40 años se han usado verdaderos ríos de tinta para clarificar, obscurecer, debatir e intentar avanzar sobre la noción de clausura. Personalmente creo que la hora de la discusión puramente conceptual ya pasó y que el verdadero avance se producirá cuando se muestre como esta noción (la clausura) sirve en el mundo del biólogo experimental. Mi apuesta está en hacer una amalgama efectiva entre las ideas más clásicas del estudio del metabolismo como el de uso de ecuaciones diferenciales o el estudio del estado estacionario con nociones quizás más abstractas pero potentes como la clausura metabólica.

Por lo tanto, el camino que se explorará en esta tesis plantea un compromiso entre los modelos ya descritos (Conjuntos RAF, COT, Sistemas (M.R) y Autopoiesis; tratando de destilar lo máximo posible el análisis algebraico) y la incorporación del tiempo y la lógica circular. Una carencia de todos los modelos mencionados anteriormente es la incapacidad de incorporar en forma natural y simultánea la dinámica temporal del sistema y las ideas de clausura. Esta dirección es importante pues el estudio del tiempo y sus nociones asociadas como: flujo a través de una reacción, estados estacionarios, transiciones y estabilidad; nos permitirán acercarnos al mundo experimental. Mi esperanza es que contemplando tanto la dinámica como la clausura, se abrirá una puerta a futuros resultados cada vez más operativos.

Más aun así quiero resaltar que, un resultado técnico en esta interesante intersección entre topología y dinámica, resultaría conveniente para los biólogos del metabolismo quienes están interesados principalmente en “entidades” temporales, particularmente en su

variación bajo distintas condiciones experimentales. Ejemplos de estas entidades o parámetros son: concentración de metabolitos en distintas condiciones, mediciones de flujos y su distribución en la red metabólica, etc... En general muchas de las preocupaciones (experimentales) se pueden resumir en:

*Dada una red de reacciones, qué ocurre con la distribución de flujos si la concentración de un catalizador (enzima) cambia?* Este es el tipo de consideraciones experimentales que debería poder responder una teoría del metabolismo para considerarse útil y este es el tipo de preguntas que pretendo empezar a responder echando mano a formalismos matemáticos.

## **Teorías formales**

En esta sección revisaré someramente las teorías formales atinentes a la clausura metabólica y el metabolismo en general. En los siguientes dos capítulos profundizaré bastante más en los detalles técnicos, pero creo que es importante primero contar con un *sabor* de los distintos formalismos.

## **Conjuntos Auto catalíticos**

Hacia 1980 la idea de clausura metabólica ya se había presentado bajo varios ropajes (Sistemas  $(M, R)$ , Autopoiesis, el Chemoton de Tibor Ganti, los Hiperciclos de Manfred Eigen y Peter Schuster), pero nadie había intentado hacer una teoría que se refiriese a un conjunto específico de reacciones y donde los catalizadores fuesen específicamente producidos por la red. Stuart Kauffman (1939-), un biólogo teórico que ya se había hecho un renombre con la idea de redes de genes (genetic networks) vistos como redes de interruptores de dos estados, atacó frontalmente el problema de la clausura metabólica con su noción de *conjuntos auto catalíticos* (Kauffman, 1986). Es importante estudiar este trabajo

ya que anuncia varias de las técnicas usadas más tarde.

### Autocatálisis a Nivel de una Reacción

Una simple reacción es auto catalítica cuando algún producto es un catalizador de la propia reacción. Esta es una situación poco común en el mundo orgánico pero se han descrito varios casos en el mundo inorgánico. Esta situación es esencialmente inestable ya que la velocidad de consumo de los sustratos crece exponencialmente y por ende termina consumiendo todos los sustratos.



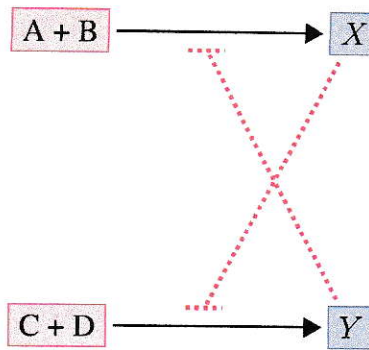
**Figura 3:** La idea de *auto catálisis*. El producto de una reacción (i.e.  $X$ ) cataliza a la misma reacción que la produce (.....). Esta reacción crece exponencialmente en velocidad. Este crecimiento exponencial es un problema esencial de todos los sistemas auto catalíticos.

### Autocatálisis a Nivel de Conjunto de Reacciones

A Kauffman le interesaba una forma más sutil de auto catálisis; la idea de un conjunto de reacciones donde ninguna reacción sea auto catalítica (en el sentido del párrafo anterior) pero sí que todos los catalizadores sean producidos por el sistema. Esta condición de *catálisis cruzada* define a un conjunto auto catalítico y se denomina *catalytic closure*.

Catalytic “closure” must be achieved and maintained. That is, it must be the case that every member of the autocatalytic set has at least one of the possible last steps in its formation catalyzed by some member of the set, and that connected sequences of catalyzed reactions lead from the maintained “food set” to all members of the autocatalytic set. (...) the main purpose of this article is to suggest that achieving catalytic closure can be a high probability event.





**Figura 4:** La idea central de *clausura auto catalítica* o *auto catálisis colectiva*. El producto de una reacción (e.g.  $X$ ) cataliza a alguna otra reacción en la red (.....). Además, esta reacción tiene un catalizador producido por la red en otra reacción. Esta interacción cruzada hace que, aunque ninguna reacción es *per se* auto catalítica, todas las reacciones son catalizadas por moléculas producidas por la propia red

Hasta aquí sólo se trata de definiciones, la idea comienza a ser interesante cuando se supone que existe una probabilidad no nula ( $p$ ) de que el producto de una reacción cualquiera del conjunto de reacciones catalice alguna reacción en la red. Con que probabilidad  $p$  y con qué cantidad de reacciones  $N$  tengo una red auto catalítica sustentable? Increíblemente, con herramientas simples de combinatoria y suponiendo sólo reacciones de condensación entre péptidos logra un resultado interesante. Si se parte de dos monómeros  $A$  y  $B$  (el llamado *Food Set* de la red ya que reflejan los elementos inyectados desde el exterior) y que se generan polímeros de la forma  $AABAB..ABA$  y que cada polímero tiene una probabilidad  $p = 10^{-9}$  de catalizar una reacción de condensación... se necesitan sólo 18000 tipos de polímeros para tener una red con clausura catalítica. Pero esta conclusión no estuvo exenta de críticas ya que las estimaciones de Kauffman no consideraban varios efectos “secundarios” que pueden transformarse en verdaderas barreras. Un ejemplo es el rol de ácido débil que juega el agua y que termina por degradar cadenas proteicas largas (hidrólisis). Otra crítica se refiere al requerimiento de que una nueva molécula tenga la *misma* probabilidad  $p$  de catalizar cualquier reacción. Esta es una suposición muy fuerte y encontró fuertes dudas (Lifson, 1997).

Pero el intento de Kauffman es interesante por los siguientes puntos:

- Define la idea de clausura catalítica a nivel de un conjunto de reacciones
- Obtiene cotas del número de componentes y reacciones para lograr clausura
- No usa la idea de matriz estequiométrica o de listado de reacciones. Sólo argumentos probabilísticos y combinatorios
- No usa ningún razonamiento de cinética química
- El parámetro que caracteriza a la red es la razón entre *aristas* (i.e. reacciones) y *nodos* (i.e. moléculas)
- Cuando ese parámetro alcanza el valor de 0,5 se produce una *transición de fase* y la red deviene catalíticamente clausurada.
- Anuncia todos los análisis que se basan en *leyes de conectividad en grafos* del estilo *pequeño mundo*. No por nada menciona en su bibliografía el trabajo de Paul Erdős (1913-1996) *On Random Graphs*
- Las cotas al número de componentes y reacciones encontradas, siendo altas, son razonables para el escenario en el que se enmarca esta teoría
- Abrió un área de investigación que continua hasta hoy (2013)
- Escribió un libro, que en los 90s tuvo gran impacto; *Origins of Order: Self-Organization and Selection in Evolution* (Kauffman, 1993) que establece que los sistemas auto-organizados tienen una ventaja adaptativa que pre-existe a la selección natural. Esta idea viene directamente del resultado de que las redes con clausura catalítica son inevitables, a pesar de que la probabilidad  $p$  sea muy baja, a medida que aumenta el número de reacciones. Para esta tesis este es un resultado muy importante ya que

nuestro resultado principal se refiere justamente a este problema... pero creo que el resultado es más claro que la derivación de Kauffman.

- Otro aspecto interesante es que Kauffman en su trabajo de 1986 no cita ni a los sistemas  $(M, R)$ , ni a la autopoiesis, pero si cita generosamente a la noción de hiperciclos. Esto no es de sorprender ya que Kauffman ve a sus conjuntos auto catalíticos como la justificación teórica de la estabilidad de los hiperciclos de Eigen y Schuster (Eigen y Schuster, 1978). La dinamica de citas entre las distintas teorías y autores es un tema muy interesante en este campo.

La manera de aproximarse a la clausura metabólica de Kauffman (clausura catalítica), respecto de los enfoques de la Autopoiesis y sistemas  $(M, R)$  puede parecer artificial (asignar al “voleo” una probabilidad  $p$  de capacidad catalítica al producto de cualquiera reacción) pero produce un primer resultado cuantitativo al estimar el “tamaño” que debe tener una red auto catalítica<sup>1</sup>. Pero ahora aparece un nuevo problema: ¿Confrontado con una red específica de reacciones, como se determina si tiene clausura metabólica? Las teorías siguientes retoman este punto.

## Conjuntos RAF

En 1998 un matemático neozelandés, Mike Steel, decidió visitar de nuevo el teorema de Kauffman y ver si se podrían hacer mejoras o clarificaciones. Pero el espíritu matemático de Steel, tal vez sin quererlo, empezó un nuevo camino al poner rigor en la *definición* de la red de reacciones. Su modo de análisis es usar, bien!, la teoría de conjuntos para realizar una descripción de la conectividad de una red de reacciones. Como resultado de este ejercicio emergen los conjuntos RAF, un formalismo que si bien aún vive en las sombras, cada vez gana más potencia. En el capítulo II, se dará una revisión técnica a esta

---

<sup>1</sup>¿El primer teorema de la nueva biología?

teoría.

Los conjuntos RAF son interesantes ya que una de las motivaciones iniciales fue generar algoritmos que pudiesen investigar si un conjunto de reacciones químicas era auto catalítico. Dicho de otra manera: ¿Cual es el método para saber si me enfrento a una red donde todas las reacciones son catalizadas por moléculas provenientes de la misma red. Este es un problema de búsqueda en un árbol de *items* en una base de datos. De hecho el primer trabajo de Steel y cols. es justamente la descripción de este método de búsqueda. Este problema, que corresponde a una búsqueda exhaustiva, es un problema que demanda mucho poder de computo y por eso es interesante buscar algoritmos eficientes. Entonces, en realidad tanto la búsqueda necesaria para saber si un conjunto auto catalítico es RAF o para saber si un conjunto de reacciones es una *organización* es una manipulación especial en una base de datos del metabolismo.

## **Teoría de la Organización Química**

Esta teoría, llamada en inglés *Chemical Organization Theory* (COT), fundada por Peter Dittrich y colaboradores (2007) (Dittrich y di Fenizio, 2007; Centler y cols., 2010), resulta interesante pues representa el metabolismo como una serie de reacciones químicas, a la RAF, pero incorporando la estequiometría. Aunque su manera de enfocar el problema de la clausura tiene muchas similitudes con conjuntos RAF, estos autores se reconocen herederos de la tradición de *Artificial Chemistry*, es decir, lidian con sistemas formales de reacciones químicas introducidos por investigadores del área de la Vida Artificial como Walter Fontana a comienzos de los 90s. En el capítulo siguiente, una vez tenga el formalismo de los conjuntos RAF será muy fácil introducir los detalles de esta teoría.

Hay que subrayar que la construcción de los sistemas RAF y COTs proveen algo esencial para esta tesis: operacionalizan la noción de clausura metabólica. Como se ha mostrado en el capítulo I, en ambas teorías, existen nociones de clausura que no son idénticas,

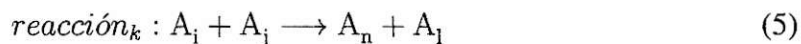
pero si similares. Lo interesante es que la clausura se define como una cadena particular de flechas y nodos, y esto nos lleva nuevamente a considerar que lo realmente importante para estudiar el metabolismo real es manipular las bases reales de datos genéticos, y del metabolismo.

Un punto interesante respecto de esta tesis, es que no diseñaré métodos para saber si una red metabólica cumple o no una cierta cualidad. Esta investigación usa un argumento totalmente distinto, en vez de buscar qué propiedades topológicas tienen nuestras redes, les exigiré una propiedad topológica y veré que propiedades cinéticas se derivan de esta conectividad.

## **Lo que se debe saber sobre la Teoría de las Redes de Reacciones Químicas**

En biología se usa hablar de *criptozoología* para referirse a aquellas especies semi-míticas (*Okapi*, *Calamar Gigante*) que son difíciles de ver pero que son reales. En la teoría de las redes metabólicas existe una *criptoteoría* que, aunque tiene ya 40 años ha estado tan oculta como *Homo floriensis*: la *Chemical Reaction Network Theory* (CRNT) o Teoría de las Redes de Reacciones Químicas en español.

La CRNT ataca un problema fundamental de la Ingeniería Química, ¿Que pasa si en un recipiente cerrado se tienen  $(A_1, A_2, A_3, \dots, A_p)$  especies químicas (moléculas) que interaccionan de acuerdo a  $q$  reacciones del tipo:



Además, para hacer todo más fácil, se supone que todas las reacciones siguen la ley de

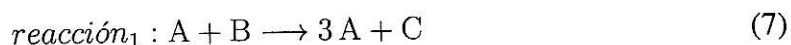
acción de masas. Así la reacción definida arriba tiene una velocidad

$$v_k = K \cdot A_i(t) \cdot A_j(t) \quad (6)$$

Obviamente si se tienen  $q$  reacciones acopladas resulta un sistema de ecuaciones diferenciales (no-lineales) acopladas que no tienen solución analítica. Aparentemente es imposible continuar más allá de este formalismo básico. Pero hacia 1970 Martin Feinberg (1941-) elaboró la CRNT para tratar de predecir cualitativamente como variarían las concentraciones de los distintos  $A_i$ . Claro que él, para poder avanzar, hace algunas restricciones. Para entender porque se debe, al me saber de la existencia de la CRNT, revisaremos su teorema más famoso, y hay que dejar en claro que CRNT no es una teoría para entender los secretos de la vida, es una teoría para predecir concentraciones (Feinberg, 1987).

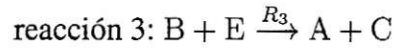
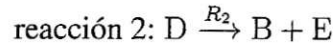
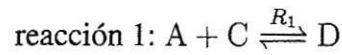
### **The Zero Deficiency Theorem, ZDT**

La idea intuitiva para entender el ZDT es que se tiene un reactor, cerrado, perfectamente homogeneizado y en el tiempo cero se introducen  $p$  especies (moléculas) ( $A_1, A_2, \dots, A_p$ ). Entre estas especies ocurren  $q$  reacciones según reacciones como:

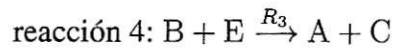
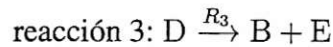
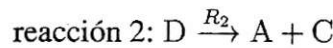
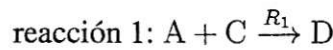


Desde el punto de la vista de la química en esta reacción participan 3 especies ( $A, B, C$ ) y dos complejos ( $C_1 = (A+B); C_2 = (3A+C)$ ). Toda la CRNT se centra en los complejos, no en las especies individuales. Por lo tanto, un conjunto interactuante de moléculas se ve como un conjunto interactuante de complejos y las reacciones son entre complejos, no entre moléculas. Así se define una red de reacciones química (*chemical reaction network*) como un grafo  $G = (C, R)$ , donde los vértices ( $C$ ) son los complejos y las aristas (que son orientadas en este caso)  $R$ .

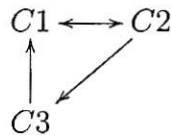
Veamos un ejemplo concreto:



Aquí el conjunto de especies  $S = A, B, C, D$  tiene 5 elementos, mientras que el conjunto de complejos  $C = C1 = (A + C), C2 = (D), C3 = (B + E)$  tiene tres elementos. Pero no hay tres reacciones, como pareciera sugerir el listado de más arriba, ya que la “reacción 1” se considera como dos reacciones separadas; por lo tanto hay 4 reacciones. La red finalmente posee 4 reacciones:



Visto de esta manera  $G = (C, R)$  es el siguiente grafo <sup>1</sup>



Para analizar el devenir temporal de las concentraciones de los 5 metabolitos se puede escribir una ecuación diferencial

$$\frac{\partial x}{\partial t} = f(x) \quad (8)$$

<sup>1</sup>Como se puede ver aquí aparecen algunas ideas de Rosen de 1958, pero de una manera más precisa

que, en caso de estas reacciones se puede escribir en términos de las matrices estequiométricas.

En todo caso lo realmente relevante es que este sistema de ecuaciones diferenciales no se puede resolver en forma analítica ya que es no-lineal. En su desarrollo aparecen términos del tipo  $A(t)B(t)$  y similares. Pero, y aquí debe estar el genio de Feinberg, hay varias cosas que se saben; por ejemplo, ninguna especie puede tener concentraciones negativas. Si el sistema tiene soluciones en estado estacionario, todos los componentes son mayores que cero. Además, los cambios están ligados por los coeficientes estequiométricos (cada vez que desaparece un  $D$  aparecen  $1A, 1C, 1B, 1E$ ). Esto hace que la matriz estequiométrica  $S$  que representa al sistema

$$S = \begin{pmatrix} -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ -1 & 1 & 0 & 1 \\ 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

tenga algunas dependencias lineales entre sus filas. Además se define el número de clases de ligamiento (*linkage classes*) como el número de subgrafos conexos que existen en la red (en nuestro ejemplo sólo hay un grupo de ligamiento). Con estos conceptos Feinberg define la *deficiencia* de una red como:

$$\delta = |C| - l - s = -2 \quad (9)$$

$3 - 1 - 4$

donde:  $l$  = número de clases de ligamiento,  $|C|$  = cardinalidad del número de complejos,  $s$  = dimensión del espacio de columnas de matriz  $S$ .  $\delta$  = deficiencia de la red



## Definición de clausura en CRNT

y define que una red es **débilmente reversible** si, para cada reacción  $r_{(i,j)}$  entre los complejos  $C_i$  y  $C_j$  existe un camino de reacciones que partiendo en  $C_j$  (es decir el complejo de término de la reacción  $r_{(i,j)}$ ) termina en el complejo de partida (i.e.  $C_i$ ). Con todos estos conceptos Feinberg (1974) nos regala ~~con~~ un teorema sorprendente:

Sea una red química con conjunto de especies  $S$ , reacciones  $R$  y complejos  $C$ . Además supongamos que la red sea débilmente reversible y que su deficiencia sea igual a cero ( $\delta = 0$ ) entonces **para cualquier conjunto de valores  $k_i$  de las reacciones** siempre existe un punto de equilibrio estable de concentraciones de las especies en  $S$ .

Este teorema es casi increíble ya que toma un problema de ODE (no lineales) y obtiene un resultado generico (existencia de un punto estable) sin importar el valor de las constantes cinéticas. Para esta tesis es particularmente importante ya que:

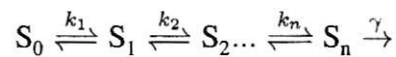
- usa grafos para representar reacciones químicas acopladas
- usa nociones de clausura
- no depende de los detalles de como es la conectividad, sólo que se respete la reversibilidad débil.
- demuestra la existencia de un estado estable ante cualquier valor de parámetros externos
- usa argumentos geométricos (y del álgebra lineal) para obtener un resultado sobre sistemas dinámicos.
- desde un punto de vista histórico es interesante, ya que mezcla ideas de los sistemas  $(M, R)$  (uso de grafos), con COT (ambas teorías tratan de obtener resultados sobre estados estacionarios)... y fue una teoría largamente ignorada entre 1970 y 2000.

- Desde hace 10 años ha existido una verdadera explosión de trabajos desarrollando CRNT y con muchas aplicaciones a la biología de sistemas.

Hay que insistir que lo que realmente se desearía poder generar con todas estas teorías de clausura es poder generar un teorema equivalente al ZDF.

## Análisis del Control Metabólico

El análisis del control metabólico es otra teoría que se gestó de forma contemporánea a la autopoiesis y, aunque no trata de encontrar “el secreto” de los seres vivos, es relevante para esta tesis. Fue escrita independientemente por H. Kacser y J. Burns (Kacser y Burns, 1973), y R. Heinrich y T. Rapoport (Heinrich y Rapoport, 1974). Estos autores modelaron matemáticamente una secuencia de reacciones, es decir, un metabolismo muy simple como el que se muestra a continuación:



Suponiendo cinética de acción de masas e introduciendo la noción de coeficiente de control, fueron capaces de derivar múltiples teoremas cuyos resultados incluso hoy no son incluidos en la formación de biólogos. Se puede agrupar los teoremas en dos ramas:

- Los teoremas de la sumación, que indican que los flujos metabólicos son propiedades sistémicas y por lo tanto su control es compartido por todas las reacciones del sistema; y
- Los teoremas de conectividad, que establecen una conexión entre las propiedades cinéticas de las reacciones individuales y las propiedades sistémicas de la vía metabólica.

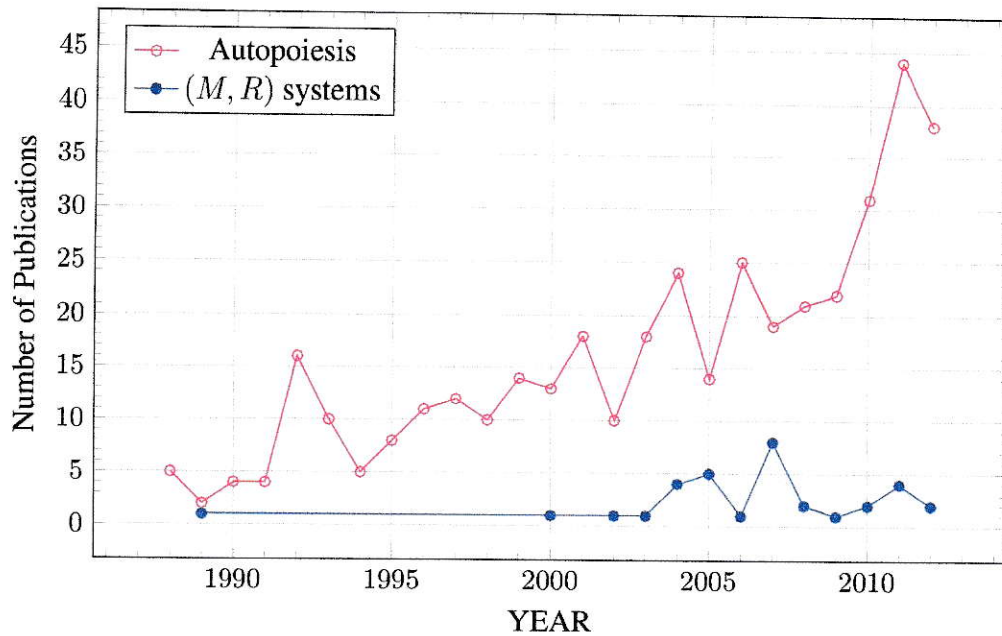
Este análisis puede ser interpretado como una expresión matemática de la clausura, pues indica que el sistema debe ser estudiado como tal en su totalidad por un lado, y

por otro que el origen de todo cambio es debido a la dinámica interna del sistema. Por desgracia existe un elemento faltante en la constelación de ideas manejadas dentro de esta teoría, y es incorporar las reacciones que generan a los catalizadores dentro del análisis. Por esto la clausura no es totalmente captada bajo este formalismo.

Si bien esta teoría se acerca tangencialmente al tema de la clausura, es de gran valor pues al día de hoy es la única que ha tenido algún grado de éxito en la biología experimental. Por ello su formalismo, basado en matrices y sistemas de ecuaciones diferenciales podría contener perspicacias que iluminen el camino que se debe seguir en el campo más ambicioso de la biología teórica.

## **Impacto relativo de estas teorías**

Para cerrar este capítulo es útil hacer un análisis sobre la importancia (e impacto) relativa que estas teorías han tenido sobre el entendimiento teórico del metabolismo de los sistemas vivos. No es nuestra intención de hacer un estudio acabado de las citas que han generado, pero si es interesante tener algún nivel de información sobre cual ha sido el impacto real de todas estas teorías en la biología de las últimas 5 décadas. Es notable cuan fértil fue el periodo alrededor de 1970. Entre 1970 y 1975 vieron la luz: CRNT, Autopoiesis, MCA, Hiperciclos y el Chemotón. El impacto que han tenido estas teorías ha sido muy desigual y también ha sido variable en el tiempo. Por ejemplo CRNT, creada hacia 1974, permaneció por largos años como un conocimiento esotérico de un puñado de ingenieros químicos y sólo en 2000 comenzó a popularizarse y hacia 2005 comenzó a ser mencionado en los círculos de Systems biology. Un similar periodo de invisibilidad tuvo MCA en su primera década de existencia para después convertirse en el fundamento del *Metabolic Flux Balance* (MFB) que es uno de los pilares del análisis cuantitativo del metabolismo propugnado por *Systems Biology*. La figura 5 es iluminadora: Hasta 2008 se publicaba



**Figura 5:** Evolución temporal de trabajos, en el sistema ISI, centrados en Autopoiesis o Sistemas  $(M, R)$ . Fuente: análisis propio del sitio Web of Science

muy poco en sistemas  $(M, R)$ , yo estimo que el cambio de dinámica que aparece el 2007 se debe a trabajos (del 2003 y 2006) que relacionaban ambas teorías. Este gráfico muestra como estas ideas comienzan a ser aceptadas. No hay que olvidar que en la “web” el nivel de tráfico relacionado con ambos conceptos es muy grande, con varios sitios dedicados a su análisis y difusión.

## El Juego de las Citas

Como siempre es interesante analizar cuáles son las referencias usadas en el primer trabajo que introduce los conjuntos RAF. Obviamente se cita el trabajo de 1986 de Kauffman... pero ahora desaparecieron las citas a los hiperciclos. Siguiendo con la aparente tradición no se hace ninguna referencia a los sistemas  $(M, R)$  o a la Autopoiesis. En las referencias usadas en el primer trabajo que introduce COT, sorprendentemente no se cita el trabajo de 1986 de Kauffman... pero se siguen citando a los hiperciclos. Nuevamente aquí no se

hace ninguna referencia a los sistemas  $(M, R)$  o a la Autopoiesis, pero tampoco a los conjuntos RAF. Esta dinámica de citas muestra la insularidad (¿intrínseca?) de este campo y resulta hasta inexplicable a sabiendas de que Peter Ditttrich está muy interesado y ha reflexionado largamente sobre los sistemas autopoieticos.

Por razones que no se pueden explicar aquí fácilmente la Autopoiesis ha generado un interés en una variedad inimaginable de campos. Esto puede reflejar la necesidad que tiene la ciencia moderna de tener una verdadera “teoría general de sistemas” con resultados que se parezcan más a la termodinámica que a la Teoría General de Sistemas de Ludwig von Bertalanffy (1901-1972). Los razonamientos “sistémicos” de la autopoiesis, por eso, han cautivados a muchos. Pero aún falta un primer resultado técnico. Como he dicho, creo que un paso importante sería derivar alguna conclusión sobre la estabilidad o robustez de los sistemas biológicos y esta búsqueda ha motivado esta tesis. He presentado un resumen de las distintas visiones y de este resumen he rescatado los resultados que apuntan hacia nociones de estabilidad. Ahora, al final de este largo –pero necesario– capítulo es bueno recapitular las ideas principales sobre estabilidad:

- La *Biología Relacional* de Rashevsky no dice nada sobre estabilidad.
- Los sistemas  $(M, R)$ , dado que su existencia es tan inusual (dependen de propiedades más bien misteriosas de un operador aún no entendido  $(\beta)$ ) son intrínsecamente estables. Pero esta estabilidad es una estabilidad algebraica (debido a  $\beta$ ) y no en el tiempo... una variable o parámetro que no aparece en la teoría. Los sistemas  $(M, R)$ , en su versión más clásica, son atemporales.
- La formulación primera de *De máquinas y seres vivos* postula, como hechos evidentes, que los sistemas autopoieticos son autónomos, estables y pueden sufrir el proceso de evolución orgánica de acuerdo a pautas precisas donde cambie la estructura y no la organización. Esta supuesta “estabilidad” intrínseca ha sido una de las

razones por la cual la autopoiesis ha sido usada por teóricos en el campo del derecho como justificación de que los sistemas sociales son obligatoriamente conservadores; ver *Autopoietic Law: A New Approach To Law And Society* (Teubner, 1988)

- Los conjuntos auto catalíticos de Kauffman anuncian un teorema de estabilidad ya que afirman que *siempre* es posible encontrar redes estables si se espera lo suficiente para que sean lo suficientemente grandes en número de reacciones. Las contribuciones de los conjuntos RAF son, desde el punto de vista de la estabilidad, son mejoras en definir el tamaño de la red y la probabilidad a la cual auto-aparecen redes con clausura.
- Una organización, *a la COT*, es un metabolismo estable! Claro que la demostración no deja contento a muchos ya que se supone que debe existir un punto de equilibrio y luego se demuestra que es estable
- Sin lugar a dudas que el mejor teorema proviene de la CRNT, que ni siquiera se puso como objetivo (inicial) entender a los sistemas vivos. EL *ZDT* es genuinamente sorprendente al demostrar que siempre existe, para un subtipo muy especial de sistemas químicos con clausura (los que tienen *deficiencia zero*, una solución estable. Después de permanecer oculto del interés general el *ZDT* ha invadido la biología de sistemas y nuevos teoremas y refinamiento aparecen continuamente. Por ejemplo, a comienzos del 2013 se mostró como los estados estables anunciados por el *ZDT* tienen una estructura, técnicamente se dice que forman una variedad (Brogio-li, 2013).
- Los teoremas de MCA, otra teoría que no quiso descubrir “el secreto de la vida”, son esenciales ya que se refieren a propiedades sistémicas de cómo una red opera en el organismo.

# Antecedentes topológicos o relacionales del metabolismo

El análisis matemático del metabolismo parte por encontrar representaciones formales de nuestros objetos de estudio (metabolitos, reacciones químicas, redes de reacciones y flujos). Esto, que parece fácil, es una tarea repleta de problemas que se ha iniciado múltiples veces y a pesar de grandes esfuerzos en la práctica aún no existe un consenso de cómo hacerla en forma efectiva. En este capítulo compararé las distintas formalizaciones que se han hecho del metabolismo a nivel topológico, es decir, en lo que se refiere a cómo se conectan los distintos elementos que componen el metabolismo, sean estos metabolitos relacionados por reacciones químicas o conjuntos anónimos de mayor abstracción. Este enfoque se conoce también como análisis estático (algunos hasta lo llaman Biología Relacional), ya que no considera la dinámica temporal asociada al sistema. Además, caracterizaré algunos sistemas (e.g., RAF, RAO, Organizaciones) que podrían resultar de interés para un análisis dinámico, así como revisaré teoremas que pueden guiar nuestra búsqueda en el entendimiento de las propiedades de sistemas con clausura.

## Reacciones y redes de reacción

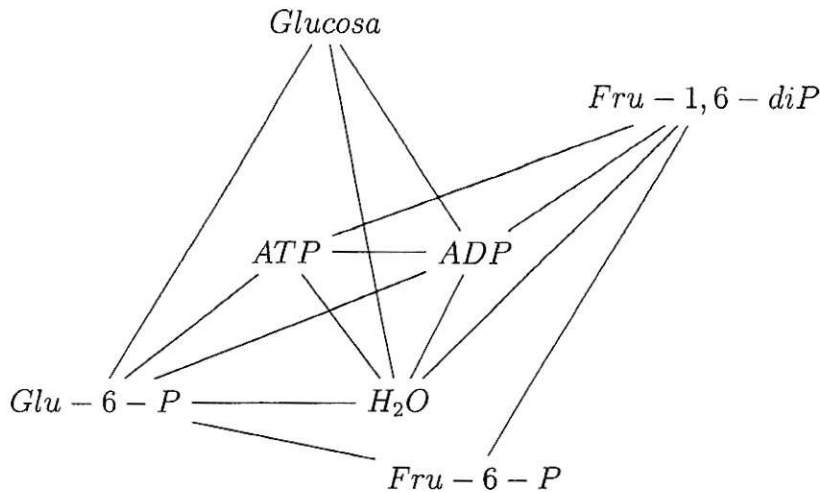
La primera tarea es *delimitar* el problema de la representación del metabolismo. Desde hace más de 100 años estamos todos de acuerdo que la naturaleza de los seres vivos pasa inicialmente por conjuntos de reacciones químicas. Pero hay elementos no obvios en este nivel de descripción, a continuación doy una lista de elementos/componentes que se deben considerar para hacer una adecuada descripción de un metabolismo.

- Reacciones Químicas
- Catalizadores
- Grafos de reacciones
- Niveles de conectividad
- Limitantes termodinámicas
- Flujos
- Elasticidades y datos cinéticos
- Genes y Circuitos de control
- Datos empíricos

La figura 6 muestra la versión dual de como presentar el metabolismo presentado en la figura 7. Una representación está centrada en las enzimas y otras en las moléculas o en las reacciones. Pero en realidad, para el estudio del metabolismo *in toto* o *in silico* es irrelevante que representación se escoja. En efecto, puede ser que para estudiar pequeñas vías o circuitos, una representación como la figura 6 sea superior a otra como la figura 7. Pero para estudiar el metabolismo completo (o casi completo) lo que en realidad se necesita son representaciones algorítmicas (computacionales). Hay dos tipos de representaciones generales:

Representación bio-informática. Aquí se usa una multitud de distintas jerarquías para tener árboles o cadenas de conceptos. Estas cadenas después se usan para hacer computaciones que son como contabilidades muy precisas que sirven para detectar ciclos. El ejemplo más extremo de este enfoque es la base de datos EcoCyc ([www.ecocyc.org](http://www.ecocyc.org)), escrita en LISP, y en la cual se pueden hacer preguntas tales como: “Extraiga todas las





**Figura 6:** Este grafo representa las primeras tres reacciones de la glicólisis (figura 7). Los nodos son las moléculas y las aristas son las reacciones. Comparando ambas figuras uno entiende que, para analizar metabolismos completos, se requiere necesariamente de tener una visión informatizada de toda la red de reacciones.

reacciones que usan oxalato como sustrato y cuyas enzimas tengan más de 400 aminoácidos y ordénelas por la posición del gen en el cromosoma bacteriano”. En la práctica este es el único enfoque que permite trabajar de forma cuantitativa.

Representación por Matriz Estequiométrica. Una contribución importante de MCA fue el deseo de entender los metabolismos reales. A consecuencia de ello, y usando ideas que ya existían en la química, se usó el formalismo de las matrices para representar **todo** el metabolismo. La noción fundamental es la *matriz estequiométrica*  $S$  que está a la base de la investigación actual de ingeniería metabólica o de la biología de sistemas. Con la matriz  $S$ , sumado a técnicas comunes del álgebra lineal y de la optimización se han hecho interesantes contribuciones.

- Búsqueda de estados estacionarios, del metabolismo, como los vectores de velocidades ( $v$ ) que hacen  $Sv = 0$
- Uso de técnicas como FBA (*Flux Balance Analysis*) para restringir que valores puede tomar los vectores  $v$  buscados.

- Uso de técnicas de *Optimización Matemática* para restringir  $v$  a valores cada vez más plausibles.
- Técnicas de medición de flujos, *in vivo*, usando  $C^{13}$ .
- Generación de muchas técnicas donde estas cuatro ideas principales confluyen y que, al día de hoy, constituyen el núcleo conceptual de la ingeniería metabólica; FVA (Flux Variability Analysis), MOMA (Minimization Of Metabolic Adjustments, EFA (Elementary Flux Analysis), OMNI (Optimal Metabolic Network Identification) entre otros.

Durante las dos últimas décadas ha existido una proliferación de trabajos buscando construir una teoría y una tecnología del metabolismo, hasta existe una revista, *Metabolic Engineering* dedicada a la práctica y a la teoría del entendimiento metabólico. Lo que hay que rescatar de esta intensa búsqueda de métodos es:

- Debido a la explosión de datos es ahora posible construir bases de datos exhaustivas de reacciones, metabolitos y genes.
- Se ha desarrollado un pensamiento algorítmico que permite construir grafos de reacciones
- Se han usado las herramientas del álgebra lineal para definir de manera precisa qué se entiende por *estado estacionario* y cuales son sus propiedades relevantes
- Usando conjuntamente las nociones de estados estacionarios e ideas de optimización se puede deducir como variarán algunos flujos bajo nuevas condiciones experimentales
- Además se ha comenzado a medir *in vivo* las concentraciones de varias decenas de flujos en paralelo.

- *Ninguna de estas aproximaciones usa las ideas de clausura de Rosen, Maturana y Varela, Steel o Dittrich.*

Un uso interesante de toda esta actividad ha sido la acción del grupo de Albert-Barabási que ha descrito las *redes de escala libre, o de pequeño mundo* (Reka y Barabási, 2002), las cuales son grafos donde la conectividad promedio de los nodos sigue la regla: los nodos más conectados ganan más conexiones que los nodos menos conectados. Esta idea, cuando se mostró en 2002, pareció augurar el descubrimiento de nuevas leyes de la conectividad entre reacciones y metabolitos. Pero estas *leyes de conectividad* no han dado muchos frutos en biología más allá de servir para darnos cuenta que el metabolismo es un grafo con leyes especiales de conectividad (y sugerir relaciones entre estas leyes y propiedades de los seres vivos) (Wagner y Fell, 2001). Incluso estos resultados han sido revisados bajo luces distintas y se han encontrado contradicciones (Arita, 2004).

El enfoque que tomaré es el de rescatar los puntos valiosos de las distintas teorías, por lo que a continuación haré una revisión de los formalismos que más llamaron mi atención desde el punto de vista de la descripción estática de los sistemas.

## **Rosen in a (long and hard) nutshell**

Rosen tuvo una intuición brillante, nada menos que en su tesis de doctorado, sobre la clausura metabólica. Dado que sus ideas estaban adelantadas en décadas no es extraño que no se haya diseminado su contenido. Tampoco nos debe extrañar que toda la vida de Rosen, desde su trayectoria como profesor, sus trabajos, sus libros, su temprana jubilación (a los 57 años) y su muerte por diabetes haya estado imbricada con esta intuición que, de alguna manera lo terminó cegando<sup>1</sup> a las limitaciones de su obra. Rosen, como otros antes y después, llegó a la conclusión que la clausura metabólica era el aspecto esencial de los

---

<sup>1</sup>Icaro quiso acercarse al sol y esto lo hizo sucumbir

seres vivos pero su formulación es tan extraña, tan atípica, que no se puede resumir en un solo párrafo (se requieren varias páginas). Para los interesados es necesario indicar que las ideas de Rosen tienen cuatro etapas:

- (1958-1959) enunciado de la intuición fundamental en el contexto de la teoría de grafos y posteriormente de la teoría de categorías. Serie de cuatro papers que en lo esencial son la tesis de doctorado.
- (1960-1967) tentativa de clarificación de la intuición usando el lenguaje de la teoría de autómatas (es decir las máquinas de Turing <sup>1</sup>)
- (1972-1974) reformulación de la intuición en términos algebraicos, sin usar teoría de categorías o de autómatas
- (1991) Síntesis final: la intuición inicial de 1957 como núcleo de una nueva epistemología<sup>2</sup>.

Como se puede ver, entender a Robert Rosen implica necesariamente entender su intuición, su misteriosa perspicacia, contenida en su tesis de doctorado hacia el año 1957. Por lo tanto, haré una exposición un poco extensa de ésta intuición, tal como fue inicialmente formulada. Para el desarrollo de esta tesis no es un punto increíblemente importante, pero la incluyo tanto por completitud como ejemplo de lo obscura, complicada o derechamente inabarcable que puede terminar siendo un formalismo<sup>3</sup>. Además creo honestamente que es bueno que exista una exposición completa en un documento de esta facultad ya que servirá de base para futuros estudiantes que, al igual que yo, se interesen por la teoría de la biología.

---

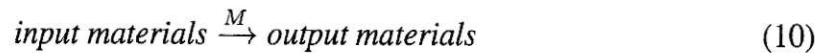
<sup>1</sup>En honor a Alan Turing

<sup>2</sup>Curiosamente no es la única teoría de la organización de lo vivo que decanta en una nueva epistemología.

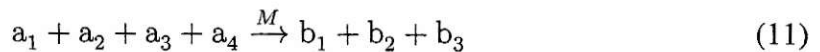
<sup>3</sup>Esta teoría es un caso curioso donde la abstracción juega en contra de la simpleza.

## Metabolismos, grafos y la primera instancia de sistemas $(M,R)$

En 1957 Rosen se adelantó 20 años al considerar que el metabolismo podía ser interpretado como grafos (Rosen, 1958a). La primera novedad fue considerar a las enzimas (o reacciones) como *operadores*  $M$  que tomaban un cierto número de *input materials* para producir otro número (no necesariamente igual) de *output materials*. Estos operadores, que en el hablar normal de la bioquímica se deberían llamar *catalizadores* son el centro de la actividad metabólica del organismo para Rosen (de ahí viene su símbolo  $M$ ).



Con más precisión, en el caso de una reacción que transforma 4 sustratos en 3 productos se puede escribir:



En este punto la exposición se hace pesada y difícil de seguir, aún para estándares modernos, ya que trata de poner la acción de  $M$  como una función con un dominio y recorrido definidos según:

$$M : (A_1 \times A_2 \times A_3 \times A_4) \rightarrow (B_1 \times B_2 \times B_3) \quad (12)$$

Esta manera de ver el metabolismo, donde se mezclan ideas de *grafos* con ideas de *funciones*, *dominios*, *recorridos*, *productos cartesianos* ha sido una fuente de confusión sin par para todos los lectores por los últimos 55 años y es fundamental que el lector de Rosen esté al tanto que lo que el usa (y defiende) como transparente (es decir que una enzima tiene un *Dominio* y un *Recorrido*) algo no tan trivial de aceptar. La razón por la cual Rosen usa este formalismo particular sólo se revela en su trabajo de 1959. Pero se debe ser muy fino en diferenciar entre  $A_1$  (las moléculas de la clase  $A_1$  y una instancia en particular ( $a_1$ ) de esa

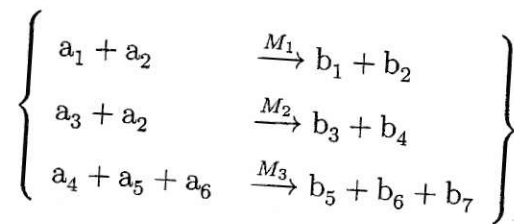
clase. Este punto (distinguir entre conjuntos y elementos) será bastante relevante cuando se analice la re-formulación de 1972, que es la exposición que ha sido más difundida de las ideas de Rosen.

Esta hipergeneralización del metabolismo, cuando se piensa en el operar de las enzimas, es un tanto artificial por dos razones:

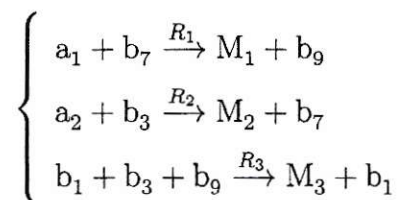
- No da espacio ni dice como describir la especificidad enzimática
- La inexistencia, en el metabolismo donde está actuando la enzima, de los posibles sustratos alternativos.

Rosen prefiere mantener este nivel de abstracción para su modelo y es por esto que ha costado aterrizarlo incluso para estar seguros de que una causa eficiente se refiere efectivamente a una enzima.

Una primera observación crucial de Rosen es indicar que los operadores  $M$ , al ser moléculas, son susceptibles de ser degradadas o destruidas por mecanismos internos (i.e. proteasas) o externos (interacción con metales pesados por ejemplo). Por lo tanto asociado a cada operador  $M$  debe existir otro operador  $R$  que repara (repair) al operador  $M$ , de ahí el símbolo  $R$  y el nombre de la teoría:  $((M,R)$  systems. Es importante clarificar estas ideas, y ponerlas en un lenguaje moderno. Consideremos un sistema  $(M,R)$  muy simple de 3 componentes  $M$  ( $M_1, M_2, M_3$ ). Veamos los componentes  $M$



Veamos los componentes  $R$



Como se puede ver todos los  $M$  son producidos por algún  $R$ . Este es un sistema muy simple, pero ya se puede ver como se construyen las relaciones de clausura. Hay que destacar como una reacción del sistema ( $a_2 + b_3 \xrightarrow{R_2} M_2 + b_7$ ) tiene reacciones que la anteceden ( $a_4 + a_5 + a_6 \xrightarrow{M_3} b_5 + b_6 + b_7$ ). La reacción  $R_3$  es aún más interesante ya que tiene 3 reacciones que la anteceden ( $M_1, M_2, R_1$ ).

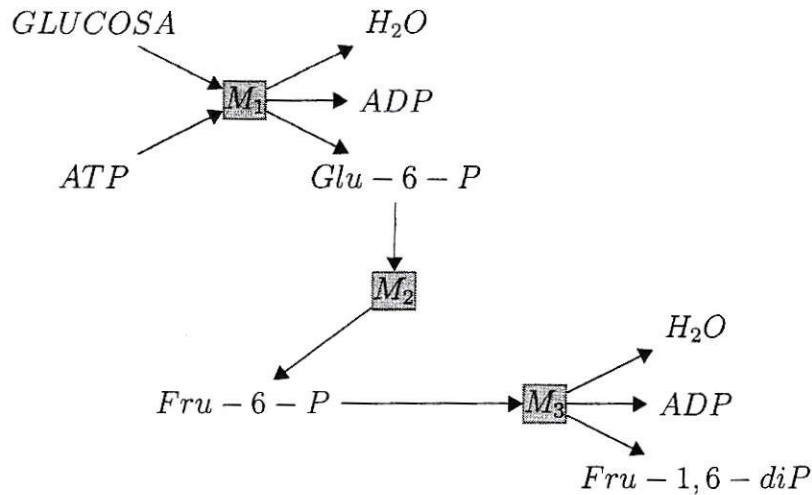
## Reparación vs. Reemplazo

Rosen llamó a sus sistemas teóricos (*Metabolism-Repair systems*), el uso de la palabra “repair” es desafortunado ya que evoca una función que no ocurre. En el metabolismo, las enzimas son degradadas, pero nunca son “reparadas”. Lo que ocurre es que existen procesos (esencialmente la biosíntesis de proteínas) que “re-emplazan” a las enzimas que son destruidas por el propio metabolismo. Por eso, y en el contexto que esta tesis fue hecha en el laboratorio del Dr. Letelier se prefiere denominar la función de los operadores  $R$  como de “reemplazo” y así los  $(M,R)$  systems son metabolism-replacement systems.<sup>1</sup>

Claro que los operadores  $R$  son distintos, no son simples moléculas, sino que son de por sí conjuntos de reacciones bioquímicas que permiten reemplazar a los operadores  $M$  que se van inactivando. Así una secuencia de reacciones metabólicas (Figura 7) se puede escribir como un grafo como lo muestra el siguiente ejemplo.

---

<sup>1</sup>Este cambio de vocabulario no ha sido aceptado por algunos ex-alumnos de Rosen como Aloisius Louie



**Figura 7:** Conjunto de reacciones interpretado como un grafo simple. Cada reacción está representada por un operador  $M_i$  que toma materiales de entrada y produce materiales de salida.

### Un primer teorema de los sistemas $(M,R)$

En el primer trabajo donde se introducen los sistemas  $(M, R)$  Rosen hace una demostración de que en estos sistemas siempre debe existir un componente (algún  $M_j$ ) que de ser inhibido (es decir que si a un tiempo dado todas las moléculas de tipo  $M_j$  son extraídas del sistema) esto implica el colapso de todo el sistema ya que desaparecen todos los  $M$ . Esto se debe, justamente a la estructura de inter-relaciones entre todas las reacciones tipo  $M$  y tipo  $R$ . En este teorema Rosen ya anuncia modos de pensamiento que aparecerán en teorías posteriores. Pero Rosen no siguió esta manera de pensar, su trabajo se centró en perseguir la intuición que caracterizó todo su obra.

### $R$ reemplaza a los $M$ , qué mecanismo reemplaza a los $R$ ?

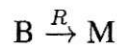
La intuición de Rosen se refiere al proceso que mantiene los  $R$  funcionando. Rosen llegó a la conclusión de que la clave para entender a los sistemas vivos residía en entender cómo es que los  $R$  eran producidos. Volvamos a introducir las ideas principales; el metabolismo



Roseneano es simple, su modelo es:



Pero el catalizador  $M$  es una molécula que se degrada, luego quién asegura que exista suficiente  $M$ ? Rosen postuló que para cada  $M$  existía un proceso que también era catalizado por algún  $R$  (llamado el reparador de  $M$ :  $R_M$ ) que “reparaba” a  $M$  según:<sup>1</sup>



Pero en este esquema aparece una pregunta natural: Quien repara a  $R_M$ ? Obviamente, invocar a un *reparador del reparador* no es una solución muy elegante. Fue justamente esta regresión infinita (*fuite en avant* o *infinite regress*) que implica que existan catalizadores para la síntesis de catalizadores lo que cautivó a Rosen y debe haber pensado intensamente en este aspecto (que es corazón de la clausura metabólica) durante su tesis de doctorado hacia 1957.

Rosen resuelve esto con una potente intuición que afirma que en algunos sistemas es *el propio sistema en su totalidad* que repara a todos los reparadores. Este es el teorema principal de Rosen y una labor de esta tesis es mostrar cómo esta idea, un tanto esotérica, se alinea con la idea de clausura. A pesar de la gran generalidad y abstracción que hacen atractiva esta teoría, ha resultado difícil, por estos mismos elementos, obtener resultados operativos de ella.

## ¿Demuestra Rosen su intuición?

Rosen expresa su resultado fundamental en 1959 y da una demostración en la cual mezcla teoría de conjuntos y la formidable teoría de categorías. Asegurándose así que ningún

---

<sup>1</sup>se recuerda que, para seguir los usos y costumbres del laboratorio del Dr. Letelier se debería usar la palabra “reemplazar”. Aquí usaré “reparar” para facilitar a los futuros lectores el leer los trabajos originales

biólogo lo va a poder leer nunca. La demostración de Rosen no es una demostración, en el sentido canónico de la palabra, es más bien una formulación técnica de un requerimiento o condición necesaria. Rosen muestra que una condición muy excepcional (se explicará más adelante) es posible de que exista. Su demostración original es confusa y, como todas las ideas cuando se presentan originalmente, no sigue un camino recto. Pero en esencia dice que, en algunos sistemas  $(M, R)$ , cuando se consideran en un mismo pie todas las reacciones de tipo  $M$  y las de tipo  $R$  sería posible encontrar que todos los  $M$  son generados por  $R$  (obvio, por la definición misma de los  $R$ ) y que *todos* los  $R$  son producidos por variadas mezclas y conjunciones de  $M$  y  $R$  (esto no es obvio y constituye la intuición de Rosen). Además, hay que resaltar otra elección desafortunada de terminología; Rosen llama a los sistemas  $(M, R)$  que cumplen con su propiedad *replicative*  $(M, R)$  systems. El uso de la palabra “replicative” da a entender erróneamente que se está hablando de reproducción .. cuando en realidad se está hablando de una invariancia en la organización metabólica del sistema.

### **De $(M, R)_{1959}$ a $(f, \Phi)_{1972}$**

Dada la complejidad de la formulación inicial de Rosen, no es de extrañar que haya buscado formulaciones más sencillas y elegantes. En 1972 (Rosen, 1972) elabora la versión final de su intuición y debo darla aquí ya que aparece un nuevo concepto que tiene alguna relevancia para esta tesis.

Los sistemas  $(M, R)$  son un modelo abstracto del metabolismo definido únicamente en términos relacionales. Las entidades u objetos a las que se les presta interés son:

- las relaciones de causación material (la transformación de moléculas)
- las relaciones de causación eficiente (la participación de catalizadores)

En particular Rosen planteaba que los organismos vivos realizaban fundamentalmente

una función  $f$  (que representa el conjunto de reacciones metabólicas) la cual actuaba sobre un conjunto  $A$  (que representa a los sustratos) y generaba elementos de un conjunto  $B$  (los productos).

$$A \xrightarrow{f} B \quad (14)$$

En este punto nos vuelven a atenazar las diferencias entre conjuntos ( $A$ ) e instancias ( $a$ ). Hay que recordar que en la primera formulación de Rosen (1958) para poder usar el lenguaje de las funciones se asume que el dominio (i.e.  $A$ ) define el conjunto de moléculas *permisibles* de ser transformadas; es decir  $A = a_1, a_2, \dots, a_n$ . Además, uno quiere una teoría del metabolismo que acepte, para definir y poder operar con  $f$ , pequeñas variaciones en el conjunto  $A$ , sin tener que redefinir la función  $f$ . Específicamente situaciones como la aparición de una nueva reacción debido a la existencia de una nueva molécula como en el caso del operón *lac* bajo variaciones de concentración de glucosa y lactosa. Por lo tanto la ecuación 14 se debe pensar así:

$$f : A = (A_1 \times A_2 \times \dots \times A_p) \rightarrow (B_1 \times B_2 \times \dots \times B_q) = B$$

$$\mathbf{a} = (a_1, a_2, a_3, \dots, a_p) \rightarrow f(\mathbf{a}) = (b_1, b_2, \dots, b_q) = \mathbf{b}$$

La función  $f$  representa la acción conjunta de los miles de  $M_i$  y por ello se le asocia al metabolismo. Rosen además sobre-interpreta 14 y asume que la transformación  $A \rightarrow B$  refleja la causa material del proceso metabólico y que  $f$  refleja la causa eficiente de ese mismo proceso. De esta forma Rosen planteaba que la causa material, es decir, la materia que permite la ocurrencia de  $b$  era  $a$ , pero que dado que el agente que causaba la reacción era  $f$ , este debía ser la causa eficiente.

Siguiendo su formulación inicial interpretó a  $R$  como la función  $\Phi$  (originalmente

$R$ ) que toma los metabolitos  $b \in B$  y los transformaba en  $f^{-1}$  que está en el conjunto  $Map(A, B)$ . Es decir,  $\Phi$  constituiría la causa eficiente de  $f$  y  $\mathbf{b}$  su causa material.

$$\Phi(\mathbf{b}) = f \text{ con la condición, } \mathbf{b} = f(\mathbf{a}) \text{ (para algún } \mathbf{a} \in A) \quad (15)$$

$\Phi$  se denomina *selector* y es una función que, tomando el resultado concreto del metabolismo, define (dictamina la estructura de) el metabolismo.

$$\Phi : B \rightarrow Map(A, B)$$

$$b \rightarrow \Phi(b) = f \text{ donde para algún } a \text{ se tiene } f(a) = b$$

$$\Phi \in Map(B, Map(A, B)) = \text{conjunto de selectores.}$$

En este enfoque  $b$  es producido por  $f$ ... el cual es producido por  $\Phi$ . Se tienen así los primeros pasos de un sistema  $(M, R)$  pero ahora con un formalismo basado estrictamente en funciones  $(f, \Phi)$ . Con este formalismo, **que no considera para nada la conectividad de la red metabólica subyacente que se quiere entender**, Rosen no puede pretender redemostrar el teorema de la existencia de “componentes centrales”, pero sí de clarificar su intuición fundamental. Lo que busca Rosen es encontrar un “algo” que genere al selector  $\Phi$  pero que esté definido por el mismo sistema. Ese algo, denotado por  $\beta$ , se le debe pedir lo siguiente:

$$\beta(f) = \Phi \quad (16)$$

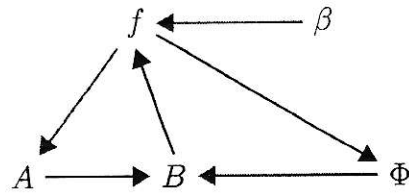
La figura 8 explica mejor la situación.

Que es este misterioso  $\beta$ ? Formalmente es una función que tiene como dominio  $Map(A, B)$  y como recorrido a  $Map(B, Map(A, B))$  y que colectivamente cumple las siguientes res-

---

<sup>1</sup>El rigor matemático exige aquí una aclaración:  $\Phi$  es una función que toma elementos de  $B$  y retorna funciones de  $A$  sobre  $B$

### Diagrama de Rosen: pre-clausura



**Figura 8:** La causa eficiente de  $f$ . Este diagrama muestra las distintas relaciones entre los operadores. EL metabolismo (representado por  $f$ ) es la causa de las moléculas que se encuentran ( $f(a) = b$ ). Existe un operador ( $\Phi$ ) que produce al metabolismo según:  $\phi(b) = f$ . la pregunta es: ¿Quién produce a  $\Phi$ ? Una manera es invocar a otro operador ( $\beta$ ) que formalmente cumple:  $\beta(f) = \Phi$ . Pero cual es la verdadera naturaleza de  $\beta$ ? Las flechas rojas denotan *causas eficientes* es decir operadores que actúan sobre elementos. Las flechas negras denotan transformaciones entre entidades.

tricciones:

$$f(a) = b, \Phi(b) = f, \beta(f) = \Phi \quad (17)$$

La intuición de Rosen se plasma aquí afirmando que  $\beta$  es posible encontrar un operador que este determinado de esta manera. Desgraciadamente nunca da un ejemplo que se pueda trabajar. La clarificación de  $\beta$  está al centro de muchos trabajos sobre esta manera de pensar, aquí recomiendo específicamente leer (Letelier y cols., 2006, 2005). Pero a la cual no le asigno mucho valor ya que creo que no es central para entender esta tesis. Hay tres aspectos que quiero reseñar:

1. Rosen plasmó, en su libro de 1991, estas relaciones en un diagrama que ha tenido amplia difusión y con una frase interesante: “Los seres vivos contienen todas sus causas eficientes”.
2. en (Letelier y cols., 2006) se afirma que la intuición de Rosen, es decir que exista un  $\beta$  con las propiedades indicadas sólo puede ocurrir si los conjuntos involucrados tienen restricciones sobre su “tamaños relativos”, según:

$$|A| \approx |B| \approx |\text{Metabolismos}| \approx |\text{Selectores}| \quad (18)$$

Interesantemente, esta simple ecuación ya ha dado cuenta de un trabajo que la usa para entender la complejidad celular o de los sistemas de manufacturas (Rietman y cols., 2011).

3. Que “es”  $\beta$ ? Una *quick and dirty way* de comprender la naturaleza de  $\beta$  es pensar en “resolver” la ecuación que la define.

$\Phi(b) = f$  ... supongamos que  $b$  fuese invertible  $b^{-1}$  así tendríamos

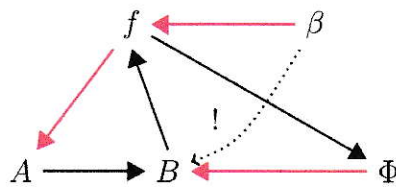
$\phi(b)b^{-1} = (f)b^{-1}$  .. y que además actuase “en ambos lados”

$\Phi[(b)(b^{-1})] = b^{-1}(f)$  pero  $(b)(b^{-1}) = Id$ , damos vuelta la ecuación y comparamos con la definición de  $\beta$

$$\beta = b^{-1} \quad (19)$$

Esta manipulación, hecha sin ninguna justificación muestra que  $\beta$  tiene algo que ver con el output del metabolismo (i.e. con  $b$ ), pero no le es idéntico. La figura 9 clarifica que relación existe entre  $\beta$  y  $B$ .

Diagrama de Rosen: intuición fundamental

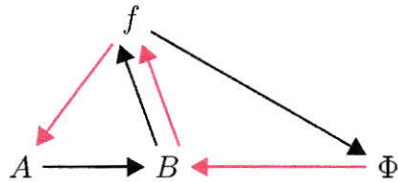


**Figura 9:**  $\beta$  tiene alguna relación con  $B$ . Rosen tuvo una intuición;  $\beta$  no es un operador externo sino, que en *algunos* sistemas  $(M, R)$  es obtenible (derivable) de un elemento de  $B$ . Así, la línea punteada refleja su intuición y su grado de magia aún no explicada (!).



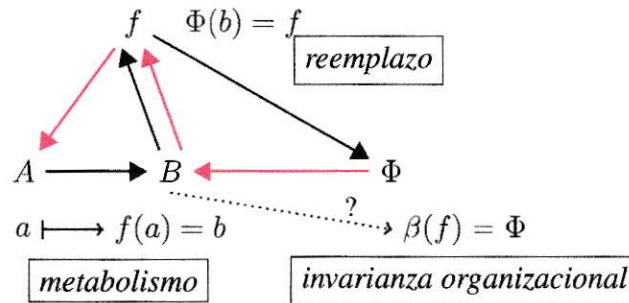
Con estas ideas Rosen generó la figura 10, un diagrama que resume su enfoque, sin embargo podría resultar provechoso dar una versión más completa y comentada, como la de la figura 11

Diagrama de Rosen



**Figura 10:** Este diagrama resume el trabajo y la intuición de Robert Rosen. La idea central es que: *Los organismos contienen todas sus causas eficientes*

Diagrama de Rosen-II



**Figura 11:** Se han adicionado las tres ecuaciones que definen los tres operadores ( $f$ ,  $\Phi$ ,  $\beta$ ) además que se muestran los nombres de los tres procesos manejados por estos operadores. Adicionalmente se muestra (en línea punteada y con un punto de interrogación) la relación existente entre  $B$  y  $\beta$ . Un desafío, es poder explicar en términos de redes metabólicas concretas cual es la naturaleza de  $\beta$ .

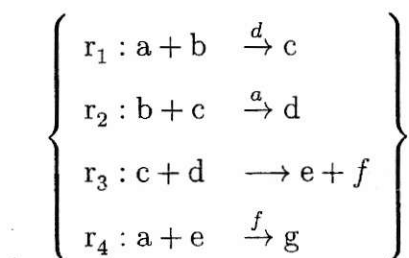
### Autopoiesis, Sistemas ( $M, R$ ) y causas

Las teorías de la autopoiesis y sistemas ( $M, R$ ), aunque causan un embrujo potente y presagian muchas potencialidades, son pobres en detalles de tipo técnico. Por eso en este capítulo, que introduce las ideas *técnicas*, no haré mucho hincapié en ellas. La definición

de clausura de Rosen o Maturana y Varela son muy generales y difíciles de llevar a la práctica. De hecho en 40 años de teoría nadie ha podido aún definir con precisión matemática que se debe entender por *organización autopoietica*. Eso sí, las ideas de conjuntos RAF y COT se pueden usar para definir clausura y es esa definición la que terminaré utilizando. En esta tesis, en su parte más técnica, buscaré clarificar que se debe entender por clausura y para empezar esta misión introduciré las teorías más técnicas sobre lo vivo.

## Conjuntos RAF

El formalismo planteado por W. Hordijk y M. Steel (Hordijk y Steel, 2004). Utiliza la Teoría de Conjuntos para realizar un análisis topológico de las reacciones de un metabolismo. Durante este análisis son de especial interés los conjuntos que son “reflexivos auto catalíticos” y “ $F$ -generados”, es decir, conjuntos de moléculas que son capaces de catalizar todas las reacciones que generan al conjunto a partir de un conjunto denominado  $F$ . Los metabolismos RAF se consideran al menos topológicamente viables, es decir, tienen la potencialidad de mantenerse en el tiempo, sin embargo este último aspecto (temporalidad) es ignorado completamente por este formalismo. A pesar de esta importante falencia, esta teoría contiene una definición matemática de clausura que se asemeja a la de las organizaciones químicas y establece además la importante distinción de la existencia de catalizadores. Así es como su noción de clausura se acerca a la biología.



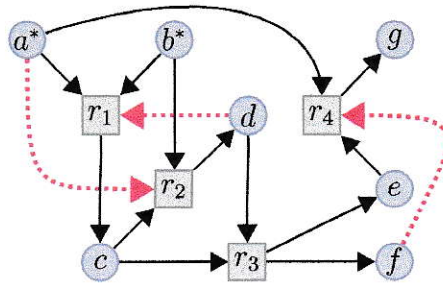
Este micro-metabolismo consta de:



- $X = \{a, b, c, d, e, f, g\}$  Conjunto de moléculas presentes
- $R = \{r_1, r_2, r_3, r_4\}$  Conjunto de reacciones
- $C = \{(d, r_1), (a, r_2), (f, r_4)\}$  Conjunto de *catalizaciones*
- $F = \{a, b\}$  Conjunto de alimentos (Food items)

En la figura 12 se ve una representación gráfica de este metabolismo. Obviamente las ideas presentadas aquí en forma simple y gráfica aceptan definiciones algorítmicas:

- sea  $X$  el conjunto total de moléculas que operan en la red
- una reacción  $r$  es un par ordenado  $(A, B)$  donde:
  - $A, B \subset X$
  - $A \cap B = \emptyset$
- si  $A = \{a_1, \dots, a_n\}$  y  $B = \{b_1, \dots, b_m\}$  entonces  $r$  representa una reacción química donde las moléculas de  $A$  (reactantes) se combinan en ciertas proporciones para producir las moléculas de  $B$
- $F$  (por “food”) es el conjunto de moléculas que alimentan a la red de reacciones. Es decir no son creadas por ninguna reacción de la red sino que incorporadas desde el medio.
- Sea  $R$  el conjunto de reacciones (es decir  $R = \{(A_1, B_1), (A_2, B_2), \dots, (A_r, B_r)\}$ ) se denomina a una *catalization* a un par  $(x, r)$  donde  $x \in X$  y  $r \in R$ . Se dice que  $x$  cataliza a la reacción  $r$ . Se denota  $C \subset X \times R$  al conjunto de catalizaciones.
- si  $r = (A, B) \in R$  se definen  $\rho(r) = A$  y  $\pi(r) = B$  y además consideremos un sub-conjunto  $R_1$  de  $R$ . Con estas nociones generalizamos la idea de los reactantes de una reacción  $r$  (es decir  $\rho(r)$ ) a la de los reactantes de un sub-conjunto de reacciones



**Figura 12:** Representación gráfica de  $Q = (X, R, C)$  para el conjunto de reacciones especificadas. Las cuatro reacciones ( $r_1, \dots, r_4$ ) (cuadrados), las 7 moléculas (círculos), los dos componentes del Food-set ( $a^*, b^*$ ) y las tres catalizaciones (flechas rojas) configuran un  $\text{CRS} = Q = (X, R, C)$

como:  $\rho(R_1) = \bigcup_{r \in R_1} \rho(r)$  podemos hacer lo mismo para los productos:  $\pi(R_1) = \bigcup_{r \in R_1} \pi(r)$  y así definir el soporte de  $R_1$  como:  $\text{supp}(R_1) = \rho(R_1) \cup \pi(R_1)$ . Así, el conjunto  $\text{supp}(R_1)$  (el soporte de  $R_1$ ) es el conjunto de todas las moléculas de  $X$  que tienen algo que ver (es decir son consumidas o producidas) por al menos una reacción en  $R_1$ .

- el sistema  $(X, R, C)$  se llama un *Catalytic Reaction System* sobre  $F$  (o CRS sobre  $F$ ). Algunas veces se encuentra escrito como  $Q = (X, R, C)$

## Definición de Clausura<sup>1</sup>

Con los elementos formales definidos arriba se puede definir muy precisamente lo que es la clausura. Sea un sub-conjunto  $R'$  de  $R$  (es decir un sub-conjunto de las reacciones totales), y  $X'$  de  $X$  (es decir un sub-conjunto de las moléculas totales). Se define la **clausura** de  $X'$  relativa a  $R'$  ( $cl_{R'}(X')$ ) como el sub-conjunto mínimo  $W \cap X$  que contiene a  $X'$  y que satisface la condición que para cada reacción  $(A, B) \in R'$  se tiene:

$$A \subseteq (X' \cup W) \implies B \subseteq W$$

Informalmente se tiene que  $cl_{R'}(X')$  es  $X'$  más todas las moléculas que pueden ser

<sup>1</sup>esta pequeña sección es importante

sintetizadas a partir de  $X'$  por la aplicación repetida de reacciones en  $R'$ . Nótese que  $cl_{R'}(X')$  está bien definida y  $cl_{R'}(X') \subset (\pi(R') \cup X')$ .

Dado un CRS:  $Q = (X, R, C)$  y un sub-conjunto  $R'$  de  $R$  se dice:

- $\mathcal{R}'$  es *reflexivo auto catalítico* (RA) en  $Q$  si: para todos los  $r \in \mathcal{R}'$  existe un  $x \in \text{supp}(\mathcal{R}')$  tal que  $(x, r) \in C$ .
- $\mathcal{R}'$  es *F-generado* si  $\rho(\mathcal{R}') \subseteq cl_{\mathcal{R}'} F$
- $\mathcal{R}'$  es *reflexivo auto catalítico y F-generado* (RAF) en  $Q$  si  $\mathcal{R}'$  es RA en  $Q$  y F-generado. En este caso se dice que  $R'$  es un conjunto RAF para  $Q$ .

Todas estas relaciones de conjuntos se pueden traducir en lo siguiente: Un sub-conjunto  $R'$  es RAF si:

- para cada una de sus reacciones existe un catalizador producido por alguna reacción en  $R'$
- cualquier reactante en  $R'$  puede ser construido recursivamente aplicando reacciones de  $R'$  a partir del conjunto de moléculas de  $F$ .

Se puede ver que la idea de **conjunto RAF** captura la idea de un sistema vivo a partir de algunos componentes externos (conjunto  $F$ ) y por su habilidad de auto catálisis cruzada puede autogenerarse. Como se verá más adelante las ideas de los conjuntos RAF, aunque interesantes (sobre todo porque representan el metabolismo como un grafo dual), permanecieron varios años sin mayor difusión. No es de extrañar esta relativa invisibilidad ya que el esfuerzo de Hordijk y Steele está encaminado a re-analizar el resultado de Kauffman que afirma que la probabilidad de tener una red auto catalítica es 1 (es decir certeza) a pesar que la probabilidad que un producto pueda catalizar cualquier reacción en la red ( $p$ ) sea muy baja. Así, en vez de resultados analíticos, estos autores usan simulaciones computacionales. Primero crean redes con un número arbitrario de reacciones (del orden 20-40).

Para cada una de estas redes ven si existe un conjunto RAF y repitiendo este procedimiento millones de veces pueden evaluar la probabilidad de aparición de un conjunto RAF. Todo el andamiaje matemático dado aquí es sólo el fundamento de un algoritmo de búsqueda que, dada una CRS, busca si tiene en su interior un conjunto RAF (no por nada los autores lo llaman el algoritmo RAF). El interés es que este algoritmo es polinomial respecto del número de reacciones, por lo tanto es resoluble por **fuerza bruta** por un computador. Las nociones RAF serán usadas en esta tesis, pero no para buscar en espacios multidimensionales conjuntos auto catalíticos, sino más bien como una herramienta para clasificar tipos de redes con clausura metabólica. Una de sus frases de la conclusión del paper es muy iluminadora respecto a la pregunta fundamental detrás de los conjuntos RAF:

In this way [i.e. using RAF ideas] it may be possible to shed some light on the question of whether autocatalytic, self-sustaining systems rely on a very special and unlikely biochemical balance, or whether they are a likely consequence of a sufficiently complex system of biochemical reactions.

## Teoría de la Organización Química

Esta teoría comienza definiendo un metabolismo por:

- un conjunto  $\mathcal{M}$  de moléculas que se transforman
- un conjunto  $\mathcal{R}$  de reacciones. La definición es muy parecida a RAF  $r = (A, B) = (A \rightarrow B) \in \mathcal{R}$ <sup>1</sup>
- el par  $\langle \mathcal{M}, \mathcal{R} \rangle$  se denomina una química artificial<sup>2</sup>.

La idea de clausura es fácil de entender: un sub-conjunto  $C \subseteq \mathcal{M}$  es **cerrado** si para todas las reacciones  $r = (A, B) = (A \rightarrow B) \in \mathcal{R}$  cuando  $A$  tiene los elementos en

<sup>1</sup>Hay una diferencia, COT usa *multiconjuntos* y RAF usa conjuntos

<sup>2</sup>Increible coincidencia con los sistemas  $(M, R)$

$C$  entonces  $B$  también tiene sus elementos en  $C$ . Intuitivamente esta definición quiere decir; si un conjunto  $C$  de moléculas tiene la propiedad que cualquier reacción  $R$  que tiene **todos** sus reactantes en  $C$  también tiene **todos** sus productos en  $C$  se dice que  $C$  es cerrado. Además dado cualquier sub-conjunto  $S \subseteq \mathcal{M}$  siempre es posible generar su clausura, denotada por  $G_{CL}(S)$ , buscando el conjunto  $C$  más pequeño que sea cerrado. Se dice que  $S$  genera al conjunto cerrado  $C = G_{CL}(S)$  y llamamos a  $C$  la **clausura** de  $S$ .

Un conjunto  $S \subseteq \mathcal{M}$  se dice **auto-mantenido** si todas las moléculas de  $S$ , y que son usadas por reacciones, son producidas por reacciones con reactantes en  $S$ <sup>1</sup>.

Un conjunto  $O \subseteq \mathcal{M}$  se denomina una **semi-organización** si es la vez cerrado y auto-mantenido.

## Una Innovación: Introducción de la Noción de Flujos, aparece el Tiempo

Pero, aunque la idea de que una semi-organización (como los conjuntos RAF) parece capturar la idea de un metabolismo autosustentable no garantiza que en su evolución temporal no se vaya a producir que algunas moléculas desaparezcan. Por lo tanto, se deben usar ideas de velocidad de consumo y de velocidad de producción de moléculas para definir ahora conjuntos que sean **mantenedor de masa** (mass-maintaining). Sea un sistema de química artificial  $\langle \mathcal{M}, \mathcal{R} \rangle$  con  $m = |\mathcal{M}|$  igual al número total de tipos moleculares y  $n = |\mathcal{R}|$  y sea  $\mathbf{M} = (m_{i,j})$  la matriz estequiométrica de  $(m \times n)$  que describe todas las reacciones metabólicas ( $m_{i,j}$  denota el número de moléculas del tipo  $i$  producidas por la reacción  $j$ ). Un sub-conjunto  $C \subseteq \mathcal{M}$  se denomina *mantenedor de masa* si existe un vector de flujos  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  tal que para cada reacción  $v_j > 0$ . Esta condición muestra que todas las moléculas van a mantener (o incrementar) su concentración.

---

<sup>1</sup>Observemos que los conjuntos RAF no son automantenidos

## **Una Anti-innovación? No hay Noción explícita de Catálisis.**

Un elemento que falta en esta teoría es la mención explícita de los catalizadores como tales. Es decir, un catalizador en esta teoría se comprende como una molécula que se encuentra a ambos lados de una reacción. En particular no hay nada similar a la noción de catalización de los conjuntos RAF. En todo caso la teoría COT ha servido para traer la idea de cambio temporal, y estabilidad cuantificada con sistemas dinámicos, al estudio de la clausura metabólica. Por este solo detalle merece ser estudiada con detalle y de hecho un nuevo trabajo ha explotado intuiciones publicadas por nuestro laboratorio (Jaramillo y cols., 2010) para mostrar la importancia de ciclos de reacciones en determinar estabilidad (Kreyssig y cols., 2012).

En resumen y consideración de todo este capítulo, desde que se han co-desarrollado técnicas computacionales para atravesar árboles de conceptos en bases de datos, y ideas de álgebra lineal ha sido posible entender como atacar el problema de la estabilidad de los sistemas de reacciones. A continuación veremos como las ideas simples presentadas aquí pueden tener un gran impacto cuando se hable de flujos y estados estacionarios en un organismo modelo.

# Antecedentes Dinámicos

Como se ha insistido anteriormente, los aspectos dinámicos de un sistema corresponden al punto de vista más cercano al mundo experimental ya que los objetos de estudio, las concentraciones de los metabolitos y sus cambios en el tiempo, corresponden precisamente a las entidades que pueden ser medidas empíricamente. Teniendo esto en consideración, durante esta tesis hemos prestado especial atención a este enfoque, de forma que los resultados obtenidos durante este trabajo tengan potencialmente algún impacto para el experimentalista. A continuación revisaremos algunos conceptos y teorías existentes que lidian con la dinámica del metabolismo para insertar nuestros resultados en un contexto adecuado y hacer una revisión que permita contrastar la naturaleza de este trabajo con otros puntos de vista en biología teórica.

## Ley de Acción de Masas

Como nos preocupará principalmente la tasa de cambio de distintas especies que conforman un metabolismo, repasaremos brevemente cómo se suelen representar las velocidades de reacciones químicas en términos de ecuaciones diferenciales. La Ley de Acción de Masas (propuesta por primera vez entre 1864 y 1866 (Gulberg y Waage, 1864)) informa sobre la tasa de ocurrencia de una reacción. Inicialmente se enfocaba en los mecanismos que llevan a una solución en la que ocurren reacciones hasta el equilibrio, aunque posteriormente el enfoque se dirigió a la dinámica fuera del equilibrio. El planteamiento moderno se fundamenta en la física estadística y la deducción sigue más o menos los siguientes lineamientos:

Contémplese un volumen en el cual hay especies químicas capaces de interactuar a

través de la siguiente reacción química:



Considérese además que el medio de reacción se encuentra a temperatura constante y bien mezclado, es decir, es homogéneo. Bajo estas condiciones la ocurrencia de una reacción química depende principalmente de dos factores:

- El encuentro de las especies moleculares involucradas en el volumen de reacción
- Que la colisión sea suficientemente energética y con la geometría adecuada

Si la solución está bien mezclada y en ausencia de interferencia, la tasa de colisiones por unidad de tiempo depende esencialmente de la cantidad de *masa activa* de las especies que se encuentre en el volumen, es decir, de la actividad química o concentración efectiva<sup>1</sup>. El segundo factor dependerá de la temperatura de la solución, la presencia o ausencia de catalizadores y otras propiedades de las especies químicas. La actividad química suele medirse de forma empírica pues calcularlo requeriría en general de la realización de simulaciones de química cuántica (e.g. para estimar la reactividad química) y el cálculo de una cantidad de configuraciones geométricas posibles muy elevada.

Con esto podemos escribir la probabilidad de ocurrencia de una reacción R en un intervalo de tiempo  $\delta t$  como:

$$P(R) = k_R [A]^m [B]^n \delta t \quad (21)$$

donde  $k_R$  es un término constante que contiene el efecto de la temperatura, catalizadores y particularidades energético-probabilísticas de la reacción,  $[A]$  y  $[B]$  corresponden a las

---

<sup>1</sup>Por distintos efectos como la quelación o aglomeración de moléculas debe considerarse la concentración efectiva y no la concentración total.



concentraciones efectivas de estas especies químicas y  $m$  y  $n$  a los coeficientes estequiométricos de la reacción. A partir de este punto puede plantearse una ecuación diferencial estocástica para describir la concentración de las distintas especies, pues cada vez que ocurre una reacción las concentraciones cambian y esto a su vez cambia la probabilidad de ocurrencia de la reacción. Si consideramos que hay suficientes partículas como para que se cumpla la Ley de los Grandes Números, podemos tender al límite donde la concentración es continua y por lo tanto aproximar la ecuación estocástica que indica el cambio de la concentración de  $A$  como una ecuación diferencial ordinaria:

$$\frac{\delta A}{\delta t} = -mk_R[A]^m[B]^n \quad (22)$$

De esta misma forma puede describirse el curso temporal de la concentración de las demás especies. Si además se considera que la reacción inversa ocurre, obtenemos que el cambio neto en las concentraciones realmente es:

$$\frac{\delta A}{\delta t} = -mk_R[A]^m[B]^n + mk_{R^{-1}}[C]^p[D]^q = m(k_{R^{-1}}[C]^p[D]^q - k_R[A]^m[B]^n) \quad (23)$$

Para expresar la tasa de cambio de concentraciones de todo un sistema se recurre a la matriz estequiométrica  $S$  y a un vector de velocidades de reacción  $v(X, k)$ , describiendo el cambio en las concentraciones  $X$ :

$$\frac{\delta X}{\delta t} = Sv(X, k) \quad (24)$$

donde  $X$  es el vector de concentraciones de todas las moléculas.

Estas ideas son la base general de los sistemas que simulan la ocurrencia de reacciones para mecanismos simples. En particular, la ecuación 23 describe el curso temporal de las concentraciones de especies involucradas en reacciones que siguen el mecanismo plantea-

do anteriormente y es lo que se conoce como Ley de Acción de Masas. Más sobre este tema puede encontrarse en los artículos originales o de forma muy didáctica en el trabajo de Daniel Gillespie (Gillespie, 1977). Por supuesto que en biología las reacciones suelen estar sujetas a la acción de un catalizador, cuya disponibilidad no es absoluta ni constante, por lo que se generan dinámicas como la conocida ecuación de Michaelis-Menten (resuelta analíticamente recién en el 1997 (Schnell y Mendoza, 1997)) que consideran saturaciones y cambios en la concentración de catalizador. Otras complicaciones, como la existencia de compartimentalización y fenómenos de difusión serán ignoradas para simplificar el análisis, considerando siempre que el sistema se encuentra bien mezclado.

## **Estado estacionario**

Desde los tiempos de Henri Poincaré se hizo evidente que las ecuaciones diferenciales encerraban secretos difícilmente accesibles, excepto en ciertos casos particulares. Hoy resolvemos el problema de analizar sistemas complejos utilizando computadores, aunque incluso esta es una opción limitada a casos de dimensión pequeña. Para los matemáticos y físicos del momento, la solución fue desechar el rigor numérico y aproximar las ecuaciones desde un punto de vista cualitativo, momento en el cual se hizo evidente que lo que había que estudiar primero eran los llamados conjuntos límite, es decir, puntos de equilibrio y orbitas periódicas, atractores, etc...

En biología nos concentramos en los estados estacionarios, no porque consideremos que los organismos deben permanecer siempre en un estado estacionario, sino porque como primera aproximación son muy informativos ya que dada la complejidad actual de los organismos, los estados estacionarios son de lo máximo que podemos abarcar desde el punto de vista analítico; además los puntos fijos se corresponden bastante con la idea de invarianza a niveles de composición, de estructura, etc... es decir, corresponden a estados

homeostáticos donde la actividad de un sistema está en tal grado de balance o equilibrio que todo lo que es degradado es reconstruido por el sistema.

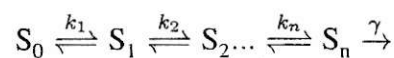
En el estado estacionario de un sistema descrito por la matriz estequiométrica  $S$  y un vector de velocidades de reacción  $v$ , se cumple lo siguiente:

$$\frac{\delta X}{\delta t} = Sv(X, k) = 0 \quad (25)$$

Como la matriz  $S$  es no nula, la atención se vuelca hacia aquellos vectores cuyo producto con la matriz es nulo, es decir, el kernel de la matriz  $S$ . Sobre las condiciones que se requieren para la existencia de estos vectores, la concentración de los distintos metabolitos en ese momento, las propiedades del punto de equilibrio y los efectos de pequeñas perturbaciones se ha escrito mucho. Una de las grandes teorías que cristaliza estos estudios es la teoría del control metabólico, la cual introduciremos a continuación.

## Análisis del Control Metabólico

La teoría del MCA es un potente estudio matemático del metabolismo que tuvo su nacimiento principalmente de la mano de al menos tres grupos de investigadores independientes (Heinrich y Rapoport, 1974; Kacser y Burns, 1973; Reder, 1988). Por supuesto existen otros autores que también contribuyeron al campo como Higgins y Savageau (Higgins, 1963; Savageau, 1976) e incluso cómicamente, la misma matemática fue resuelta en un contexto distinto por (Zhou y cols., 2005). Su preocupación inicial, o lo que podría llamarse el *toy-system*, es una cadena enzimática lineal, es decir, reacciones cuyos productos constituyen los reactantes de la siguiente reacción como puede verse a continuación:



Esta decisión obedece la motivación por estudiar el metabolismo del momento, que era esencialmente la ingeniería metabólica (en sus inicios, no su forma moderna) y por lo tanto el estudio del balance o distribución de masa en una vía o *pathway*. Esta perspectiva simplificadora no limita sus resultados, pues su modo de análisis puede extrapolarse fácilmente a redes de reacción con ramificaciones. A continuación veremos las herramientas matemáticas utilizadas por esta teoría pues resultan edificantes.

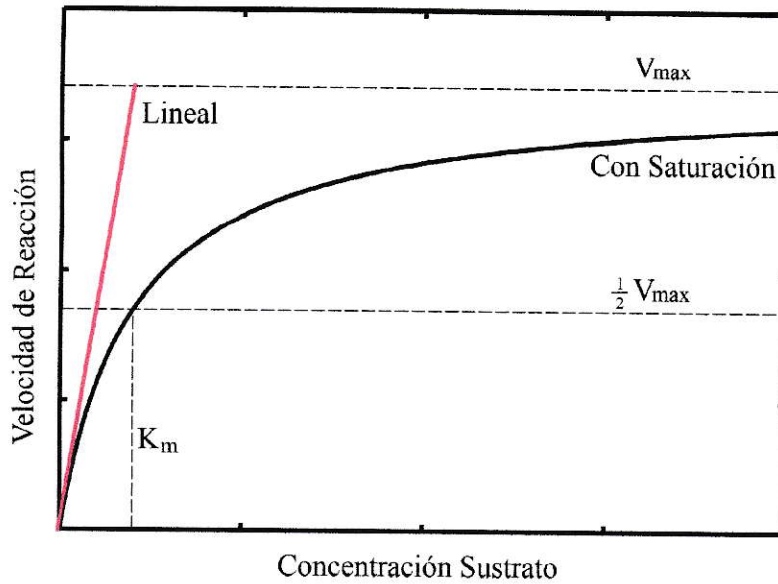
## Cinética enzimática

En el trabajo de MCA las enzimas o catalizadores se consideran elementos esencialmente estáticos, sólo susceptibles a efectores. Nuevamente esto se debe a que se intentaba considerar las concentraciones *in vivo* y modificaciones de estas. Otra simplificación introducida por estos autores, aunque reconocen que el modelo más acertado es el de Michaelis-Menten ilustrado por la ecuación 26 y la figura 13, es el de considerar que los organismos vivos funcionan probablemente con las enzimas no saturadas, es decir en la zona lineal de la cinética, particularmente cuando  $S_i < \frac{1}{2}K_m$ :

$$v = \frac{V_{max}S}{K_m + S} \approx \frac{V_{max}S}{K_m} \quad (26)$$

Esta aproximación permite hacer un análisis iluminador sin mucho trabajo dedicado exclusivamente a las ecuaciones. Así, la velocidad de reacción del sistema se describe por estas ecuaciones:

$$\begin{aligned} \frac{\partial S_1}{\partial t} &= k_1 S_0 - (k_{-1} + k_2) S_1 + k_{-2} S_2 \\ \frac{\partial S_i}{\partial t} &= k_i S_{i-1} - (k_{-i} + k_{i+1}) S_i + k_{-(i+1)} S_{i+1} \\ \frac{\partial S_n}{\partial t} &= k_n S_{n-1} - (k_{-n} + \gamma) S_n \end{aligned} \quad (27)$$



**Figura 13:** Curva típica de cinética *à la* Michaelis-Menten. En rojo se describe el comportamiento que tendría la velocidad de reacción si no hubiese saturación. Imagen original cortesía de usuario U+003F de Wikipedia

A continuación los autores reconocen que un buen punto de partida es el de buscar un estado estacionario, pues este debe reflejar el estado de la cadena enzimática durante la mayor parte del tiempo. Para simplificar los cálculos se propone el siguiente cambio de variables:

$$\tau_i = \frac{1}{k_i + k_{-i}}, \quad q_i = \frac{k_i}{k_{-i}} \quad (28)$$

Aparte de ser algebraicamente apropiado, este cambio de variables permite entender el sistema ya no en términos de las constantes cinéticas, sino que de la posición o constante de equilibrio para la reacción  $i$ -ésima  $q_i$  y su tiempo característico o de relajación  $\tau_i$ . Para encontrar el estado estacionario es necesario igualar las tasas de cambio de los distintos metabolitos a 0, ya que esa es la posición de equilibrio, obteniéndose un sistema de ecuaciones algebraicas no homogéneas (es último pues  $S_0$  y  $\gamma$  son términos que no están asociados a las variables del sistema). Luego de arduos cálculos, es posible despejar las

concentraciones al equilibrio de las distintas especies moleculares:

$$S_i = S_0 \prod_{j=1}^i q_j \frac{1 + \gamma \sum_{k=i+1}^n \tau_k (1 + q_k) \prod_{m=k+1}^n q_m}{1 + \gamma \sum_{k=1}^n \tau_k (1 + q_k) \prod_{m=k+1}^n q_m} \quad (29)$$

Ahora que contamos con esta base teórica, podemos dedicar nuestra atención a nociones mas informativas, como lo son el flujo y los coeficientes de control y que constituyen los resultados más importantes de esta línea de trabajo, la cual si bien es la base fundamental de la biología de sistemas actual, no ha visto grandes avances desde su planteamiento original. Por ejemplo, el flujo de masa a través de la totalidad de la cadena enzimática en el estado estacionario corresponde a  $v_g = \gamma S_n$ , es decir:

$$v_g = \frac{\gamma S_0 \prod_{j=1}^n q_j}{1 + \gamma \sum_{k=1}^n \tau_k (1 + q_k) \prod_{m=k+1}^n q_m} \quad (30)$$

De esta ecuación se puede obtener información interesante, por ejemplo puede verse que el concepto de enzima limitante de reacción, presente incluso en los textos modernos de bioquímica (Devlin, 2004) es una aproximación que desecha los aspectos sistémicos pues el flujo depende de los parámetros de todas las enzimas de la vía. Para ser justos, si consideramos que la mayoría de las reacciones son reversibles ( $q_i = 1$ ) el flujo se simplifica a

$$v_g = \frac{S_0}{\frac{1}{\gamma} + 2 \sum_{k=1}^n \tau_k} \quad (31)$$

de donde se desprende que la enzima con el tiempo de relajación más grande es aquella que tiene mayor influencia en la vía. Otro aspecto interesante que se puede derivar de la

ecuación 30 es el caso de las reacciones irreversibles, es decir cuando  $q_s \gg 1$ :

$$v_g = \frac{S_0 \prod_{j=1}^{s-1} q_j}{\tau_s + \sum_{k=1}^{s-1} \tau_k (1 + q_k) \prod_{m=k+1}^{s-1} q_m} \quad (32)$$

dejando en claro que si existe una reacción esencialmente irreversible el flujo sobre toda la vía queda determinado esencialmente por esa enzima y todas las que la preceden. A esta ecuación todavía le quedan un par de vueltas de tuerca, pero para ello necesitamos introducir los coeficientes de control, los elementos verdaderamente fuertes detrás de MCA.

### Coefficientes de control

Como ya mencioné anteriormente, el contexto en el que se gestó esta teoría es el de la ingeniería de metabolismo, por lo que la pregunta natural es: Si vamos a modificar la existencia de una enzima ¿cuál será aquella que traiga mayores beneficios? En particular, cual modificará en mayor medida el flujo? La respuesta analítica es tomar la derivada del flujo respecto de la velocidad de una reacción aislada, así, nombrando a esta magnitud coeficiente de control de flujo se puede escribir:

$$C_i^{v_g} = \frac{\partial v_g}{\partial v_i} \quad (33)$$

donde  $v_i$  es la velocidad de reacción de esa reacción aislada, la cual en la mayoría de las cinéticas enzimáticas aumenta conforme aumenta la concentración de la enzima. Considerando una aproximación lineal podemos descomponer la ecuación anterior:

$$C_i^{v_g} = \frac{\frac{\partial v_g}{\partial \tau_i}}{\frac{\partial v_i}{\partial \tau_i}} \quad (34)$$

Siguiendo la cinética expuesta en MCA, un aumento de la cantidad de enzima presente significa esencialmente un aumento de las constantes  $k_i$  y  $k_{-i}$ . En términos de nuestro cambio de variables,  $q_i$  no se ve modificado, pues es relativo a la posición de equilibrio, así que valdría la pena explorar el cambio en el flujo respecto de  $\tau_i$ . Considerando que la enzima reacciona con los sustratos en sus concentraciones en el estado estacionario:

$$v_g = v_i = \frac{1}{\tau_i(1 + q_i)}(q_i S_{i-1} - S_i) \quad (35)$$

y luego la variación del flujo respecto a  $\tau_i$  es:

$$\frac{\partial v_g}{\partial \tau_i} = \frac{1}{1 + q_i} \frac{\partial}{\partial \tau_i} \left[ \frac{1}{\tau_i} (q_i S_{i-1} - S_i) \right] \quad (36)$$

Obteniendo así uno de los términos de la ecuación 34. Para la velocidad de alguna enzima en particular la expresión es básicamente la misma. Así, los autores de MCA llegan a una expresión analítica para  $C_i$ , haciendo notar varias propiedades sobre este coeficiente, en especial que  $0 \leq C_i \leq 1$  y que mientras mayor el coeficiente, más grande es el impacto de un cambio en esta enzima sobre el flujo de la vía. Una derivación similar puede llevarse a cabo para el coeficiente de control de metabolitos, es decir, como cambia la concentración de un metabolito en el estado estacionario cuando cambiamos la velocidad de la reacción  $i$ -ésima. Anotaremos este coeficiente  $C_i^{S_i}$ . Luego de estas definiciones matemáticamente escabrosas podemos pasar a los resultados principales de MCA, los cuales capturan propiedades sistémicas de los metabolismos, sin embargo no capturan la idea de clausura.

## Teoremas de MCA

Respecto de los coeficientes de control, MCA presenta dos grupos de teoremas que corresponden a su contribución más importante para la comprensión del metabolismo:



**Teoremas de la Sumación:** Estos teoremas indican que la suma de los distintos coeficientes de control tiene un valor fijo, en particular

$$\begin{aligned} \sum_i^{n+1} C_i^{v_g} &= 1 \\ \sum_j^{n+1} C_j^{S_i} &= 0 \end{aligned} \tag{37}$$

lo que significa que los flujos por las vías y las concentraciones en el estado estacionario de los distintos metabolitos son ambas propiedades sistémicas, es más, su control se encuentra compartido por todas las reacciones del metabolismo, independiente de si estas son anteriores, o posteriores a la formación del metabolito. Por supuesto esto no significa que los coeficientes de control sean fijos, al contrario, indica que cuando una velocidad de reacción cambia y posiblemente su coeficiente de control también se vea afectado, entonces la ganancia o pérdida de control será compensada por el resto de la red, que respectivamente perderá o ganará control.

**Teoremas de Conectividad** Los teoremas de conectividad, correspondientes al segundo grupo de teoremas de MCA no son realmente relevantes para esta tesis y requieren la explicación de otras magnitudes conocidas como los coeficientes de elasticidad, por lo que no se incluirán aquí más que esta mención por completitud.

Al menos los teoremas de sumación son notables por su respuesta sistémica, la cual ha sido validada experimentalmente (Torres y cols., 1986; Fell, 1992) pero MCA sufre una deficiencia seria para entender los organismos vivos. El hecho que limita esta teoría como plataforma para estudiar el metabolismo es que en este caso se ha olvidado el hecho fundamental de que las enzimas, en el caso del metabolismo, constituyen también metabolitos, lo que conduce a dinámicas altamente no-lineales en las que además la cinética fuera y alrededor del equilibrio es responsable de gran parte del comportamiento. En otras palabras, se ha olvidado la clausura.

## Teoría de la Organización Química

La teoría de la organización química comparte muchas definiciones con los Conjuntos RAF, en particular las ideas de reacción (aunque COT utiliza multiconjuntos para mantener la estequiometría explícita), soporte y notablemente clausura. Para evitar la repetición de contenido, en este capítulo nos limitaremos a revisar las nociones dinámicas que se han desarrollado en COT.

La primera definición relevante en esta teoría es la de conjunto **auto-sostenido**. Esta definición estudia la relación entre la matriz estequiométrica y el vector de velocidades de reacción para examinar la posibilidad de que el sistema sea capaz de mantener su masa en el tiempo. El producto  $Sv$  corresponde a la tasa neta de aparición o desaparición de las distintas especies de un sistema en el cual las reacciones ocurren a las velocidades descritas por el vector  $v$ . Así, un conjunto auto-sostenido es un conjunto de moléculas, las cuales reaccionado a alguna tasa  $v^+$ , es capaz de mantener o aumentar la masa total de todas sus moléculas. Si bien la función detrás del vector  $v^+$  puede tomar cualquier forma, considerando que en general la cinética de las reacciones bioquímicas es altamente no lineal, es necesario contar con definiciones básicas que delimiten los objetos de interés, por lo que esta primera definición no trabaja al nivel de detalle de la forma funcional que  $v$  pueda tener. La idea de conjunto auto-sostenido extiende la noción de un conjunto semi-sostenido, el cual es un requerimiento similar a éste pero sin la noción de tiempo como condición, es decir, la condición de conjunto auto-sostenido es más estricta que la de auto-semi-sostenido. Resumiendo, un sistema químico se dice auto-sostenido si existen vectores positivos  $v^+$  tales que:

$$Sv^+ \geq 0 \quad (38)$$

La siguiente definición, que está al centro de la teoría COT, es la de **Organización**. Una Organización se define como un conjunto cerrado (condición estática) y auto-sostenido

(condición dinámica). Es decir, una Organización es un sistema de reacción químico en el cual la masa de todas las moléculas se mantiene o incrementa, mientras que el tipo de moléculas presentes se mantiene constante, pues ningún tipo molecular aparece o desaparece. Esta es una característica importante para los sistemas vivos, pues les permite mantener las relaciones entre componentes dándoles una continuidad material. Si se violase la primera condición, aparecerían tipos moleculares nuevos todo el tiempo, por lo que se generaría interferencia, o interacciones complejas y novedad, cambiando constantemente así la organización del caso. Por otro lado si una de las especies moleculares agota su existencia, esta no estará disponible para participar de las reacciones, forzando una reestructuración de la red y por lo tanto violando la invarianza organizacional, propiedad característica de los seres vivos.

## **Dinámica y COT**

Si bien COT ha sido utilizada amplia y exitosamente para estudiar una multitud de sistemas de naturaleza diversa Centler y Dittrich (2007); Matsumaru y cols. (2005, 2006), hasta ahora su fuerte ha sido el enfoque jerárquico que provee en términos de las Organizaciones y sub-Organizaciones que contiene un sistema. Aunque estos aspectos son sumamente interesantes desde el punto de vista biológico, pues nos informan sobre los sub-sistemas que podrían existir en un metabolismo, nos aleja de los aspectos dinámicos en los que nos queremos concentrar en este trabajo. A continuación revisaremos un par de teoremas que sí atacan directamente preguntas de origen dinámico y por lo tanto son atingentes al desarrollo de esta tesis.

## COT y Conjuntos Límite

Como se mencionó en la introducción, los puntos cualitativamente interesantes en un sistema dinámico corresponden a los llamados conjuntos límite, que pueden corresponder a

- Puntos fijos
- Órbitas periódicas
- Ciclos límite
- Atractores extraños

y que en general son conjuntos de puntos hacia los cuales tienden las órbitas de los sistemas dinámicos cuando se los deja evolucionar por un tiempo infinito. Así, los puntos fijos están constituidos por un conjunto con un punto único, mientras que las órbitas periódicas son trayectorias cerradas y los ciclos límite corresponden a órbitas espiraladas que van cerrándose cada vez más sobre sí mismas. Por su lado los atractores extraños tienen una estructura interna más compleja. Los conjuntos anteriores describen tendencias o tipos de trayectorias que podría seguir un sistema dinámico. Si además estos conjuntos son estables tendrán una cuenca de atracción, lo que significa que es altamente probable encontrar a un sistema en un estado perteneciente a un conjunto límite.

Al respecto, COT tiene un teorema inicialmente demostrado para puntos fijos Dittrich y di Fenizio (2007) pero que se ha extendido Peter y Dittrich (2011) para los conjuntos límite acotados que no incluyan los ejes del sistema (pues allí ocurre un cambio cualitativo). Este teorema indica:

**Teorema 1** *Si la dinámica de un sistema de reacción tiene conjuntos límite acotados que no contengan a los bordes del sistema, entonces aquellos subconjuntos de moléculas pre-*

*sentas en estos estados que no posean subconjuntos propios entre las moléculas presentes corresponden a organizaciones.*

Con ayuda de este teorema podemos, al menos a groso modo, decir que el comportamiento a largo plazo de un sistema químico acotado tenderá a una organización. Es importante que no hemos hecho suposiciones de ningún tipo respecto del comportamiento de la cinética, más allá de que una reacción cuyos reactantes se encuentran presentes efectivamente ocurra con alguna tasa. Esto significa que el teorema se aplica a una gran clase de sistemas químicos, implicando entonces que la mayoría de los sistemas químicos efectivamente, al ser observados, se encontrarán conformando una organización.

Si bien el recíproco no es cierto, este teorema es útil a la hora de buscar sistemas o subsistemas que pudiesen tener puntos de equilibrio. Además de poder desarrollar propiedades para los equilibrios que se desprendan del hecho de que estos son Organizaciones y quizás restricciones adicionales que se le imponga al sistema, pues en realidad ser una Organización no es una condición fuerte. Notemos por ejemplo, que no todas las Organizaciones son estables. Las Organizaciones sólo tienen asegurado que podría algebraicamente existir un vector de flujo, ni siquiera se hace mención a su alcanzabilidad dentro del sistema químico. Es por esto que la estabilidad de un sistema debe ser tratada por otros medios.

## **COT y Viabilidad**

Para profundizar aún más en los aspectos dinámicos de COT, el grupo de investigadores que desarrolla principalmente COT plantea (Peter y cols., 2010) un refinamiento importante sobre la idea de Organización para lo cual primero introducen una forma funcional

para  $v$  basada en la Ley de Acción de Masas:

$$(v(X, k))_i = k_i \prod_{j=1}^m X_j^{a_{ij}} \quad (39)$$

Donde  $a_{ij}$  es el número de apariciones de la molécula  $s_j$  en el soporte de la reacción  $i$ -ésima. Luego, una organización se dice *viable* si, de forma similar a la condición de auto-sustentación, existe un vector positivo  $X$  tal que:

$$Sv(X, k) \geq 0 \quad (40)$$

para cualquier selección de  $k$ . Esta condición es sumamente fuerte y relevante, pues una organización viable es independiente de la cinética de las distintas reacciones.

Considerando que probar esta condición para todas las selecciones de  $k$  sería un despropósito, contamos con el siguiente teorema (el cual es válido bajo la cinética mencionada anteriormente):

**Teorema 2** *Una organización es viable si y sólo si*

$$\log_{\beta}(Ker_{\geq 0}^{>0}(S)) - Im(A) = \mathbb{R} \quad (41)$$

para cualquier  $\beta > 0$ , con las siguientes definiciones:

$$\begin{aligned} Ker_{\geq 0}^{>0}(S) &= v \in \mathbb{R}_{>0}^n : Sv \geq 0 \\ Im(A) &= y : y = Ax, x \in \mathbb{R}^m \end{aligned} \quad (42)$$

y el logaritmo aplicado elemento a elemento.

Este teorema indica, quizás de forma poco clara, cómo es que la estructura de las matrices  $A$  y  $S$  restringe los valores que  $k$  puede ocupar a modo de que la organización

se mantenga como tal, e incluso bajo ciertas condiciones indica que cualquier elección de  $k$  es válida. El tener esta última condición, la viabilidad, nos permite trabajar ignorando los parámetros cinéticos de un sistema y concentrarnos sólo en su topología, lo cual es sumamente liberador pues desde el punto de vista empírico no podemos realmente elegir el vector  $k$  del sistema.

Sin embargo, existe una limitante a la hora de determinar si una organización es viable o no: No existen métodos automáticos para poder probar la condición requerida por el teorema anterior, pues los métodos del álgebra lineal no pueden tratar con términos del tipo  $\log_{\beta} Ker_{\geq 0}^{>0}$  pues no son lineales.

## COT y una Descomposición

El último teorema que mencionaré sobre COT fue desarrollado recientemente en la Facultad de Ciencias de la Universidad de Chile (Veloz y cols., 2011). Este teorema, cuyos detalles requieren un trabajo entero por sí sólo, define una descomposición de un sistema en los siguientes conjuntos:

- Conjunto **no reactivo**: Especies que no están involucradas en reacción alguna
- Conjunto **de catalizadores**: Definido como las especies que actúan como catalizadores
- Conjunto **reactivo**: Especies que no pertenecen a los conjuntos anteriores

Esta descomposición no parece agregar información sustancial, pero el teorema sigue refinando la descomposición del último conjunto en:

- Conjunto de especies **sobreproducidas**: Aquellas que pueden ser producidas sin reducir la concentración de otras especies en el sistema

- Conjunto del **ciclo activo potencial**: Moléculas pertenecientes al conjunto reactivo que no son sobreproducidas

Desde aquí se descompone el PAC en sub-PACs disjuntos (que no se sobrelapan), de forma que la condición de auto-sustentación de una organización puede ser evaluada estudiando la auto-sustentación de los sub-PACs. En particular, el teorema de la descomposición afirma lo siguiente:

**Teorema 3** *Sea  $X$  un conjunto de moléculas y  $N$ ,  $E$ ,  $F$  sus conjuntos no-reactivo, catalizadores y de especies sobreproducidas respectivamente, y  $D_i$  sub-PACs que no se sobrelapan, es decir*

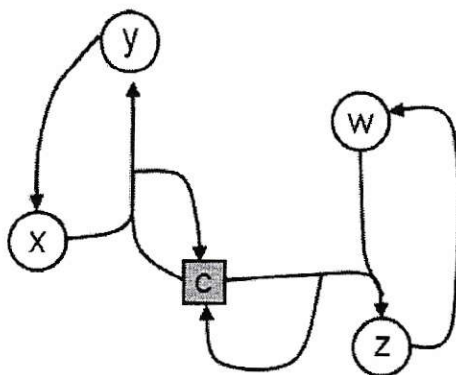
$$X = N \cup E \cup F \cup D_1 \cup \dots \cup D_d \quad (43)$$

*entonces  $X$  es auto-sostenida si y sólo si  $N \cup E \cup F \cup D_i$  es auto-sostenida para cada  $i = 1, \dots, d$*

Para ilustrar el sentido de este teorema hemos incluido un caso de ejemplo en extremo simple que sin embargo provee todos los elementos para comprender el alcance del teorema de la descomposición. En la figura 14 puede verse una organización que contiene 2 sub-PACs y un catalizador. Con este caso se puede ver que la dinámica de un sub-PAC (conformado por  $x, y$ ) es independiente de la dinámica del otro sub-PAC ( $z, w$ ) y que ambos PACs, en unión con el catalizador, deben ser auto-sostenidos para que todo el conjunto de moléculas también lo sea.

Este teorema es de fundamental importancia pues establece que la potencial estabilidad de un sistema complejo puede estudiarse descomponiéndolo en subsistemas más pequeños y más simples, los cuales tienen su cinética débilmente acoplada entre sí. En el capítulo siguiente nosotros describiremos un patrón de catalizaciones que está presente en todas las Organizaciones y por ello, en combinación con este teorema, plantea un enfoque interesante.





**Figura 14:** Sistema simple ejemplificando la descomposición de una organización en distintos conjuntos. Las flechas indican reacciones: las colas inician en los reactantes y las puntas apuntan a los productos.  $c$  actúa como catalizador por lo que constituye el conjunto de catalizadores.  $x, y$  corresponden a un PAC pues si se alcanza un estado estacionario para estas moléculas, ninguna es sobreproducida ni podría serlo (la estequiometría no lo permite). De la misma forma  $w, z$  conforman otro PAC. Así, la auto-sustentación de  $x, y, c$  y la de  $w, z, c$  son independientes, y si se confirman ambas condiciones, puede asegurarse que toda la red es auto-sostenida.

## COT y Conjuntos RAF

Como hemos visto hasta ahora, los Conjuntos RAF contienen definiciones, algoritmos y teoremas interesantes sobre lo que consideran sistemas vivos (incluso por consideraciones de espacio he debido dejar fuera ideas recientes como los conjuntos irreducibles), pero deja los aspectos dinámicos fuera. Por otro lado COT es un formalismo que presenta mucho avance en lo que se refiere a traducir propiedades topológicas a consecuencias dinámicas, sin embargo su falta de foco en la catálisis lo hace débil a la hora de analizar metabolismos. Notando que ambos formalismos comparten muchas de las definiciones a nivel topológico, decidimos extender e investigar el nivel de solapamiento entre ambas teorías. Nuestro primer resultado en esta dirección es el siguiente teorema:

**Teorema 4** *Si un sistema de reacción catalítico es  $F$ -generado, entonces es una organización.*

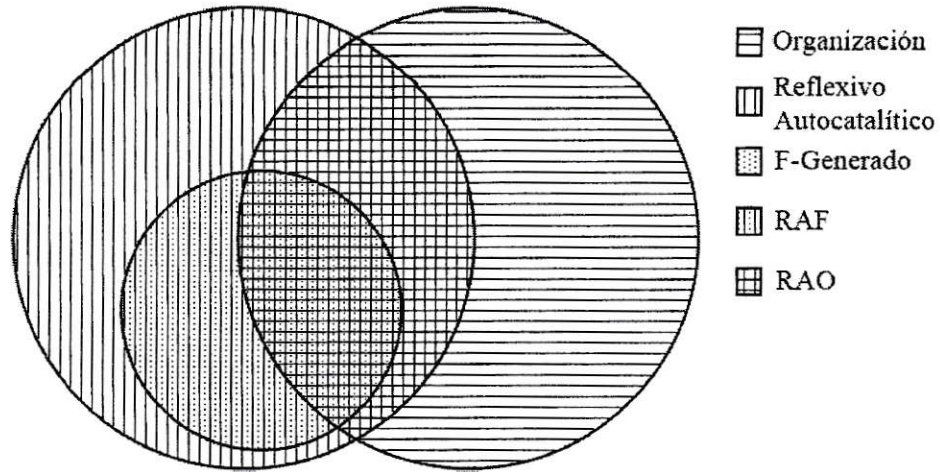
Este teorema, cuya demostración puede encontrarse en (Contreras y cols., 2011), hace una potente conexión entre los Conjuntos RAF y COT, pues indica que todas las herra-

mientas conceptuales desarrolladas para el análisis de las Organizaciones pueden aplicarse también a los sistemas RAF. Es justo preguntarse si el teorema valdrá en la dirección opuesta, es decir, si toda organización es un sistema RAF. La respuesta es esencialmente negativa, pues la condición de ser F-generado es más fuerte, al exigir que el sistema pueda construirse incrementalmente a partir del conjunto F y las reacciones que pueden gatillarse, mientras que las Organizaciones sólo requieren que el sistema, descrito a nivel topológico permita la existencia de vectores de flujo de forma que la masa se mantenga. Esta diferencia entre ambas teorías puede parcharse si se elige un conjunto F definido distinto a lo que sugeriría COT: correspondiente a las moléculas dadas por el ambiente. Estas ideas se exploran a mayor profundidad en (Contreras y cols., 2011), pero aquí reflexionaré sobre los aspectos más sobresalientes de este tópico:

Cualquier conjunto es trivialmente F-generado si elegimos el conjunto *food* suficientemente grande. En este sentido, sólo son interesantes los casos donde F es minimal (o si acaso existen múltiples conjuntos, quizás disjuntos, que puedan jugar el rol de *food*). Si nos remitimos a considerar dentro del conjunto *food* a aquellos elementos dados por el ambiente, la figura 15 muestra que solo algunos sistemas RA son F-generados; de la misma forma sólo algunas organizaciones son F-generadas.

Un segundo refinamiento que me gustaría introducir aquí como amalgama entre los Conjuntos RAF y las Organizaciones es la idea de sistema RAO, es decir un sistema reflexivo auto catalítico que además es una organización. Se ha decidido caracterizar estos sistemas pues poseen las propiedades que se esperarían encontrar en los organismos vivos hoy, es decir, sistemas que catalizan las reacciones que los componen como sistemas químicos y que además tienen la posibilidad de mantenerse materialmente en el tiempo. Esta clasificación tendrá frutos en el siguiente capítulo, donde desarrollaremos una consecuencia importante que trae consigo pertenecer a la clase de los sistemas RAO. En particular se prestará atención a los sistemas cRAO, sistemas RAO con clausura eficiente, es decir,

sistemas que producen todos sus catalizadores, pues estos capturan y definen a nivel técnico modelos como el de Rosen o el de Maturana y Varela. Así, el panorama de las distintas clases de sistemas puede verse en la figura 15.



**Figura 15:** Diagrama de Venn de las distintas clasificaciones de sistemas metabólicos realizadas en esta tesis. Puede verse como los RAO son sistemas reflexivos auto catalíticos que simultáneamente son organizaciones, sin embargo no siempre son *F*-generados, es decir, su origen no se encuentra restringido a un conjunto de moléculas iniciales.

## Hipótesis

La hipótesis central de esta tesis es que la clausura metabólica permite deducir teoremas acerca de la existencia y propiedades de estados estables en el metabolismo. Idealmente estos teoremas no debiesen depender de la estructura local de algunas pocas reacciones sino que de la organización circular y de propiedades sistémicas que derivan de la clausura. Ambiciosamente me gustaría poder demostrar, para una definición formal de la clausura metabólica y un uso biológico de la idea de robustez (e.g. resiliencia, estabilidad, adaptabilidad), la siguiente implicancia:

$$\text{Clausura Metabólica} \Rightarrow \text{Robustez}$$

En términos técnicos aterrizaré la idea de robustez como estabilidad del sistema dinámico que describe las trayectorias de las concentraciones de los metabolitos. Demostrar esta implicancia sería un avance importante para la teoría de los sistemas autopoieticos pues explicaría desde primeros principios como emerge, en los sistemas autopoieticos, la robustez dinámica; una de las propiedades más notorias.

## Objetivos

El objetivo general de esta tesis es generar un marco útil y unificado del metabolismo desde el cual se puedan derivar teoremas sobre los estados estacionarios que sean de interés para la biología experimental. Este marco, a diferencia de las teorías centrales sobre clausura, incorporará las nociones de tiempo y de cambio; además se tratará de mostrar bajo qué condiciones ocurren la estabilidad y multiestabilidad. Es decir, el formalismo a desarrollar analizará las consecuencias para la dinámica sistémica al ejecutar la clausura metabólica.

### **Los objetivos específicos son:**

1. Rescatar y **amalgamar** teorías circundantes (sistemas  $(M, R)$ , conjuntos RAF, COT, conjuntos auto catalíticos, etc...) a la idea de clausura, para destacar enfoques y resultados relevantes que puedan tener importancia para la teoría de los sistemas con clausura.
2. Usando a la vez conceptos provenientes de los conjuntos RAF y de COT, modelar un marco formal que permita describir los sistema con clausura en términos técnicos precisos y hacer alusión a su dinámica temporal.
3. Usando la precisión ganada en el punto anterior, obtener teoremas sobre conectividad que iluminen qué significa, desde el punto de vista técnico, tener clausura metabólica, tanto para la topología como para la dinámica sistémica.
4. Desde este resultado técnico, derivar un teorema sobre la estabilidad de los estados estacionarios utilizando como punto de partida la idea de clausura. Así se unirá el espacio existente entre teoría y experimento. Se pondrá especial atención en estudiar si es que el valor de las constantes cinéticas afecta las propiedades de los estados estacionarios.
5. Realizar simulaciones que corroboren la validez del resultado en sistemas simples biológicamente inspirados, con el fin de tener una apreciación más directa del significado de los teoremas obtenidos.

# Resultados

Los principales resultados de esta tesis se dividen en dos grupos:

- Caracterización de una nueva clase de organización/metabolismo (cRAO) que involucra aspectos topológicos y dinámicos. Con esta clasificación se aunan, bajo un mismo esquema, ideas provenientes de la Teoría de Organización Química (COT), Sistemas  $(M, R)$ , Autopoiesis y Conjuntos RAF. Si bien esta clasificación construye sobre trabajos anteriores, constituye una intersección que no ha sido estudiada con anterioridad.
- Teoremas sobre la existencia y estabilidad de un patrón de catálisis en los sistemas cRAO. En particular, se obtuvieron teoremas sobre la dinámica temporal y existencia de estados estables de los sistemas cRAO. Estos teoremas están basados en la noción de clausura metabólica y destacan la importancia de esta propiedad a través de consecuencias matemáticas conceptualmente simples que podrían verificarse experimentalmente. Además, estos teoremas corresponden a un novedoso y prometedor punto de partida que considero valioso en el estudio de los sistemas con clausura.

## **Organización de lo vivo: una descripción precisa y técnica**

Como se ha hecho explícito en esta tesis, el tema de la organización de lo vivo es un aspecto fundamental en toda la discusión sobre el origen, la estabilidad y el cambio de las redes metabólicas. Todos los enfoques analizados afirman tener ese secreto resuelto, pero muchas veces hay poco rigor técnico en sus argumentos. Además, hay interrogantes relevantes que pocas personas han atacado; por ejemplo, cuál es la relación entre los distintos tipos de sistemas? Como se ha cultivado la nula interacción entre los subcampos, pareciese

ser que los distintos tipos de sistemas viven en universos disjuntos... por ejemplo: ¿ Son las organizaciones *à la* COT sistemas (M,R)?

En este trabajo intersecté dos formalismos potentes –los Conjuntos RAF y la Teoría de la Organización Química – para generar así los sistemas RAO, es decir, sistemas reflexivos auto catalíticos que además son organizaciones. Esta decisión proviene del hecho de que ambos formalismos nos parecen interesantes pero al mismo tiempo insuficientes pues no capturan en su totalidad aspectos del metabolismo. Por un lado COT falla en el aspecto de incluir a las enzimas o catalizadores, por lo que es difícil obtener deducciones de la inclusión del metabolismo como producto mismo del sistema. Dicho de otra forma, COT no maneja el concepto de clausura bajo causas eficientes. Los conjuntos RAF por su lado ignoran completamente la dinámica temporal y estequiometría de los sistemas. Es por éstas falencias que planteo los RAO, un subconjunto de sistemas que vive en la intersección de ambos formalismos, y que, por lo tanto exige y complementa el análisis con características de ambos. En particular me referiré a los cRAO, sistemas RAO que tienen clausura eficiente.

Resumiendo, la clasificación RAO corresponde a sistemas:

- **Reflexivos Auto catalíticos** Todas las reacciones son catalizadas por moléculas del sistema (F-generadas o generadas por el sistema)
- **Cerrados** El conjunto de tipos de moléculas que componen al sistema corresponden exactamente al tipo de moléculas que son generados por el sistema.
- **Autosustentados** Existen tasas de reacción que permiten que el sistema como un todo se mantenga o crezca en términos cuantitativos.

y los cRAO además tienen la siguiente propiedad:

- **Clausura eficiente** Todos los catalizadores son productos de reacciones del sistema (a diferencia de ser F-generadas).

Como puede verse al comparar estos puntos con lo dicho en el capítulo 2, ser RAO es una condición similar a RAF, pero no impone restricciones sobre el origen (un sistema RAF debe ser capaz de construirse a partir de un conjunto  $F$ ) sino que sobre su viabilidad dinámica. A pesar de esto existe una estrecha relación entre los sistemas RAF y los RAO que fue mencionada en el capítulo anterior y es explorada con detalle en el apéndice (Contreras y cols., 2011).

Haré hincapie en los cRAOs, no por capricho o deuda histórica con Rosen, sino porque la clausura eficiente es una propiedad sumamente necesaria para la demostración de los siguientes teoremas aunque acarrea consigo el precio de tener que abandonar la discusión sobre el origen de la vida. Esto último ya que posiblemente la vida en sus inicios no generaba sus propios catalizadores sino que los capturaba o encontraba en el medio en el que se desenvolvía.

Esta clasificación (los cRAO) es un paso crucial hacia los verdaderos resultados de esta tesis, pues como es común en matemáticas, mientras más particular el sistema al cual uno se refiere, más propiedades sobre estos sistemas uno puede demostrar. Por supuesto nuestra intención es que la caracterización de nuestros sistemas se mantenga lo suficientemente general como para abarcar una gran extensión del fenómeno de la vida y al mismo tiempo caracterizar este fenómeno de forma que sea posible decir algo más sobre estos sistemas.

## **Teoremas de Clausura**

El resultado más importante de esta tesis es un teorema sobre la estabilidad de los sistemas con clausura. En particular, el teorema indica que todos los sistemas que dada su organización pueden ser clasificados como autopoieticos contienen subgrafos de reacción (denominados lazos de catalizaciones) que en términos dinámicos generan atractores estables. Es más, estos lazos están involucrados directa o indirectamente en la dinámica del



resto del sistema, por lo que sus propiedades son centrales a la hora de considerar el sistema completo. A continuación veremos los pasos y supuestos necesarios para demostrar este teorema a modo de comprender de forma más acabada su alcance.

## Control entre Moléculas

Para desarrollar este primer teorema debemos introducir una nomenclatura adecuada para describir una serie de relaciones que pueden darse entre catalizadores. La primera de estas es una relación binaria entre dos moléculas que se refiere a cuando una molécula ( $A$ ) cataliza una reacción que contiene entre sus productos a otra molécula ( $B$ ). Diremos en este caso que  $A$  *controla* a  $B$ . Para representar esta relación utilizaremos la siguiente notación

$$A \vdash B \quad (44)$$

indicando que la molécula  $A$  cataliza una reacción que incluye a  $B$  como producto. Esto se puede reescribir en términos del formalismo RAF (ver sección RAF en Capítulo 2) como:

$$A \vdash B \equiv \exists(A, r) \in C | B \in \pi(r) \quad (45)$$

Esta es una relación relevante pues implica que el estado o concentración de  $A$  influye directamente sobre el estado o concentración de  $B$ , es decir, las dinámicas de ambas moléculas están directamente acopladas.

A continuación definiremos una *cadena de catalizaciones* como una serie de moléculas en las cuales una controla a la siguiente, la cual controla a la siguiente, etc... como puede verse en la figura 16 A. Para esta situación anotaremos, siguiendo lo dicho anteriormente:

$$A \vdash B \vdash C \vdash \dots \vdash D \quad (46)$$

Esta relación resulta natural ya que si  $A$  influye en la dinámica de  $B$  y esta a su vez sobre  $C$ , por transitividad  $A$  influye sobre la dinámica de  $C$ . Aquí es válido distinguir que existen dos tipos de control, el primero que fue introducido anteriormente, el cual llamaremos control directo y el segundo, que corresponde al control indirecto y se refiere a que la molécula  $A$  controla indirectamente a  $C$  si y sólo si  $C$  es controlada directamente por  $A$  o por una molécula  $B$  que es controlada indirectamente por  $A$ . Como puede verse en esta definición recursiva, una molécula controla a otra indirectamente si es que existe una cadena de catalizaciones que las incluya a ambas, con la primera antecediendo a la segunda. Esta relación de control indirecto es interesante pues posee la propiedad de transitividad que lo cual tiene consecuencias desde el punto de vista teórico. Para referirnos a la situación de control indirecto utilizaremos la siguiente notación:

$$A \models D \tag{47}$$

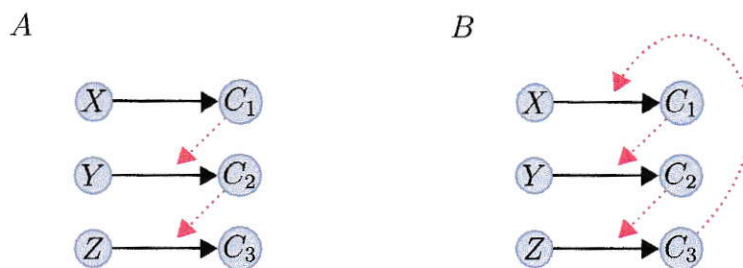
indicando que  $A$  controla indirectamente a  $D$ .

De la definición es evidente que

$$A \vdash B \Rightarrow A \models B \tag{48}$$

En este punto es necesario hacer notar que esta idea de cadena es distinta de la cadena lineal de reacciones estudiada por MCA pues involucra a los catalizadores como elementos articuladores de la cadena incorporando así las nociones de clausura.

La última definición necesaria para demostrar el primer teorema es la de *lazo de catalizaciones*, que corresponde a una cadena de catalizaciones circular como puede verse en la figura 16 B. Dicho de otra forma, un lazo de catalizaciones es una serie de moléculas en la cual la primera molécula se controla indirectamente sí misma a través del resto de moléculas del lazo. Esto es verdadero para todas las moléculas que componen el



**Figura 16:** Diagrama de dos estructuras de control distintas. **A** Una cadena de catalizaciones en la cual no hay un ciclo de control. **B** Un lazo de catalizaciones en el cual el último elemento controla al primero formando un ciclo.

lazo. La autocatálisis constituye un ejemplo trivial de este tipo de lazos. Existen distintas posibilidades para anotar un lazo de catalizaciones, como por ejemplo

$$A \vdash A \quad (49)$$

o

$$A \vDash A \quad (50)$$

o incluso siendo más explícitos:

$$A \vDash B \vDash C \vDash D \vDash A \quad (51)$$

Utilizaré esta última pues considero que es la que comunica mejor la situación y provee la información necesaria para caracterizar el lazo al cual uno se refiere (podría ocurrir que una molécula se controle a sí misma a través de múltiples cadenas). Así evitaremos molestas confusiones con casos de autocatálisis.

Con esto, ya podemos dar inicio al primer teorema.

## Teorema de la Existencia del lazo

**Teorema 5** *En un sistema cRA debe existir al menos un lazo de catalizaciones.*

**Demostración** La premisa es que todos los catalizadores son generados por el sistema (condición de clausura eficiente), es decir, constituyen productos de reacciones. Dado que el sistema es reflexivo auto catalítico podemos indicar que todas las reacciones son catalizadas. Con esto hemos demostrado que todos los catalizadores son controlados por algún otro catalizador, o sea:

$$\forall c \in \mathbb{K} \exists z \in \mathbb{K} | z \vdash c \quad (52)$$

De esto podemos concluir que siempre es posible ir hacia atrás en una cadena de catalizaciones (aunque algunas veces exista más de una opción, es posible elegir una sin pérdida de generalidad) perteneciente a un sistema cRA. A continuación elijamos un catalizador  $c_n$  cualesquiera y construyamos hacia atrás la cadena de catalizaciones; ésta tendrá la siguiente estructura en el paso n-ésimo:

$$c_1 \vdash c_2 \vdash \dots \vdash c_{n-1} \vdash c_n \quad (53)$$

Por supuesto que aquí encontramos un símil al problema de *infinite regress* de Rosen, pues esta cadena no está completa, es decir, siempre será posible interrogar al sistema para agregar un catalizador más que controle directamente a  $c_1$ . La base de la demostración es que considerando que el número de catalizadores es finito (y sin ignorar el hecho de que algunos catalizadores pueden catalizar más de una reacción), a lo más pueden darse  $\#\mathbb{K}$  pasos hacia atrás sin repetir un catalizador. Ya que tenemos asegurado por las condiciones iniciales que siempre podemos dar pasos hacia atrás, en algún momento en la cadena deberá repetirse un catalizador (no necesariamente  $c_n$ ). Contaremos entonces con una cadena de catalizaciones que contiene a una sub-cadena circular, y es por lo tanto un lazo de

catalizaciones. ■

**Corolario** En un sistema cRAO debe existir al menos un lazo de catalizaciones.

Este es un teorema de existencia que además indica como detectar el lazo, aunque algoritmos de mayor eficiencia son comunes en teoría de grafos. El teorema no garantiza que exista un único lazo, ni que este sea no-trivial (una reacción auto catalítica), pero dada la complejidad actual de los sistemas vivos es poco probable que efectivamente este sea el caso.

Notemos entonces algunos detalles adicionales acerca de los lazos y la relación de control indirecto: Si el sistema no es cerrado eficientemente, podría ocurrir que no sea posible seguir la cadena de catalizaciones hacia atrás, es por esto que se ha introducido este requerimiento. Es decir, la condición de que todos los catalizadores sean generados por el sistema responde a una necesidad para dar fundamento sólido a la matemática de la demostración, sin embargo no creemos que corresponda a una exigencia fuerte dentro del mundo biológico, pues la clausura eficiente es el caso para la mayoría de sistemas vivos observados, al punto que según algunos autores es reconocidamente una de las propiedades fundamentales para que un sistema se denomine vivo.

Otra observación pertinente es que la condición de ser reflexivo auto catalítico obliga a que haya al menos tantos catalizadores como reacciones o que los catalizadores participen como tales en más de una reacción, hecho común y conocido en biología como *moonlightning* (e.g. la piruvato carboxilasa).

Por otro lado, la relación de control indirecto presenta propiedades novedosas bajo esta luz:

- Es su propia clausura transitiva, es decir, es la relación minimal que es cerrada sobre  $\mathbb{K}$ , transitiva y contiene a la relación de control indirecto

Esto abre la puerta al establecimiento de nuevas relaciones o representaciones del sistema utilizando la reducción transitiva, operación que conserva la relación de alcanzabilidad entre los distintos nodos de un grafo. Esta puede ser una forma interesante de reducir la complejidad del sistema en términos de sus relaciones internas, y al mismo tiempo mantener un invariante de modo de simplificar el análisis de dinámica. Además podría utilizarse esta reducción para analizar los distintos componentes de la dinámica, dando espacio a una primera (y algo grosera) descomposición. Otro caso en el que este tipo de relaciones sería interesante es si se lograra reemplazar las relaciones de control aquí planteadas por funciones, pues en ese caso podrían emplearse teorías más específicas como la teoría de alegorías.

Además, si se consideran sólo los catalizadores que participan en lazos, la relación obtiene estas propiedades:

- **Simétrica**, es decir, que  $A \models B \Leftrightarrow B \models A$
- **Reflexiva**, es decir se cumple que  $A \models A$

Por lo tanto, la relación de control indirecto se transforma en una relación de equivalencia entre los distintos elementos del lazo, donde todos los elementos de un mismo lazo pertenecen a una clase de equivalencia, mientras que elementos pertenecientes a lazos distintos corresponderían a clases de equivalencia distintas, a no ser que estos lazos estuviesen unidos formando un lazo adicional. Además la relación de control indirecto en este caso obtiene clausura simétrica y clausura reflexiva, lo cual abre la puerta a reducciones basadas también en estas propiedades. Nuevamente, estas reducciones son interesantes pues quizás permitirían simplificar el análisis del sistema. Además, aquí se puede empezar a plantear la red metabólica como una categoría (en particular, la relación de control indirecto se presta para una excelente comparación con la estudiada categoría de caminos).

Por último notaremos que los catalizadores que controlan indirectamente a elementos

de un lazo y que no forman parte de él, deben formar parte de otro lazo, el cual en su totalidad controla a los elementos del primer lazo. Esto permite establecer una jerarquía de lazos basado en la cinética, que si bien se limita a términos de catalizadores, valdría la pena explorar.

Si ahora nos enfocamos a la relación existente entre los elementos (catalizadores) que pertenecen a un lazo y aquellos que no, veremos que estos últimos no controlan indirectamente a elementos pertenecientes al lazo, mientras que en la dirección opuesta se da el caso contrario. Esto revela que la relación sería asimétrica, es decir,  $A \models B \Rightarrow B \not\models A$ . Una relación transitiva y asimétrica es un orden estricto. De esta forma podemos ordenar las moléculas y así formalizar la subordinación del estado de los elementos exteriores al lazo a la dinámica del lazo. Además, en términos de catalizadores, se indica que el lazo es suficiente para generar (en el sentido de conjuntos RAF) al resto del sistema. Esta observación nos lleva al siguiente teorema.

## **Teorema de la Hegemonía del Lazo**

Notemos que del teorema anterior se desprende un corolario prometedor, cuyo alcance aún no medimos. Llamamos a la siguiente proposición el teorema de la Hegemonía del Lazo.

**Teorema 6** *Todos los catalizadores de un cRA son controladas indirectamente por al menos un lazo.*

**Demostración** Dado que el sistema es reflexivo auto catalítico, todas las reacciones contenidas en él deben ser catalizadas. Elijamos un catalizador cualesquiera y realicemos el procedimiento utilizado en la demostración anterior, notando que efectivamente llegamos a algún lazo en algún número de pasos finito, es decir, el catalizador, sea cual sea, contiene a un lazo río arriba y por lo tanto está bajo control indirecto de todas las especies pertenecientes a ese lazo. ■

**Corolario** Todos los catalizadores de un cRAO son controladas indirectamente por al menos un lazo.

Una curiosa consecuencia de este teorema es que si el sistema esta constituido por reacciones más alla del lazo, catalizadores pertenecientes a los lazos deben catalizar algunas de estas reacciones. Esto lleva a la conclusión de que algunos catalizadores del lazo tendrían que catalizar más de una reacción.

La importancia de este teorema es que destaca el impacto que tiene la dinámica de los lazos para el comportamiento del resto del sistema, debido a que indica que todas las reacciones tienen su dinámica (indirectamente) subordinada a la dinámica de algún lazo. Dicho de otra forma, la dinámica de todos los catalizadores y por lo tanto del sistema en sí esta fuertemente acoplada y posiblemente dominada por los lazos. Por este motivo resultaría interesante conocer la dinámica de los lazos, los distintos regímenes en los que pueden entrar y como los afectan perturbaciones. Esto nos conduce directamente al tercer teorema.

## **Teorema de la Estabilidad del Lazo**

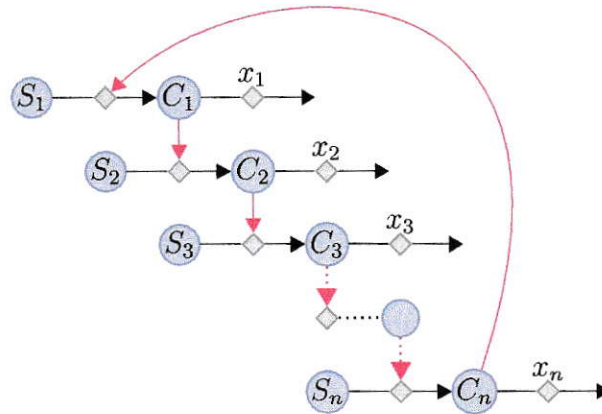
El teorema que demostraremos a continuación es la culminación de este trabajo de tesis, sin embargo lo consideramos sólo la punta del iceberg pues de seguro será posible extender estos resultados a otros tipos de sistemas y demostrar condiciones más fuertes sobre los sistemas cRAO. Este teorema se trata sobre la dinámica del lazo, en particular sobre el vecindario de sus equilibrios.

Como se acaba de enunciar, la dinámica del lazo influye en el resto del sistema y no así en sentido contrario (al menos en lo que respecta a los catalizadores). Consideraremos entonces un sistema constituido por un único lazo para simplificar el análisis aunque los resultados son ampliables a la existencia de más lazos y otras reacciones. Este tipo de



simplificaciones son cruciales puesto que las topologías de las redes de reacción de los sistemas vivos pueden ser arbitrariamente complejas. Veamos entonces que se puede decir sobre las propiedades dinámicas del lazo de catalizaciones.

Estudiaremos un sistema con la topología de la figura 17 sistema, en el cual tenemos  $N$  catalizadores ( $C_i$ ), los cuales sufren degradación de forma proporcional a su concentración. Básicamente esto indica que como la mayoría de las moléculas biológicas, tienen una estabilidad dada y que en promedio, después de una vida media que depende de su estabilidad, la mitad de los catalizadores se habrán degradado. Estas moléculas catalizan la reacción de formación del siguiente catalizador a partir de un sustrato  $S_i$ , el cual se mantiene constante. Esto forma entonces un único lazo de catalizaciones de largo  $N$ .



**Figura 17:** Sistema constituido únicamente por un lazo de catalizaciones. Los catalizadores  $C_i$  se degradan o decaen de forma proporcional a su concentración y los sustratos  $S_i$  se mantienen constantes.

Consideremos que la dinámica de cada una de las reacciones involucradas en la red de catalizaciones sigue Ley de Acción de Masas descrita en el capítulo 3. Dado que en este caso consideramos a los catalizadores como parte misma del sistema (nuestra diferencia principal con la mayoría de los análisis dinámicos en biología teórica) es necesario realizar una modificación a la fórmula desarrollada anteriormente. En particular, plantearemos la reacción como si correspondiese a una reacción trimolecular, por lo que las concentracio-

nes  $C_i$  podrían describirse en el tiempo de acuerdo a la siguiente ecuación diferencial:

$$\frac{\partial C_i}{\partial t} = k_i S_i C_{i-1} - \bar{k}_i C_i X_i \quad (54)$$

Por lo tanto, la velocidad de las reacciones depende linealmente de la concentración de sus catalizadores. Además hemos obviado la reversibilidad de las reacciones pues la concentración de las moléculas al borde del sistema (acervos moleculares involucrados en la parte reversible de las reacciones) se mantienen constantes, e incorporarlas no sería mayor complicación que la de modificar  $\bar{k}_i$ . Realizaremos ahora un cambio de notación para hacer más explícito este hecho, reemplazaremos:

$$\begin{aligned} a_i &= \bar{k}_i X_i \\ b_i &= k_i S_i \end{aligned} \quad (55)$$

Bajo esta notación, tanto  $a_i$  como  $b_i$  son positivos y el sistema queda escrito entonces:

$$\frac{\partial C_i}{\partial t} = b_i C_{i-1} - a_i C_i \quad (56)$$

Como es usual en el análisis de sistemas metabólicos, escribiremos este conjunto de ecuaciones en términos de vectores y matrices, en particular el vector de concentraciones de los catalizadores  $C$  y la matriz  $J$  que contiene los coeficientes de las ecuaciones diferenciales y corresponde al Jacobiano del sistema:

$$\frac{\partial C}{\partial t} = JC \quad (57)$$

con la matriz  $J$  siendo:

$$J = \begin{bmatrix} -a_1 & 0 & 0 & \dots & 0 & b_1 \\ b_2 & -a_2 & 0 & \dots & 0 & 0 \\ 0 & b_3 & -a_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -a_{N-1} & 0 \\ 0 & 0 & 0 & \dots & b_N & -a_N \end{bmatrix} \quad (58)$$

## Valores y Vectores Propios

En este momento invertiremos algo de tiempo en recordar y entender algunas propiedades de los valores y vectores propios, pues las secciones que siguen hacen uso extenso de esta materia. Los valores propios de una matriz  $A$  son números (posiblemente complejos) que tienen vectores propios  $v_i$  asociados tales que se cumple la siguiente ecuación:

$$Av_i = \lambda_i v_i \quad (59)$$

que también se suele escribir como

$$Av_i - \lambda_i v_i = 0 \quad (60)$$

De aquí se obtiene una importantísima propiedad que indica que si  $\lambda$  es un valor propio de  $A$  entonces

$$(A - \lambda I)v = 0 \quad (61)$$

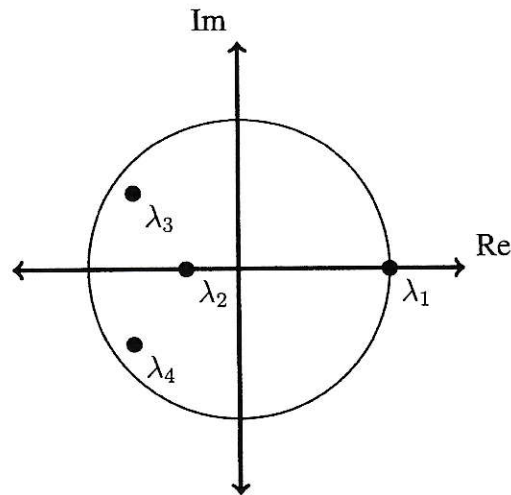
y por lo tanto

$$\det(A - \lambda I) = 0 \quad (62)$$

Para una matriz  $A$  cualesquiera, se denomina  $\sigma(A)$  al espectro de  $A$ , es decir, al conjunto de sus valores propios  $\lambda_i$ . Para ayudar a la comprensión graficaremos los valores propios en el plano de Argand o plano complejo, como se puede ver en la figura 18. Es posible que exista un valor propio cuya magnitud sea superior a la de los demás valores propios, en cuyo caso llamamos a este el valor propio dominante y hacemos referencia a su magnitud como el radio espectral, anotándolo  $\rho(A)$ :

$$\rho(A) = \max_i |\lambda_i| \quad (63)$$

En la figura 18 puede verse que el  $\rho(A)$  corresponde al radio del círculo más pequeño que incluye en su interior a todos los valores propios de  $A$ .  $\lambda_1$  es el valor propio dominante.



**Figura 18:** Plano de Argand con algunos valores propios reales y complejos dibujados. Visualizar los valores propios en este plano permite entender sus movimientos y comprender a mayor profundidad lo que sucede con el sistema bajo transformaciones que afectan a los valores propios.

Prestaremos especial atención a dos elementos importantes del álgebra lineal. El primero es el polinomio característico, un polinomio asociado a matrices cuadradas que contiene información relevante acerca de los valores propios de la matriz, su determinante y

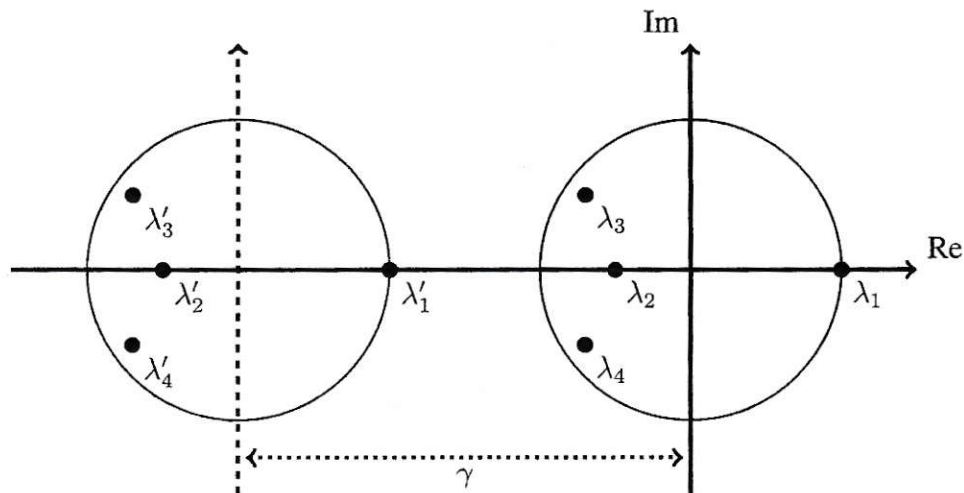
su traza. En particular el polinomio característico para la matriz  $A$  se escribe

$$p_A(\lambda) = \det(A - \lambda I) \quad (64)$$

Con esto queda claro que las raíces de  $p_A$  corresponden a los valores propios de  $A$ , por lo que se puede desprender información relevante de estos últimos a partir de  $p_A$ , e.g., si un valor propio de una matriz real tiene parte imaginaria, entonces su conjugado también es valor propio. La segunda consecuencia de la ecuación 62 es el teorema del corrimiento espectral, que indica que sumar una matriz diagonal  $\gamma I$  a  $A$  *corre* los valores propios de  $A$  en  $\gamma$  pues se puede ver que para la nueva matriz  $A + \gamma I$  se tiene:

$$\det((A + \gamma I) - \lambda I) = \det(A - (\lambda - \gamma)I) = 0 \quad (65)$$

donde  $\lambda$  corresponde a los valores propios de la matriz sumada y  $(\lambda - \gamma)$  corresponde a los valores propios de la matriz original. Por ello puede verse que la relación entre los valores propios de la matriz nueva pueden obtenerse sumando  $\gamma$  a los valores propios de la matriz antes de la suma. Esta relación se ilustra en la figura 19.



**Figura 19:** Teorema del corrimiento espectral. Si  $\lambda_i$  corresponden a los propiovalores de una matriz  $A$ , entonces los propiovalores  $\lambda'$  de  $A' = A - \gamma I$  se disponen de igual forma pero *corridos* en  $\gamma$ .

Finalmente en esta sección quiero dar una interpretación de lo que representan los valores y vectores propios para el sistema biológico en cuestión. Los valores y vectores propios de una matriz están asociados a un cambio de base o de variables, de modo que el estado del sistema podría expresarse no como una serie de concentraciones de metabolitos, sino que como una serie de coordenadas generalizadas compuestas cada una por combinaciones lineales de las concentraciones de los metabolitos. Los vectores propios indican como debe componerse cada una de estas coordenadas generalizadas en función de las concentraciones de metabolitos. Entonces, la evolución temporal del sistema puede comprenderse como cambios en estas coordenadas más naturales del sistema; donde los valores propios indican si las distintas coordenadas decaen, se mantienen invariantes o crecen exponencialmente.

## Dinámica del Lazo

Habiendo terminado el pequeño interludio algebraico, podemos proseguir con el análisis de nuestro sistema. A este sistema le pediremos tan sólo una condición adicional, y es que tenga al menos un estado estacionario distinto del trivial. Es decir, que exista un vector de concentraciones  $C^*$  tal que:

$$\frac{\partial C}{\partial t} = JC^* = 0, C^* \neq 0 \quad (66)$$

Esta condición implica directamente la existencia de un vector propio con valor propio nulo, o dicho de otra forma, un *kernel* no nulo, lo que significa que existe un subespacio de dimensión al menos 1, en el cual el sistema se encuentra en equilibrio. Nos detendremos ahora para estudiar el significado de esta condición para poder justificar más acabadamente su existencia y posteriormente nos concentraremos en las propiedades de estos equilibrios.

Notemos en primera instancia que como el sistema es lineal, si no existe un valor

propio nulo, no existen otros equilibrios más que el trivial. Además los *bordes* del sistema (cuando  $C_i \approx 0$ ) son repulsores ya que cuando el sistema se encuentra en su vecindario tiende a alejarse. Esto se debe a que la ecuación diferencial que gobierna la concentración de las moléculas, cuando

$$C_i \approx 0 \quad (67)$$

entonces

$$\frac{\partial C_i}{\partial t} \approx b_i C_{i-1} \quad (68)$$

y como sabemos que  $b_i$  y  $C_{i-1}$  son magnitudes positivas, significa que la concentración de cualquier molécula, cuando es baja, tiende a subir. Por esto, es bastante claro que el sistema, si no cumple con la condición de un kernel no nulo, parece tener sólo dos destinos posibles: Crecer infinitamente o decaer hacia el equilibrio en el origen, es decir, existen tan sólo dos subespacios invariantes los cuales generan el espacio de fase, el sub-espacio estable  $E^s$  y el sub-espacio inestable  $E^u$ . Aquí debo al lector una explicación del sentido de los subespacios invariantes. Los subespacios invariantes son espacios vectoriales, o sea, conjuntos de puntos o posiciones en el espacio de fase, que son generados por los vectores propio del sistema y que por esto mismo las soluciones al sistema de ecuaciones diferenciales que inicien en estos espacios, permanecen dentro de estos espacios, de allí que se los llame invariantes. Este hecho es trivial ya que por los métodos que conocemos para la resolución de sistemas de ecuaciones diferenciales lineales homogéneas sabemos que las soluciones  $C(t)$  tienen la siguiente forma:

$$C_i(t) = \sum_{i=1}^N n_i e^{\lambda_i t} \quad (69)$$

donde  $n_i$  es una constante que mide la proyección de la condición inicial sobre el vector propio  $i$ -ésimo. En particular si la condición inicial corresponde al espacio generado por

sólo uno de los vectores propios, el sistema evolucionará exponencialmente siguiendo dentro de ese espacio y la tasa de crecimiento o decaimiento corresponderá al valor propio asociado.

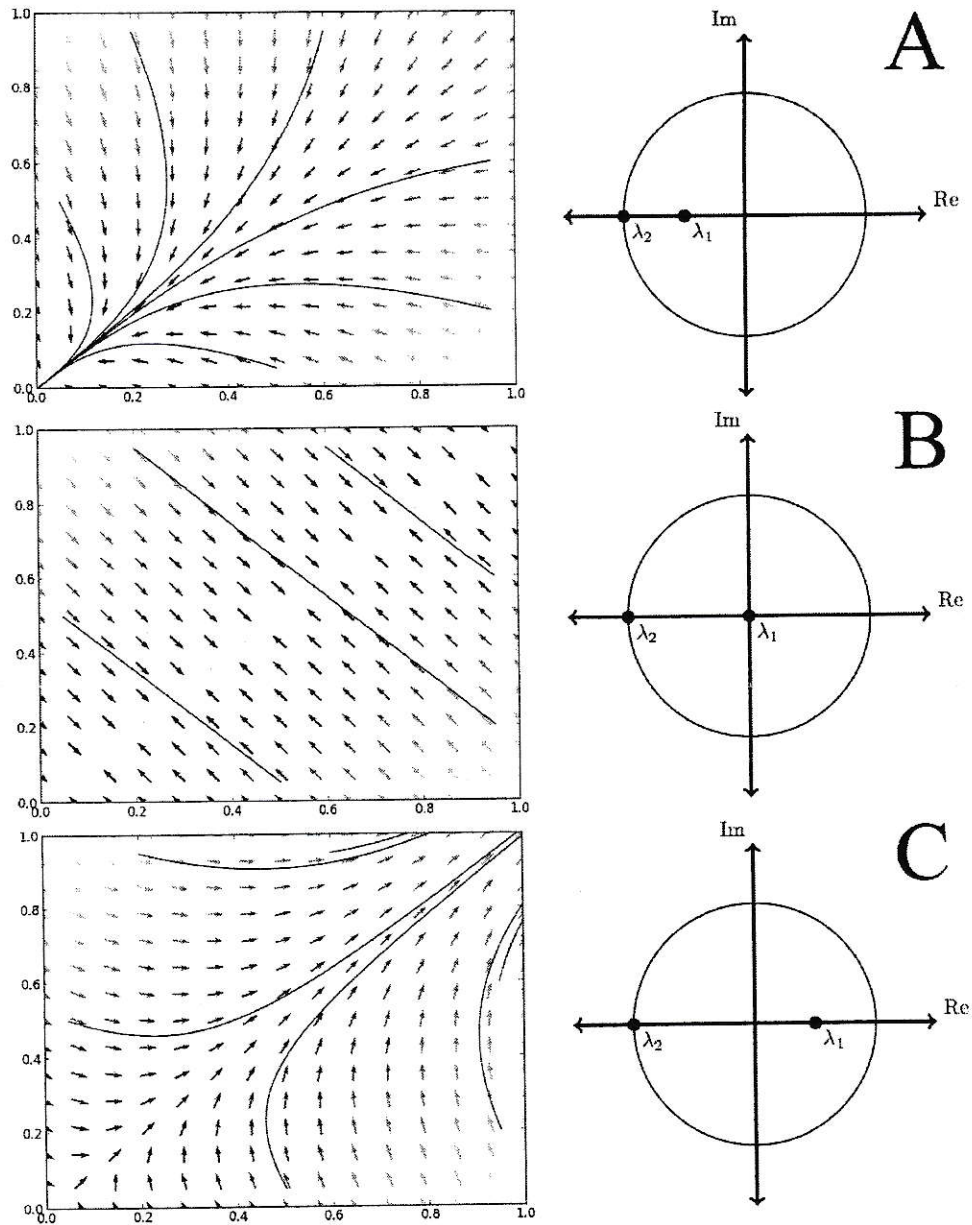
En general contaremos con posiciones iniciales que correspondan a una combinación lineal de los vectores propios del sistema, pero dado el principio de superposición con el que cumplen los sistemas lineales, es posible estudiar lo que sucederá con la componente estable  $C^s$  y la componente inestable  $C^u$  de una solución inicial. Las órbitas de estas componentes tendrán el siguiente comportamiento asintótico:

$$\begin{aligned}\lim_{t \rightarrow \infty} C_i^s(t) &= \lim_{t \rightarrow \infty} \sum n_i e^{\lambda_i^- t} = 0 \\ \lim_{t \rightarrow \infty} C_i^u(t) &= \lim_{t \rightarrow \infty} \sum n_i e^{\lambda_i^+ t} = \infty\end{aligned}\tag{70}$$

donde  $\lambda^-$  y  $\lambda^+$  denotan los valores propios negativos y positivos respectivamente. Así queda en evidencia que si el sistema ha de decaer hacia el origen, debe hacerlo a través de  $E^s$  ya que de otra forma el comportamiento asintótico de las soluciones obedece a los modos de crecimiento exponencial. Es decir, el decaimiento del sistema hacia el origen requiere la inexistencia de  $E^u$  o la existencia de una órbita muy particular en la cual todas las concentraciones disminuyen en la proporción exacta de manera que todas las moléculas se extingan simultáneamente. Una ilustración de los distintos factores y trayectorias posibles se puede observar en los campos vectoriales de la figura 20. Notemos que para sistemas como el que estamos analizando, la traza siempre es negativa por lo que asegura la existencia de al menos un valor propio negativo. En el caso particular en la figura 20, ya que se trata de un sistema de dimensión 2 (y dada la forma de la matriz), no hay espacio para valores propios complejos, por lo que existen tan sólo 3 combinaciones distintas para los signos de los valores propios. Cabe destacar que el subespacio  $E^s$  generado por sólo un valor propio no se aprecia pues el vector propio que le corresponde existe en regiones del sistema que no tienen sentido, en particular en los cuadrantes II y IV, por lo que aunque



tengamos asegurada matemáticamente la existencia de estos subespacios, esto no implica que sean posibilidades para el decaimiento efectivo del sistema.



**Figura 20:** Distintos regímenes dinámicos para un sistema bidimensional que cumple con la forma de nuestra matriz. También se ilustran las combinaciones de valores propios que los generan. Se han marcado 6 trayectorias en cada diagrama para ayudar a la visualización. **A:** Ambos valores propios son negativos, por lo que el origen es un atractor estable. **B:** Un valor propio nulo y uno negativo, por lo que existe una recta de puntos de equilibrio estables. **C:** Un valor propio es negativo y el otro es positivo, por lo que el sistema es inestable y crece exponencialmente.

## Estados estacionarios

Consideremos ahora que el sistema sí cumple con la condición impuesta por lo que tenemos un valor propio nulo, y recordando la conocidísima fórmula para el determinante de una matriz, tenemos que

$$\det(J) = \prod_{i=1}^N \lambda_i = 0 \quad (71)$$

Por otro lado, y considerando la elegante forma de esta matriz, es posible hacer un arreglo inteligente de sus elementos para así escribir el determinante en términos de los  $a_i$  y  $b_i$  como sigue:

$$\det(J) = \prod_{i=1}^N a_i - \prod_{i=1}^N b_i = 0 \quad (72)$$

con lo que concluimos que si pedimos la existencia de estados estacionarios no triviales, realmente estamos exigiendo que las medias geométricas de los  $a_i$  y los  $b_i$  coincidan:

$$\prod_{i=1}^N a_i = \prod_{i=1}^N b_i \quad (73)$$

Estas constantes pueden entenderse como tasas de decaimiento o generación medias. En particular si observamos sólo el componente de decaimiento veremos que para cualquier molécula la ecuación diferencial queda como

$$\frac{\partial C_i}{\partial t} = -a_i C_i \quad (74)$$

cuya solución es el conocido decaimiento exponencial

$$C_i = A e^{-a_i t} \quad (75)$$

permitiendo así realizar una simple interpretación de las constantes  $a_i$ . Las constantes  $b_i$  pueden comprenderse de forma similar, pero esta vez teniendo en consideración que el

elemento que genera el crecimiento exponencial, no es sí mismo, sino que un catalizador distinto, por lo tanto  $b_i$  tiene el valor de ser una constante de acoplamiento que indica cuanto aumenta el influjo o la generación de una especie con respecto al estado de otra.

En vista de lo anterior, la condición del kernel no nulo expresada en términos de las constantes de acoplamiento parece indicar que el decaimiento del sistema debe balancearse o contrarrestarse exactamente con los influjos causados por el sistema mismo. Notemos que la condición es tan sólo sobre el producto de todas las constantes, esta condición no exige que el flujo de cada reacción se encuentre balanceado pero sí se le pide al sistema un balance generalizado. Veamos entonces porque habría de darse este balance. Para simplificar el argumento consideraremos que el decaimiento es independiente de cualquier otro evento en nuestro universo, aunque no resulta muy complicado extendernos al caso en el que esto no fuese así. Consideremos inicialmente que el influjo es mayor al decaimiento, por lo tanto:

$$\prod_{i=1}^N a_i < \prod_{i=1}^N b_i \quad (76)$$

En este caso el sistema estará creciendo, es decir, aumentando las concentraciones o masa de las distintas moléculas. Con ello aumentará la tasa a la cual consume los sustratos  $S_i$ , los cuales hasta ahora hemos considerado constantes. Dado que ningún acervo molecular es verdaderamente infinito, el sistema podrá crecer sólo hasta que haga una merma efectiva sobre los sustratos, momento en el cual  $S_i$  tendría que empezar a disminuir, bajando así el valor de las constantes  $a_i$  hasta equilibrar los flujos.

Consideremos ahora el caso contrario, cuando se tiene

$$\prod_{i=1}^N a_i > \prod_{i=1}^N b_i \quad (77)$$

En este caso el sistema tiene la muerte asegurada, ya que no hay suficientes sustratos como para que alcance un estado estacionario distinto del origen. En el caso de los seres

vivos existen sistemas auxiliares que “movilizarían” al organismo hasta zonas más adecuadas. Estos distintos regímenes pueden entenderse revisando nuevamente la figura 20, donde los diagramas de fase muestran condiciones como las mencionadas aquí. En definitiva el futuro del sistema dependerá de los valores propios como se vio anteriormente, y demostrar que las ecuaciones 77 y 76 implican la existencia de valores propios positivos o su ausencia sería un paso interesante pero no importante en el estudio de estos sistemas. De todas formas se puede esbozar una demostración considerando el teorema del corrimiento espectral: Si uno se encuentra en la condición en la que existe un valor propio nulo, como veremos a continuación, esto implica que éste es el valor ubicado “más a la derecha”. Una disminución de los valores de  $a_i$  corre los valores propios hacia la izquierda, por lo que pasarían a ser todos negativos, llevando al sistema a la muerte. Por otro lado, un pequeño corrimiento de los  $a_i$  hacia valores más positivos corre el espectro del sistema hacia la derecha, generando inmediatamente un valor propio con parte real positiva. Si bien esta no es una prueba rigurosa es un argumento suficientemente fuerte para nuestros propósitos.

Hemos visto entonces como las constantes juegan un rol en la determinación del tamaño relativo de los distintos subespacios y como parece cuerdo que en general los sistemas vivos se encuentren con condiciones en la ecuación 73 es efectivamente verdadera, por lo que resulta interesante investigar ahora las propiedades de los equilibrios en el kernel de  $J$ .

## **Estabilidad de los Equilibrios**

La pregunta principal sobre estos estados estacionarios no triviales es si acaso son estables, es decir, si acaso puntos arbitrariamente cercanos a él (causados por perturbaciones arbitrariamente pequeñas) siguen órbitas que vuelven hacia el equilibrio. En general esta pregunta se resuelve examinando los valores propios de  $J$  ya que si todos estos tienen parte real negativa entonces el sistema es estable. Para iniciar nuestra investigación sobre los

valores propios del sistema veremos el polinomio característico, un polinomio cuyas raíces corresponden a los valores propios del sistema como se dijo anteriormente. En nuestro caso este polinomio  $p_J$  puede escribirse como:

$$p_J(\lambda) = \det(J - \lambda I) \quad (78)$$

Esta nueva matriz queda de la siguiente forma:

$$J - \lambda I = \begin{bmatrix} -(a_1 + \lambda) & 0 & \dots & 0 & b_1 \\ b_2 & -(a_2 + \lambda) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -(a_{N-1} + \lambda) & 0 \\ 0 & 0 & \dots & b_N & -(a_N + \lambda) \end{bmatrix} \quad (79)$$

Utilizando el mismo ordenamiento que vimos en la ecuación 72, obtenemos

$$p_J(\lambda) = \prod_{i=1}^N (\lambda + a_i) - \prod_{i=1}^N b_i \quad (80)$$

y justamente por la ecuación 73 podemos escribir

$$p_J(\lambda) = \prod_{i=1}^N (\lambda + a_i) - \prod_{i=1}^N a_i \quad (81)$$

$p_J$  tiene una forma especial, pues corresponde a la productoria de términos con la forma  $(\lambda + a_i)$  menos una constante. Esta constante resulta ser justamente el término libre del polinomio que resultaría tan solo de la productoria mencionada anteriormente. Este otro polinomio tiene por raíces exactamente a  $-a_i$ , así que si pudiese encontrarse una conexión entre ambos polinomios sería provechoso. Otro hecho notable acerca de  $p_J$  es que, por lo dicho anteriormente, todos los términos no libres corresponden a la suma de los polino-

mios simétricos de los elementos  $a_i$ , por lo que los coeficientes del polinomio son reales y positivos, particularidad que en combinación con la regla de los signos de Descartes resultará muy provechosa a continuación. Una última nota que debe mencionarse es que al eliminar el término libre de cualquier polinomio es posible factorizar por la variable (eliminando la raíz 0), así disminuimos en uno el grado del polinomio y nos damos cuenta de que 0 es una solución simple, es decir, de multiplicidad 1.

Examinaremos ahora las raíces del polinomio  $p_J$  pues éstas constituyen los valores propios de nuestro sistema:

$$p_J(\lambda) = 0 = \prod_{i=1}^N (\lambda + a_i) - \prod_{i=1}^N a_i \quad (82)$$

y por lo tanto

$$\prod_{i=1}^N (\lambda + a_i) - \prod_{i=1}^N a_i = 0 \quad (83)$$

Un punto importante a recordar aquí es que los valores propios  $\lambda$  son números complejos, es decir, pueden tener una parte imaginaria no nula. De otra forma sería posible, utilizando la regla de los signos de Descartes, demostrar instantáneamente que no existen valores propios positivos, es decir, contamos con el valor propio nulo que pedimos al inicio y el resto serían valores propios negativos. Más adelante veremos como se puede mostrar que en realidad los valores propios efectivamente tienen una parte imaginaria y por lo tanto este razonamiento no nos es de utilidad.

A partir de la ecuación 83 lo que se quiere demostrar es que las partes reales de  $N - 1$  raíces del polinomio  $p_J$  son negativas, pues sabemos que una de ellas es 0. Esta conjetura puede demostrarse de al menos 3 formas distintas. Nosotros realizaremos dos demostraciones pues una resulta increíblemente simple mientras que la otra es esclarecedora en algunos aspectos, por lo que consideramos que ambas son valiosas, pero antes enunciaremos el teorema que pretendemos demostrar.

**Teorema 7** *Si la dinámica de un lazo tiene estados estacionarios no triviales, estos son atractores.*

Tenemos en cuenta que estamos tratando por ahora con la dinámica de un lazo aislado, sin embargo indicaremos posteriormente como el resto del sistema podría incorporarse en la dinámica y como los resultados obtenidos aquí tienen sentido para sistemas cRAO de complejidad arbitraria.

**Demostración** Empezaremos con la ecuación 83 reescrita de otra forma:

$$\prod_{i=1}^N (\lambda + a_i) = \prod_{i=1}^N a_i \quad (84)$$

A continuación tomaremos la norma de ambos lados de la ecuación

$$\left| \prod_{i=1}^N (\lambda + a_i) \right| = \left| \prod_{i=1}^N a_i \right| \quad (85)$$

y como se sabe que la norma es multiplicativa en  $\mathbb{C}$  podemos escribir

$$\prod_{i=1}^N |(\lambda + a_i)| = \prod_{i=1}^N |a_i| \quad (86)$$

Para el lado derecho de la ecuación podemos eliminar la norma pues  $a_i = |a_i|$ . Para el lado izquierdo reescribiremos de acuerdo a la definición de norma para números complejos:

$$\prod_{i=1}^N \sqrt{\Re(\lambda + a_i)^2 + \Im(\lambda + a_i)^2} = \prod_{i=1}^N \sqrt{(\Re(\lambda) + a_i)^2 + \Im(\lambda)^2} = \prod_{i=1}^N a_i \quad (87)$$

Como mencionamos anteriormente, el caso en el que  $\lambda$  es real puede demostrarse fácilmente utilizando la regla de los signos de Descartes, por lo que consideremos aquí el caso en el que  $\lambda$  efectivamente tiene una parte imaginaria. En este caso  $\Im(\lambda)^2 > 0$  por

lo que la única forma de que la ecuación 87 se siga manteniendo es que  $(\Re(\lambda) + a_i)^2 < (\Re(a_i))^2$ , lo cual sólo ocurrirá si  $\Re(\lambda) < 0$ . La única excepción es cuando  $\lambda$  es real, caso en el cual debe ser 0. ■

Ésta es una prueba rápida que sirve a su propósito: demostrar que la conjetura es cierta y efectivamente los equilibrios del sistema son estables con excepción de una dirección, en la cual se encuentran infinitos equilibrios. Esta dirección corresponde al vector propio cuyo valor propio es nulo. Ahora veamos otra demostración que si bien puede resultar más complicada, esclarece algunos aspectos a mayor profundidad, es decir, más allá de las operaciones algebraicas. Para esta segunda demostración nos permitiremos otro interludio matemático.

## Matrices No-Negativas y de Metzler

Para realizar la segunda demostración del teorema de estabilidad del lazo necesitaremos conocer las matrices no-negativas y su relación con las matrices de Metzler. Las primeras son definidas como matrices que contienen elementos reales no-negativos mientras que las segundas, llamadas así por el economista norteamericano Lloyd Metzler que las estudio inicialmente, tienen la característica de que todos sus elementos no pertenecientes a la diagonal principal son no-negativos<sup>1</sup> y al menos uno es positivo (para evitar el caso trivial de la matriz 0). (Se recomienda revisar las sorprendentes propiedades de estas matrices en (Mitkowski, 2008) y más recientemente (Benvenuti y Farina, 2004)). Es notorio que las matrices no-negativas son casos particulares de matrices de Metzler, ya que cuentan con restricciones adicionales sobre los elementos de la diagonal. Además, podemos transformar cualquier matriz no-negativas en una matriz de Metzler restando a las primera la matriz diagonal  $\gamma I$ , lo cual, como se vio en el interludio matemático anterior, tiene el efecto de correr el espectro en  $-\gamma$ . Por supuesto, podemos elegir un  $\gamma$  suficientemente grande

---

<sup>1</sup>Esta propiedad hace que los bordes del sistema sean repulsores, como se argumento anteriormente.



de modo que todos los elementos de la diagonal se tornen negativos.

Para las matrices no-negativas existe un teorema que asegura la existencia de un valor propio real no-negativo conocido como el valor propio de Perron-Frobenius, cuya magnitud es máxima, es decir, corresponde al radio espectral. Si bien su multiplicidad puede no ser 1 y este valor puede no ser el único con esa magnitud, es interesante saber que el valor propio más positivo de una matriz no-negativa reside en la recta real. Este teorema indica que la disposición de los valores propios de una matriz no-negativa es efectivamente como el que se aprecia en la figura 18, en particular porque existe un valor propio sobre la recta real que demarca la máxima distancia a la cual pueden estar los demás valores propios del origen. Para nosotros este teorema tiene implicancias fundamentales sobre las matrices de Metzler, pues como hemos visto, podemos transformar una matriz no-negativa en una matriz de Metzler con el único efecto de correr los valores propios hacia la izquierda, aunque mantenemos la disposición relativa de estos. Con estos elementos en mente, empecemos la segunda demostración.

**Demostración** Considerese la matriz no-negativa  $P$  tal que  $P - \gamma I = J$ . Elegiremos una matriz  $P$  tal que  $\gamma = a_m = \max_i a_i$ , es decir:

$$P = J + a_m \tag{88}$$

Como  $-a_m$  es el elemento más negativo en la diagonal de  $J$  tenemos asegurado que en la diagonal de  $P$  sólo habrán 0s y elementos positivos, asegurando que  $P$  es no-negativa. Con esto hemos asegurado que el valor propio más a la derecha para  $J$  queda sobre la recta real. Lo siguiente es utilizar al regla de los signos de Descartes para verificar que no existen soluciones reales positivas para  $pJ$ , así que la raíz real más a la derecha debe ser el 0 que pedimos como condición. Esto implica que todos los valores propios de  $J$  se encuentran a la izquierda de 0, es decir, su parte real es negativa. ■

Aquí debemos recalcar el llamativo hecho de que el valor que utilizamos para correr los valores es  $a_m$ , o sea, la constante de decaimiento mayor de nuestro sistema. Considerando que el 0 corresponde al valor propio de Perron-Frobenius en  $P$ , eso significa que todos los valores propios de  $J$  se encuentran a lo más a una distancia de  $a_m$  del punto  $(-a_m, 0)$ .

Esta demostración parece un poco ecléctica pero nos permite observar no sólo la configuración de valores propios, sino que además nos permite entender que sucede con estos bajo pequeñas perturbaciones en la matriz. Es más, habiendo hecho explícita la relación entre nuestro sistema y las matrices de Metzler, las cuales aún son tema activo de investigación, hemos trazado un puente de forma que cualquier resultado sobre estas matrices podría tener impacto en el entendimiento de los sistemas biológicos.

Como el lector se habrá dado cuenta, estas son demostraciones de que un lazo (por lo tanto quizás un sistema cRAO) es estable. Estas demostraciones son basadas en argumentos de clausura y no requieren matemáticas avanzadas, al menos nada más allá de cursos de pregrado. Algo fundamental sobre el teorema que hemos demostrado aquí, sobre lo que no se ha hecho el énfasis suficiente es que la estabilidad es independiente de las constantes cinéticas de las reacciones involucradas en el lazo. Esto le da una generalidad bastante amplia al teorema y permite utilizarlo como argumento sin necesidad de contar con los detalles fisicoquímicos del sistema bajo análisis.

## **Dinámica Sistémica**

Dedicaremos lo que queda de este capítulo a fundamentar cómo es que la estabilidad del lazo podría tener un impacto relevante en la dinámica sistémica al punto incluso de que el lazo domine esta dinámica. Consideremos primero las distintas formas en las que un sistema puede ser distinto de lazo que contiene. Un ejemplo es que el lazo controle moléculas que no forman parte de él, caso en el cual la dinámica del lazo no se ve perturbada por

esta molécula. Es fácil ver que el lazo alcanzará su equilibrio, fijando las variables involucradas en la cinética de las moléculas dependientes. Así, estas alcanzarán un equilibrio entre su tasa de producción (dada por el lazo y otros elementos) y su degradación (dada por su propia concentración). Otro caso es el de moléculas que controlen al lazo pero no sean parte de él. Para estas moléculas hemos demostrado que forman parte de un lazo aún más grande, por lo que posiblemente encuentren estados estacionarios, transformándose entonces en perturbaciones constantes, las cuales el lazo puede acomodar sin problemas. Ésto nos lleva al interesante fenómeno de tener las dinámicas de más de un lazo interactuando, situación que probablemente (pero no formalmente demostrado) corresponda a una dinámica similar a la un lazo único o una versión más restringida de este (donde no todos los equilibrios están permitidos). Finalmente queda considerar interacciones no lineales, muy comunes en biología (activación, inhibición, fenómenos de cooperatividad e incluso transiciones de fase o biestabilidad); caso que se hará más claro con el argumento algebraico que sigue a continuación, el cual además nos ayudará a formalizar el resto de las intuiciones aquí vertidas.

## **Variedades Invariantes**

Hasta ahora hemos argumentado tres puntos sobre la condición que pedimos para el balance de flujos:

- Considerando al lazo inmerso en un sistema más grande, hay argumentos para pensar que la condición de equilibrio ocurrirá
- Esta condición genera una recta de equilibrios en el espacio de fases
- Estos equilibrios son estables

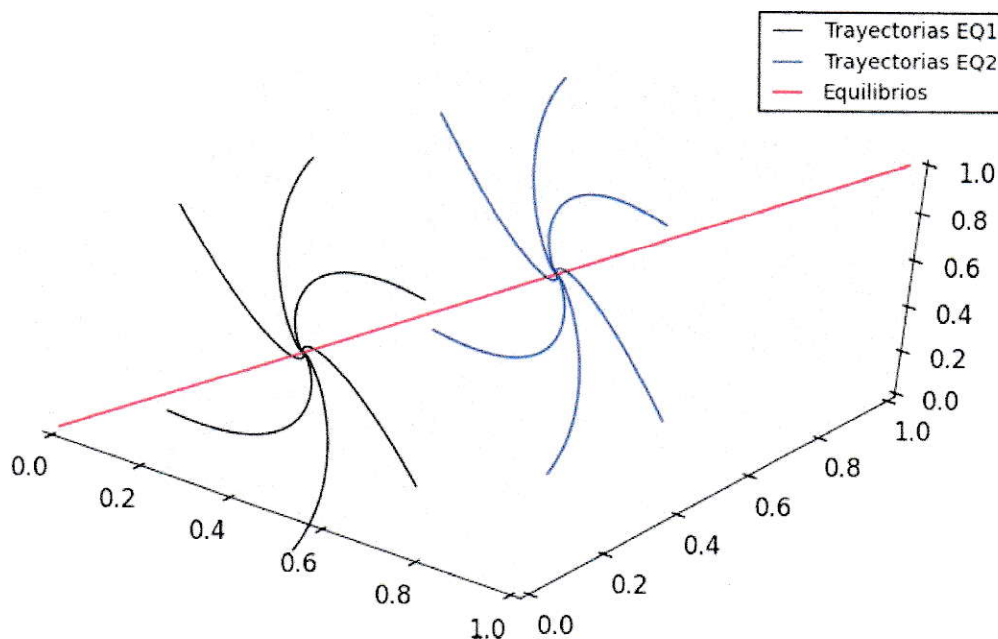
Nos gustaría agregar algo de información a lo que ocurre en el espacio de fases cuando se logra la condición de balance de flujos. Como se puede apreciar en la figura 20, el espacio de fases parece componerse de una serie de variedades invariantes. En particular, para el caso de 2 dimensiones ilustrado, estas variedades son rectas. Si bien en más dimensiones las variedades podrían complicarse (e.g. por la aparición de raíces imaginarias), dada la linealidad del sistema (y por lo tanto la invarianza de su Jacobiano en los distintos equilibrios) el campo vectorial siempre conservará esta simetría alrededor del eje generado por los equilibrios. Estas variedades invariantes corresponden a invariantes del sistema, los cuales permiten caracterizar un estado, por ejemplo, si se conoce este invariante para el sistema, puede predecirse hacia qué equilibrio tenderá. Para ilustrar este punto notemos que es posible cambiar de coordenadas el sistema para escribir el Jacobiano en su forma normal de Jordan, luego el sistema queda expresado como

$$\begin{aligned}\frac{\partial u_1}{\partial t} &= 0 \\ \frac{\partial u_i}{\partial t} &= \lambda_i^*\end{aligned}\tag{89}$$

es decir, existe una combinación lineal de las concentraciones que se mantiene invariante (la masa estequiométrica) y en general el resto de las combinaciones lineales se comportan exponencialmente alrededor del equilibrio. Con exponencial nos referimos a que, dada la forma de la matriz en el caso general donde los valores propios son simples, las soluciones tenderán a acercarse a los puntos de equilibrio ya sea siguiendo una exponencial (valores propios reales) o una espiral (en el caso de valores propios complejos). Para ayudar a entender esto se puede ver la figura 21, aunque este punto se verá más a profundidad en la siguiente sección.

---

\*si los valores propios no son simples, aparecen términos de acoplamiento entre las demás variables  $u_i$ , pero no es el caso para  $\lambda = 0$ .



**Figura 21:** Campo vectorial para un sistema en tres dimensiones. Se puede ver la recta de equilibrios y una serie de trayectorias que convergen a distintos equilibrios. Es evidente que estas trayectorias (aparentemente curvas, pues corresponden a valores propios complejos) se ubican sobre un plano, el cual dado un cambio de variables adecuado se transforma en una recta.

Por otro lado el tener invariantes permite orientar las mediciones que uno realiza, reducir la dimensión del sistema, o incluso caracterizarlo de acuerdo a este invariante... Una forma alternativa de ver estas variedades es como restricciones que el sistema siempre debe cumplir. Curiosamente este punto fue someramente investigado por Rosen, quien nombró a este tipo de sistemas, sistemas cerrados, pues uno puede plantear que el invariante tiene la forma:

$$s_1 C_1 + s_2 C_2 + \dots = \sum_i^n s_i C_i = 0 \quad (90)$$

y por lo tanto podría decirse que el sistema mantiene su masa, o se comporta estequiométricamente, siendo  $s_i$  los coeficientes estequiométricos de transformación entre una especie molecular y otra. Para más detalles se puede consultar la obra de Rosen o específicamente estos trabajos: Hearon (1953); Hyver (1972).

Finalmente tenemos que, como estas variedades son parte del comportamiento lineal del sistema, para perturbaciones no lineales de todas formas podemos asegurar que en la vecindad de los equilibrios el sistema presentará estas invariantes. Esto se debe a que como es sabido en el campo de los sistemas dinámicos, las contribuciones más importantes en un vecindario de tamaño arbitrario de los equilibrios proviene de los términos lineales.

## Teoría de la Variedad Central

Nos queda mencionar un último conjunto de teoremas, relevantes para entender el alcance de los resultados aquí derivados. La teoría de la variedad central, conformada por tres teoremas, nos ayudará a formalizar la intuición de que la estabilidad del lazo podría ser una propiedad *contagiosa*. Para entender hacia adonde apunta esta teoría es necesario recordar que un sistema no-lineal puede estudiarse cualitativamente en la cercanía de sus equilibrios linearizándolo. Este procedimiento consiste en observar el comportamiento de un sistema en el vecindario de un equilibrio considerando que el sistema es en realidad lineal. Para entender esto, consideremos el siguiente sistema:

$$\frac{\partial x}{\partial t} = f(x) \tag{91}$$

con el equilibrio  $x^*$ , es decir

$$f(x^*) = 0 \tag{92}$$

Para hacer una buena aproximación lineal de un sistema se lo reescribe de la siguiente forma:

$$\frac{\partial x}{\partial t} = Ax \tag{93}$$

donde  $A$  correspondería al jacobiano del sistema evaluado en el equilibrio, es decir

$$A = \frac{\partial f}{\partial x}(x^*) \quad (94)$$

Entonces pueden estudiarse los valores y vectores propios de  $A$ , los cuales generan tres subespacios: El subespacio estable y el inestable que se mencionaron anteriormente y el subespacio central. La teoría de la linearización asegura que para el sistema no lineal, las variedades estable, inestable y central son tangentes (en el equilibrio) a las variedades estable, inestable y central del sistema lineal. La teoría de la variedad central provee justificaciones rigurosas de porque entonces es valido reducir el sistema al estudio de la variedad central del sistema lineal. Veremos a continuación las condiciones necesarias y los tres teoremas de esta teoría, además de su pertinencia para nuestro análisis.

## Condiciones

La teoria de la variedad central puede aplicarse a sistemas como el siguiente:

$$\begin{aligned} \dot{x} &= Ax + f(x, y) \\ \dot{y} &= By + g(x, y) \end{aligned} \quad (95)$$

con  $A$  y  $B$  matrices tales que todos los valores propios de  $A$  tienen parte real nula y todos los valores propios de  $B$  tienen parte real negativa y  $f$  y  $g$  funciones diferenciables dos veces que además cumplen con  $f(0, 0) = 0$ ,  $f'(0, 0) = 0$ ,  $g(0, 0) = 0$  y  $g'(0, 0) = 0$ . Es decir, sistemas que tienen componentes lineales atractores o de equilibrio como nuestro sistema. Dada la discusión anterior puede entenderse que bajo estas condiciones, todo componente de una solución que no corresponda al subespacio central corresponde al subespacio estable (pues solo hay valores propios con parte real negativa

o nula) y estos últimos tenderán a desaparecer con el avance del tiempo, como también se vio anteriormente, quedando limitados así a comportamientos transientes.

Para nuestro caso, como se vio en la ecuación 89, es posible escribir el Jacobiano de nuestro sistema en su forma canónica de Jordan y obtener un cambio de variables que cumple con la forma de la ecuación 95. Además aquí veremos un camino más detallado para entender porqué es posible expresar el sistema así. Consideremos primero el teorema de Schur que asegura (y provee constructivamente) la existencia de una descomposición de cualquier matriz cuadrada en un producto de la forma

$$A = QTQ^{-1} \quad (96)$$

donde  $Q$  es una matriz unitaria y  $T$  es una matriz triangular superior. Dado que esto indica que  $A$  y  $T$  son similares, se puede afirmar que tienen los mismos valores propios y por esto mismo, que los elementos de la diagonal de  $T$  corresponden a los valores propios de  $A$ . De esta forma podemos escribir nuestra matriz  $J$  después del cambio de base  $T_J$  con esta forma:

$$Q^{-1}JQ = T_J = \begin{bmatrix} 0 & \times & \times & \times & \times \\ & \lambda_2 & \times & \times & \times \\ & & \lambda_3 & \times & \times \\ & & & \lambda_i & \times \\ & \mathbf{0} & & & \lambda_n \end{bmatrix} \quad (97)$$

notando que ya hemos hecho explícito el hecho de que uno de los valores propios es efectivamente nulo. Así, recordando la ecuación 57, las nuevas variables  $S_i$  (que corresponden a combinaciones lineales de  $C_i$ ) describen nuestro sistema dinámico como:

$$\frac{\partial S}{\partial t} = T_J S \quad (98)$$



con el claro caso de  $S_1$  que se puede escribir como:

$$\dot{S}_1 = 0S_1 + f(S_1, S_r) \quad (99)$$

y el resto de las variables pudiéndose escribir como:

$$\dot{S}_i = \lambda_i S_i + g(S_1, S_r) \quad (100)$$

con  $f$  y  $g$  por ahora siendo funciones lineales. Es evidente que el sistema, descrito bajo esta transformación lineal, cumple con los requisitos para ser estudiado por la teoría de la variedad central. Además, si bien por ahora hemos mantenido  $f$  y  $g$  como funciones lineales, esto no tiene porqué ser así y una forma factible de incorporar el resto de las reacciones del sistema a la descripción matricial sería agregando esos términos a las funciones  $f$  y  $g$ . Por otro lado, al establecer este vínculo entre la teoría de la variedad central y los lazos de catalizaciones, permitimos que cualquier avance en esta teoría se traduzca en resultados para el estudio de sistemas con lazos. Veamos entonces, qué es lo que la teoría de la variedad central nos puede aportar.

### **Primer teorema**

El primer teorema es un teorema de existencia de la variedad central que además la caracteriza, relacionando su flujo con la ecuación que involucra a la matriz con valor propio nulo.

*Para sistemas del tipo 95, existe una variedad central  $y = h(x)$ ,  $|x| < \delta$ , con  $h$  dos veces diferenciable. El flujo de esta variedad está gobernado por el sistema  $n$ -dimensional*

$$\dot{u} = Au + f(u, h(u)) \quad (101)$$

Tener el flujo de la variedad central es sumamente relevante pues en conjunto con el siguiente teorema tendremos que es posible conocer el comportamiento asintótico de soluciones pequeñas del sistema 95. Además de resaltar lo importante del modo correspondiente al valor propio nulo (pues son sus términos no lineales los que sobreviven), este teorema nos deja entrever que toda la información necesaria para estudiar el sistema está contenida en la ecuación 101.

## Segundo teorema

Este teorema indica que las soluciones para el origen de la ecuación 101 dictaminan el comportamiento de las soluciones en el origen para sistemas como los de la ecuación 95 y reza así:

(a) *Sea la solución en el origen para el sistema 101 estable (asintóticamente estable) (inestable), entonces la solución en el origen para el sistema 95 es estable (asintóticamente estable) (inestable)*

(b) *Sea la solución en el origen para el sistema 101 estable y sea  $(x(t), y(t))$  una solución para 95 con  $x(0), y(0)$  suficientemente pequeña. Entonces existe una solución  $u(t)$  para 101 tal que si  $t \rightarrow \infty$ ,*

$$\begin{aligned} x(t) &= u(t) + \mathcal{O}(e^{-\gamma t}) \\ y(t) &= h(u(t)) + \mathcal{O}(e^{-\gamma t}) \end{aligned} \tag{102}$$

con  $\gamma > 0$  una constante.

Como puede verse, conforme  $t \rightarrow \infty$  los componentes no pertenecientes al espacio central se desvanecen exponencialmente. Para encontrar  $h(x)$ , la variedad central, se puede reemplazar en la segunda ecuación de 95:

$$h'(x)[Ax + f(x, h(x))] = Bh(x) + g(x, h(x)) \tag{103}$$

Pero resolver este sistema es en general difícil, sin embargo el teorema que sigue indica como pueden obtenerse aproximaciones arbitrariamente precisas para  $h(x)$ .

El gran aporte de este teorema es demostrar que efectivamente es la variedad correspondiente al espacio propio nulo la que domina el comportamiento del sistema. Dicho de otra forma, podemos reducir la dimensión del análisis de  $N$  a 1, que corresponde a la dimensión de la variedad central.

### Tercer teorema

Este teorema manifiesta que es posible aproximar la variedad central, al menos en principio, a cualquier grado de precisión, por lo tanto quita el impedimento, al menos dentro de cierto orden de precisión, de tener que encontrar la variedad central explícitamente. Definamos primeramente una función de error  $M$  para una aproximación  $\phi$ :

$$M_\phi(x) = \phi'(x)[Ax + f(x, \phi(x))] - B\phi(x) - g(x, \phi(x)) \quad (104)$$

evidentemente  $M_h = 0$ . El tercer teorema puede enunciarse entonces:

*Sea  $\phi$  una función una vez diferenciable con  $\phi(0) = 0$  y  $\phi'(0) = 0$ . Supongamos que si  $x \rightarrow 0$ ,  $M_\phi(x) = \mathcal{O}(|x|^q)$  con  $q > 1$ . Entonces si  $x \rightarrow 0$ ,  $|h(x) - \phi(x)| = \mathcal{O}(|x|^q)$*

Esto indica que el error de una aproximación que cumpla con las condiciones exigidas, al menos en un vecindario del origen, corresponde a una potencia de  $x$ . Es decir, podemos aproximar  $h(x)$  como un polinomio, y mientras mayor sea el grado de este polinomio, menor será el error de nuestra aproximación.

### Significancia

En definitiva, esta teoría marca una forma en la que los resultados obtenidos para el lazo pueden extenderse al resto del sistema, es decir, introducir en la descripción matricial las

moléculas que no forman parte del lazo y que se han omitido hasta ahora. Si bien hay otros resultados que también serán útiles en este camino, la teoría de la variedad central tiene a su favor que nos permite encapsular las no linealidades y otras interacciones en las funciones genéricas  $f$  y  $g$ , y por lo tanto trabajar con sistemas más complejos. En este contexto, la teoría de la variedad central parece no decirnos mucho aún, pero este se debe esencialmente a que nuestro sistema corresponde efectivamente al sistema al cual se reduciría un sistema no lineal cuando se le aplica esta teoría <sup>1</sup>.

Otro aspecto valioso de esta teoría es que nos permite evaluar cuando ocurrirán bifurcaciones en los regímenes, lo cual provee información acerca de los puntos críticos del sistema y en el contexto de los sistemas biológicos permitiría entender fenómenos como la biestabilidad o el comportamiento más bien discreto o catastrófico (en contraposición con continuo) de algunos fenómenos biológicos. Es más, la teoría puede extenderse para considerar valores propios positivos, sin embargo esa discusión escapa a los límites de esta tesis.

Para concluir con esta sección es importante notar que si bien la teoría de la variedad central se ha planteado en referencia al equilibrio del origen, lo dicho aquí es válido para todos los equilibrios del sistema, pues el campo vectorial es esencialmente simétrico respecto del vector propio cuyo valor propio es nulo. Esta propiedad se desprende de la naturaleza lineal del sistema.

## Más allá del Lazo

Hasta el momento, este capítulo se ha centrado en el lazo de catalizaciones presente en los sistemas de nuestro interés. Habiendo demostrado que siempre existe al menos un lazo en estos sistemas y posteriormente habiendo derivado teoremas acerca de la estabilidad de

---

<sup>1</sup>Si recordamos la forma canónica de Jordan para  $J$ , se tiene que en general si el sistema es lineal  $f(x, y) = g(x, y) = 0$

los estados estacionarios del lazo, es hora de reconocer que todo el análisis de la dinámica ha ignorado completamente lo que denominamos el resto del sistema, arguyendo que la dinámica del sistema completo esta supeditada al comportamiento del lazo. Una pregunta razonable es, qué tan válido es este razonamiento? O hilando más fino aún, bajo qué circunstancias esta permitido descartar el comportamiento dinámico del resto del sistema? Si bien no existe una única respuesta, aquí hemos mencionado la teoría de la variedad central como una herramienta para extender los argumentos de estabilidad del lazo al sistema entero y justificar, con cierto nivel de rigor, porqué podría considerarse que el lazo provee estabilidad al sistema.

Otra ruta que valdría la pena explorar es realizar descomposiciones o simplificaciones al sistema, de manera de demostrar que la dinámica de sistemas complejos puede reducirse a la de sistemas más pequeños, idealmente, al lazo mismo. En esta dirección existen resultados, tanto desde la teoría de grafos, como de formalismos para redes bioquímicas, hasta enfoques principalmente dinámicos, e incluso en esta tesis se sugieren algunos caminos abiertos, sin embargo la cantidad de opciones distintas las hace inabordables dentro de los confines de este trabajo.

Nuestra apuesta es que la estabilidad conferida por el (los) lazo (s) de un sistema es lo que le provee perdurabilidad en el tiempo a los sistemas que realizan el modelo autopoietico. En otras palabras, contener un lazo permite al sistema interactuar con un ambiente perturbativo sin ser destruido ni comprometer su organización, lo cual es una propiedad fundamental de los seres vivos. Es más, el lazo como mecanismo estabilizador podría ser el centro mismo de la increíble plasticidad que exhiben los organismos, pues permite modificaciones a la estructura <sup>1</sup> manteniendo la organización <sup>2</sup>.

Quiero hacer hincapié en que este último punto abarca las escalas más diversas, por

---

<sup>1</sup>aquí vagamente definida como componentes y/o relaciones entre componentes que realizan una organización

<sup>2</sup>entiéndase por organización la macro-arquitectura de la red de reacciones que componen en sistema; en nuestro caso en particular, lo que identifica a los organismos vivos como pertenecientes a la clase cRAO

lo que el lazo, con su dinámica estabilizadora, podría sustentar fenómenos biológicos a escalas muy diversas, desde el metabolismo hasta la ecología y desde la sobrevivencia a una perturbación hasta la evolución biológica, pero los alcances reales de este punto de vista por ahora no son claros y estas declaraciones son de naturaleza especulativa.

# Aplicaciones

En este capítulo se mostrará cómo los resultados del capítulo anterior pueden traducirse a un entendimiento más profundo de los sistemas biológicos. Empezaremos con simulaciones que ilustren el comportamiento de un lazo y un sistema estructuralmente similar, luego veremos un caso pedagógico, y finalmente veremos la aplicación del lazo a sistemas celulares.

## Cadena y Lazo de Catalizaciones

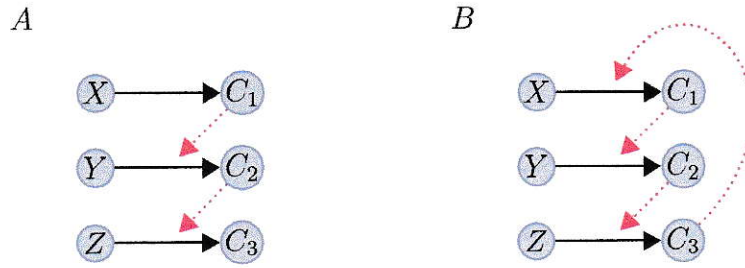
Dado que el propósito central de este capítulo es ilustrar como es que el lazo podría estabilizar o brindar adaptabilidad al sistema en el cual esta imbuido, dedicaremos esta sección a entender cualitativamente la dinámica de un lazo y un sistema estructuralmente muy similar pero que no consiste de un lazo. La estructura general de los sistemas que simularemos se puede ver en la figura 22 y particularmente trabajaremos con el caso de 2 dimensiones para poder graficar los campos vectoriales. Las ecuaciones diferenciales que se integrarán para llevar a cabo esta simulación corresponden a

$$\frac{\partial C_i}{\partial t} = \begin{bmatrix} -1 & 1 \\ 0,5 & -0,5 \end{bmatrix} C \quad (105)$$

para el sistema con lazo y

$$\frac{\partial C_i}{\partial t} = \begin{bmatrix} -1 & 0 \\ 0,5 & -0,5 \end{bmatrix} C + \begin{bmatrix} 0,5 \\ 0 \end{bmatrix} \quad (106)$$

para la cadena de catalizaciones sin lazo.



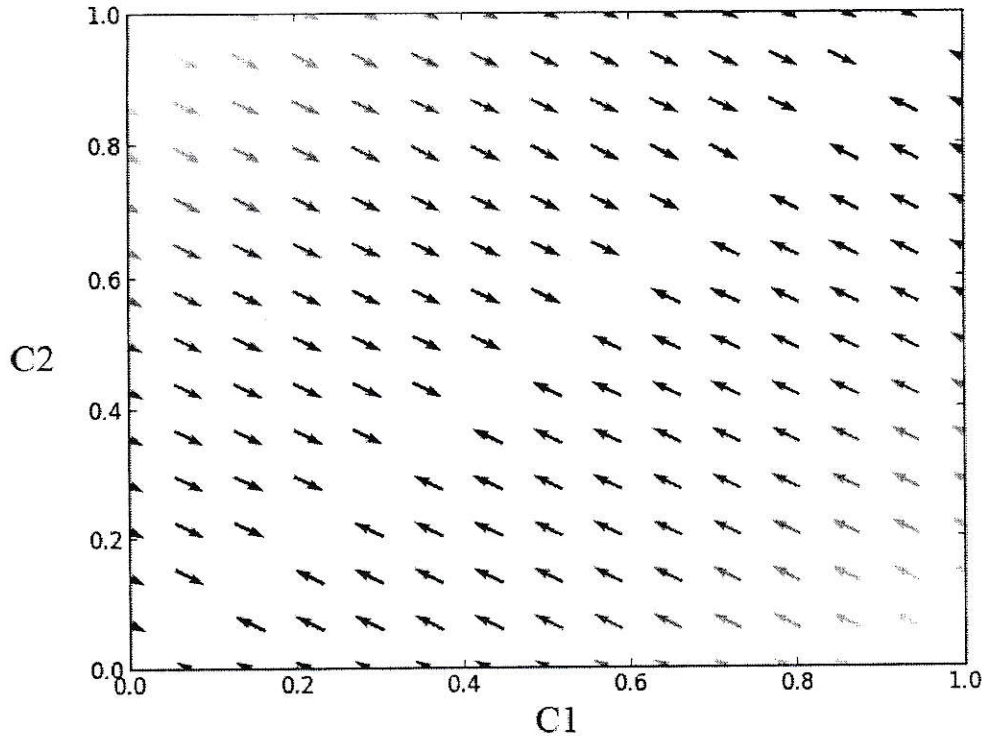
**Figura 22:** Diagrama de dos estructuras de control distintas. **A** Una cadena de catalizaciones en la cual no hay un ciclo de control. **B** Un lazo de catalizaciones en el cual el último elemento controla al primero formando un ciclo.

Podemos comprobar que en la ecuación 105 el determinante es nulo, por lo que existe un vector propio nulo (cuyo espacio generado se nota en la diagonal que cruza la figura 23). Además, como se predijo, esta diagonal (que corresponde al invariante de la variedad central) es atractora. Es decir el sistema contiene una infinitud de puntos de equilibrio estable distintos.

En el segundo caso, donde hemos desacoplado la generación del  $C_1$  de la concentración de  $C_2$  (por eso existe un componente lineal y uno constante), podemos calcular rápidamente que el sistema estará en equilibrio sólo en el punto (0,5 0,5). Si linearizamos el sistema en este punto, dado que el Jacobiano es triangular inferior, es trivial demostrar que este equilibrio único es estable. Esto trae como consecuencia que sin importar las perturbaciones que afecten al sistema, este vuelve siempre al mismo punto, incluso si se lleva al sistema al origen!

Un simple argumento heurístico puede ayudar a iluminar este segundo caso más allá del álgebra. Consideremos que el sistema inicia en un estado inicial arbitrario. Dado que en este caso la concentración de  $C_1$  esta controlada únicamente por su tasa de entrada y salida, siendo la primera fija y la segunda monótona creciente con la cantidad de  $C_1$ , es



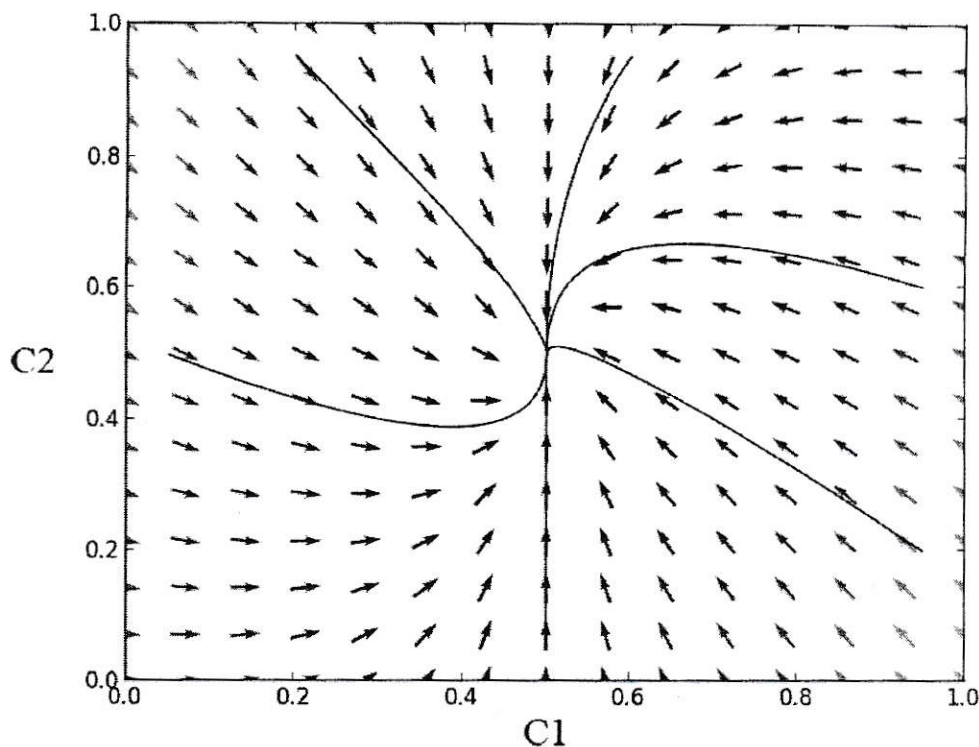


**Figura 23:** Campo vectorial de un sistema que contiene un lazo. Puede apreciarse una recta de infinitos equilibrios atractores correspondientes al kernel.

evidente que se llegará a un equilibrio único de la siguiente forma:

$$\frac{\partial C_1}{\partial t} = 0 \iff -C_1 + 0,5 = 0 \iff 0,5 = C_1 \quad (107)$$

Una vez este equilibrio es alcanzado, lo mismo ocurrirá para  $C_2$ , pues ahora la entrada es fija (ya que depende de  $C_1$  y de otros términos constantes) y por lo tanto también encontrará una concentración de equilibrio. Siguiendo este argumento puede verse que existe un equilibrio estable que además es atractor global (todas las trayectorias tienden en el infinito a él) para el sistema (independiente de su dimensión). Como el equilibrio es único, el comportamiento exhibido por este sistema es, independiente de las condiciones iniciales o perturbaciones, el de volver a este punto de equilibrio. Un comportamiento tan reducido es cualitativamente opuesto al exhibido por el sistema con lazo y dándole un sentido sistémico, no puede acomodar perturbaciones ni “aprender” o adaptarse pues su



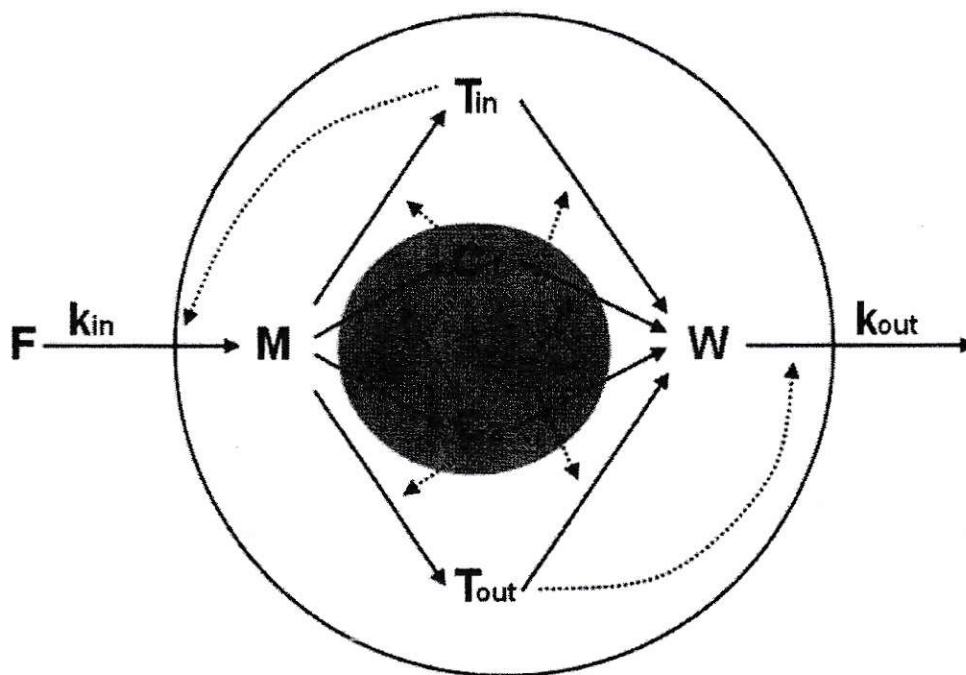
**Figura 24:** Campo vectorial de un sistema estructuralmente similar a un lazo, pero se ha eliminado una reacción. Puede apreciarse un único punto de equilibrio atractor, exhibiendo por lo tanto un comportamiento asintótico más simple y limitado.

único modo de lidiar con perturbaciones es disiparlas. Esta falta de histéresis es inusual en los sistemas biológicos y probablemente lleva asociada una gran fragilidad estructural, por lo que consideramos este tipo de estructura poco relevante para la biología. Otro elemento que demuestra que este tipo de sistemas son en realidad impracticables es que la primera reacción no es catalizada, por lo cual es de esperarse que ocurra a tasas bastante inferiores comparada con las del resto del sistema, limitando así el tamaño de éste.

## Un caso pedagógico

Como primera aplicación estudiaremos un sistema simple, casi minimal, que podría entenderse como un metabolismo abstracto. El sistema tiene la estructura que se ve en la figura 25 y pueden distinguirse de entre sus componentes un par de catalizadores que es-

tán efectivamente formando un lazo pues  $C_1 \vdash C_2 \vdash C_1$ . De hecho, este sistema es un sistema RAF, además es un sistema RAO (recordemos que estas dos condiciones no son excluyentes) e incluso un cRAO.



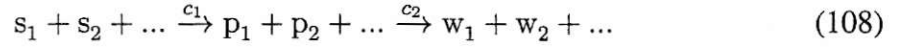
**Figura 25:** Ejemplo de sistema RAF y cRAO a la vez. Los elementos del conjunto  $F$  son incorporados al sistema hacia el metabolito abstracto  $M$ .  $M$  es transformado por los catalizadores  $C_1$  y  $C_2$  en sí mismos y en un par de catalizadores más,  $T_{in}$  y  $T_{out}$  los cuales catalizan las reacciones de influjo y eflujo de masa hacia y desde el sistema. Además, todos los componentes del sistema, con excepción de  $M$ , tienen reacciones catalizadas que los transforman en  $W$ , un metabolito de salida. Puede observarse un lazo de catalizaciones formado por  $C_1$  y  $C_2$ .

Como podemos notar de la figura 25, las moléculas  $T_{in}$  y  $T_{out}$  controlan los flujos de entrada y salida del sistema. Las moléculas  $C_1$  y  $C_2$  conforman un lazo como se mencionó anteriormente. Ni  $M$  ni  $W$  controlan molécula alguna (no son catalizadores), por lo que se descarta su participación en lazos. todas las moléculas del sistema pueden generarse a partir de la molécula  $F$  (la cual, por sí sola, constituye el conjunto  $F$ ). Con esto podemos ver que el sistema es  $F$ -generado. Por otro lado, es evidente que todas las reacciones son catalizadas por componentes del sistema, de lo que se desprende que el sistema es reflexivo auto catalítico y tiene clausura eficiente. En este sistema además hemos incluido que las

reacciones de degradación sean catalizadas.

## Primeras Simulaciones

Para poner a prueba la idea de que el lazo es un elemento estabilizador, que provee adaptabilidad y sobre todo necesario para la existencia de estados estacionarios no triviales, realizamos simulaciones numéricas de este sistema simple con y sin la dinámica del lazo. Para simular la condición sin lazo pero mantener la esencia estructural del sistema, fijamos las concentraciones de los catalizadores eliminando las reacciones de formación y degradación de estas moléculas. Las simulaciones consideraron un modelo cinético de acción de masas como se ha descrito anteriormente en esta tesis, es decir, para reacciones del tipo:



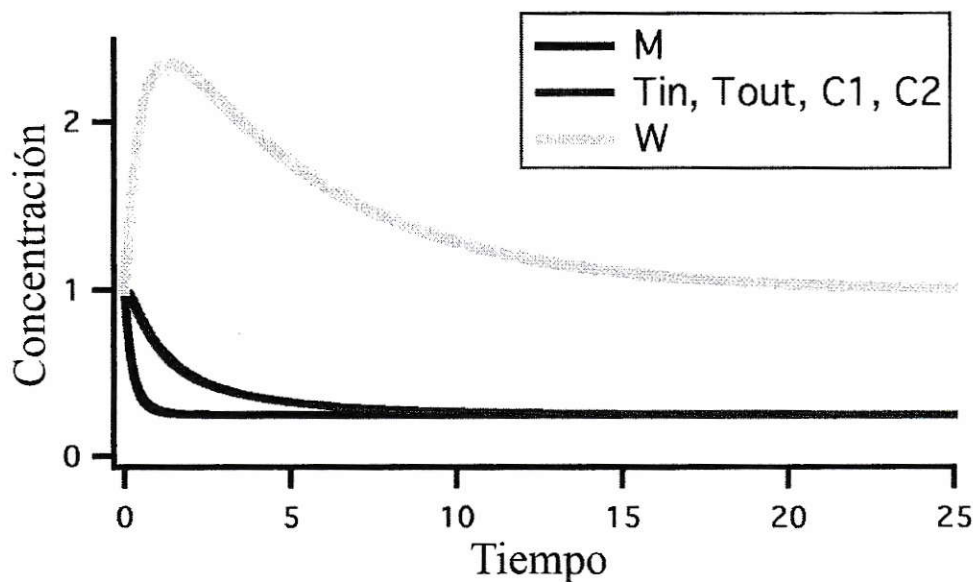
se utilizó ecuaciones diferenciales de este tipo:

$$\frac{d[p_k]}{dt} = k_1[c_1] \prod_i [s_i] - k_2[c_2] \prod_j [p_j] \quad (109)$$

Así el sistema entero puede describirse como:

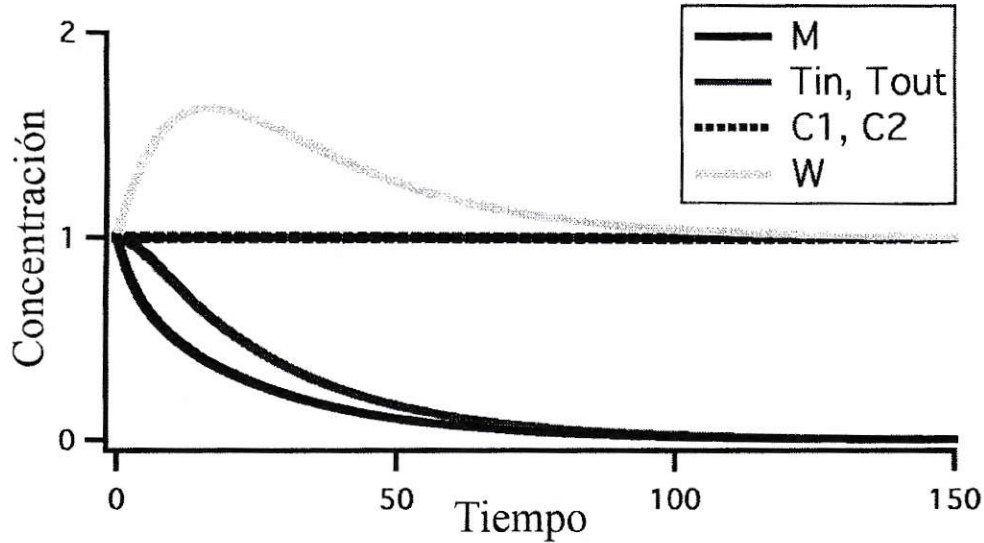
$$\begin{aligned} \frac{\partial M}{\partial t} &= k_{in}T_{in} - M[C_1(k_{T_{in}} + k_{C_2}) + C_2(k_{T_{out}} + k_{C_1})] \\ \frac{\partial T_{in}}{\partial t} &= k_{T_{in}}MC_1 - k_{T_{in}}^-T_{in}C_2 \\ \frac{\partial T_{out}}{\partial t} &= k_{T_{out}}MC_2 - k_{T_{out}}^-T_{out}C_2 \\ \frac{\partial C_1}{\partial t} &= k_{C_1}MC_2 - k_{C_1}^-C_1C_2 \\ \frac{\partial C_2}{\partial t} &= k_{C_2}MC_1 - k_{C_2}^-C_1C_2 \\ \frac{\partial W}{\partial t} &= C_2[C_1(k_{C_2}^- + k_{C_1}^-) + k_{T_{out}}^-T_{out} + k_{T_{in}}^-T_{in}] - k_{out}T_{out} \end{aligned} \quad (110)$$

Las ventajas de simular sistemas relativamente simples son que permiten entender lo que ocurre, las concentraciones pueden graficarse sin mucho problema y no hay un gran número de parámetros a considerar. Los resultados de estas simulaciones se encuentran en las figuras 26 y 27.



**Figura 26:** Dinámica del sistema simple con lazo. El sistema alcanza un estado estacionario no trivial. Las concentraciones de los distintos metabolitos varían en armonía, ajustándose mutuamente hasta encontrar un estado estacionario.

El efecto que tiene la dinámica del lazo es notorio y decidor: Si se interrumpe el lazo de catalizaciones el sistema tiende al origen, es decir, se degrada hasta desaparecer. Observando detalladamente el caso con lazo podemos ver que para obtener un estado estacionario no-trivial estable es necesario que las concentraciones de los catalizadores disminuyan cuando lo hacen el resto de las moléculas, caso contrario la tasa de destrucción de estas otras moléculas  $T_{in}$  y  $T_{out}$  excede la tasa de generación y el sistema termina por colapsar. Aquí queda en plena evidencia que el lazo, junto con su cinética estabilizadora es una fuente de plasticidad o adaptabilidad, sin la cual los sistemas no pueden lidiar con perturbaciones o condiciones iniciales fuera de algún punto de equilibrio (pues estos son inestables). Para insistir sobre este punto realizamos múltiples simulaciones utilizando

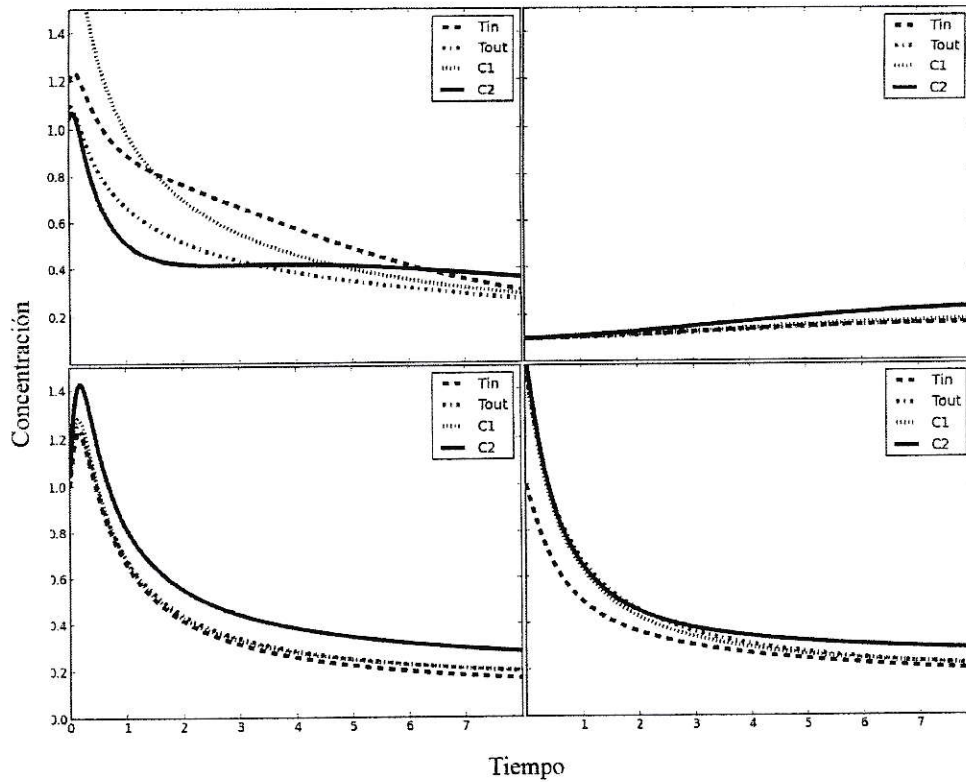


**Figura 27:** Dinámica del sistema simple sin el lazo. A pesar de que las concentraciones de  $C_1$  y  $C_2$  se mantienen constantes y no nulas, puede verse como el sistema llega a un equilibrio en el cual el sistema se desvanece, dejando en evidencia que no sólo es necesario contar con la acción de los catalizadores, sino que la dinámica que estos siguen juega un rol fundamental pues proveen de estabilidad a equilibrios no triviales.

distintas posiciones iniciales, obteniendo cada vez resultados similares como puede verse en la figura 28.

Un examen analítico de la condición de equilibrio realizado con el software Mathematica® indica que existe un único punto fijo que corresponde al siguiente estado:

$$\begin{aligned}
 M^* &= \frac{F k_{C_1} \bar{k}_{C_2}^2 k_{in} k_{T_{in}}}{k_{C_2} \bar{k}_{T_{in}} (k_{C_1} k_{C_2} (\bar{k}_{C_1} + \bar{k}_{C_2}) + k_{C_1} \bar{k}_{C_2} k_{T_{in}} + \bar{k}_{C_1} k_{C_2} k_{T_{out}})} \\
 C_1^* &= \frac{F k_{C_1}^2 \bar{k}_{C_2}^2 k_{in} k_{T_{in}}}{\bar{k}_{C_1} k_{C_2} \bar{k}_{T_{in}} (k_{C_1} k_{C_2} (\bar{k}_{C_1} + \bar{k}_{C_2}) + k_{C_1} \bar{k}_{C_2} k_{T_{in}} + \bar{k}_{C_1} k_{C_2} k_{T_{out}})} \\
 C_2^* &= \frac{F k_{C_1} \bar{k}_{C_2} k_{in} k_{T_{in}}}{\bar{k}_{T_{in}} (k_{C_1} k_{C_2} (\bar{k}_{C_1} + \bar{k}_{C_2}) + k_{C_1} \bar{k}_{C_2} k_{T_{in}} + \bar{k}_{C_1} k_{C_2} k_{T_{out}})} \\
 T_{in}^* &= \frac{F k_{C_1}^2 \bar{k}_{C_2}^3 k_{in} k_{T_{in}}^2}{\bar{k}_{C_1} k_{C_2}^2 \bar{k}_{T_{in}}^2 (k_{C_1} k_{C_2} (\bar{k}_{C_1} + \bar{k}_{C_2}) + k_{C_1} \bar{k}_{C_2} k_{T_{in}} + \bar{k}_{C_1} k_{C_2} k_{T_{out}})} \\
 T_{out}^* &= \frac{F \bar{k}_{C_1} \bar{k}_{C_2} k_{in} k_{T_{in}} k_{T_{out}}}{\bar{k}_{T_{in}} \bar{k}_{T_{out}} (k_{C_1} k_{C_2} (\bar{k}_{C_1} + \bar{k}_{C_2}) + k_{C_1} \bar{k}_{C_2} k_{T_{in}} + \bar{k}_{C_1} k_{C_2} k_{T_{out}})} \\
 W^* &= \frac{F k_{C_1}^2 \bar{k}_{C_2}^2 k_{in} k_{T_{in}} \bar{k}_{T_{out}}}{\bar{k}_{C_1}^2 k_{C_2}^2 k_{out} \bar{k}_{T_{in}} k_{T_{out}}}
 \end{aligned} \tag{111}$$



**Figura 28:** Múltiples simulaciones del sistema simple con lazo. El sistema alcanza siempre el mismo estado estacionario no trivial, independiente de las condiciones iniciales.

Notemos que este punto existe y es no-trivial mientras las constantes y concentraciones involucradas sean positivas, condición que tenemos asegurada. Si bien no corresponde a una serie de estados de equilibrio como esperábamos para el lazo, esto es explicable considerando que los elementos fuera del lazo introducen restricciones adicionales ya que el equilibrio del resto de las especies es más frágil, pues al estar fuera del lazo no tienen una dinámica acomodadora fuerte. Es llamativo que el sistema tienda siempre a este punto de equilibrio, por lo que decidimos estudiar los valores propios del Jacobiano en el equilibrio. Dada la alta dimensionalidad del sistema, decidimos incluir primero una restricción adicional:

$$k_{C_1} = k_{C_2} = \bar{k}_{T_{in}} = \bar{k}_{T_{out}} = \bar{k}_{C_1} = \bar{k}_{C_2} = 1 \quad (112)$$

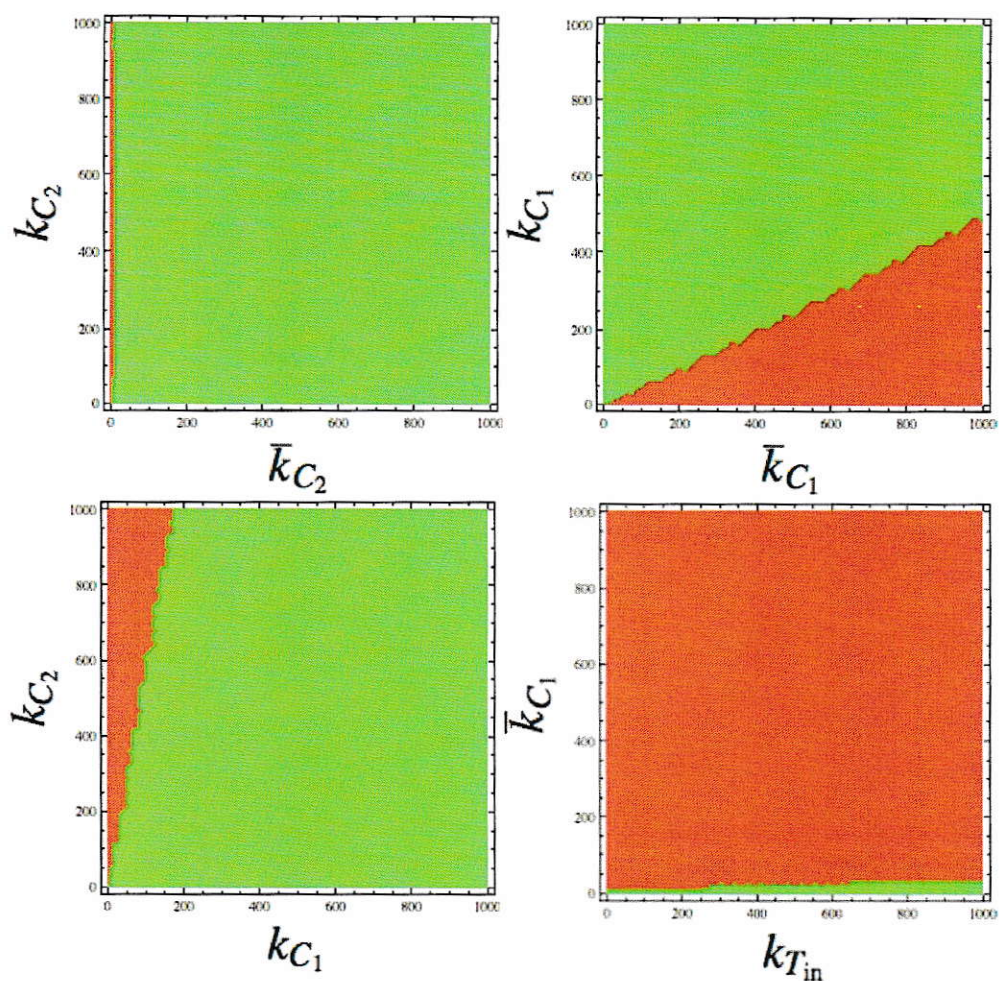
que nos permite, con agradable sorpresa escribir los valores propios así:

$$\begin{aligned}
 \lambda_1 &= -Fk_{in}k_{T_{in}} \\
 \lambda_2 &= -\frac{Fk_{in}k_{T_{in}}}{k_{T_{in}} + k_{T_{out}} + 2} \\
 \lambda_3 &= -\frac{Fk_{in}k_{T_{in}}}{k_{T_{in}} + k_{T_{out}} + 2} \\
 \lambda_4 &= -\frac{Fk_{in}k_{T_{in}}}{k_{T_{in}} + k_{T_{out}} + 2} \\
 \lambda_5 &= -\frac{Fk_{in}k_{T_{in}}}{k_{T_{in}} + k_{T_{out}} + 2} \\
 \lambda_6 &= -\frac{Fk_{in}k_{out}k_{T_{in}}k_{T_{out}}}{k_{T_{in}} + k_{T_{out}} + 2}
 \end{aligned} \tag{113}$$

Es trivial ver que dadas las definiciones de nuestras constantes y de  $F$ , el equilibrio resulta siempre estable, lo que indica que podemos mover libremente  $F$ ,  $k_{in}$ ,  $k_{out}$ ,  $k_{T_{in}}$  y  $k_{T_{out}}$  sin perturbar la estabilidad del sistema. Finalmente, llevaremos a cabo un análisis de bifurcación, en el cual examinaremos el efecto que tiene sobre la estabilidad del sistema cambiar dos parámetros a la vez, para tener una idea de cuál es la superficie (11-dimensional) de estabilidad del equilibrio. Estos resultados se resumen en la figura 29. Para construir esta figura realizamos la simulación cinética anterior con una combinación amplia de parámetros y observamos si acaso el sistema tendía a un régimen estable como el de la figura 26 o por el contrario, tendía a desaparecer como en la figura 27.

Puede verse que en general el sistema es bastante robusto frente a elecciones de constantes cinéticas. Para poner a prueba nuestra hipótesis de que el lazo confiere estabilidad, también realizamos este análisis con el sistema sin lazo. Este resultado es impactante por su contraste: Si bien el sistema con lazo es estable para la mayoría de las combinaciones, para el sistema sin lazo no se pudo encontrar un punto en el cual el sistema sobreviviese. Esto explica que en la figura 29 sólo se muestren los resultados para el primer caso. Del estudio de esta figura se desprende una conclusión adicional, y es que el sistema de-





**Figura 29:** Análisis de bifurcación para el sistema cRAO. Se realizaron simulaciones con distintas combinaciones de las constantes cinéticas y se determinó si el sistema encontraba un estado estacionario no trivial (verde) o decaía hasta desaparecer (rojo). El panorama para la mayoría de las combinaciones se parece al caso  $k_{c_2}$  vs.  $\bar{k}_{c_2}$  y por eso no fue incluido aquí excepto con la notable excepción de los gráficos que involucran  $k_{c_2}$  indicando que este parámetro debe ser tratado con precaución. Esto probablemente sigue de que  $C_2$  es el catalizador degradador para la mayoría de las moléculas. Este mismo análisis para el sistema sin lazo muestra que no existen zonas de estabilidad.

be guardar ciertas proporciones entre las constantes como era de esperarse. Esto podría corresponder tanto al requerimiento de que el lazo tenga un determinante nulo (y por lo tanto el equilibrio corresponda al equilibrio generado por el lazo) como podría responder a restricciones adicionales que imponen las moléculas fuera del lazo.

Viendo lo fructífero que es este sistema como modelo de estudio decidimos realizar más perturbaciones sobre él, a modo de profundizar nuestro entendimiento y cimentar nuestras intuiciones. A continuación presentamos este análisis.

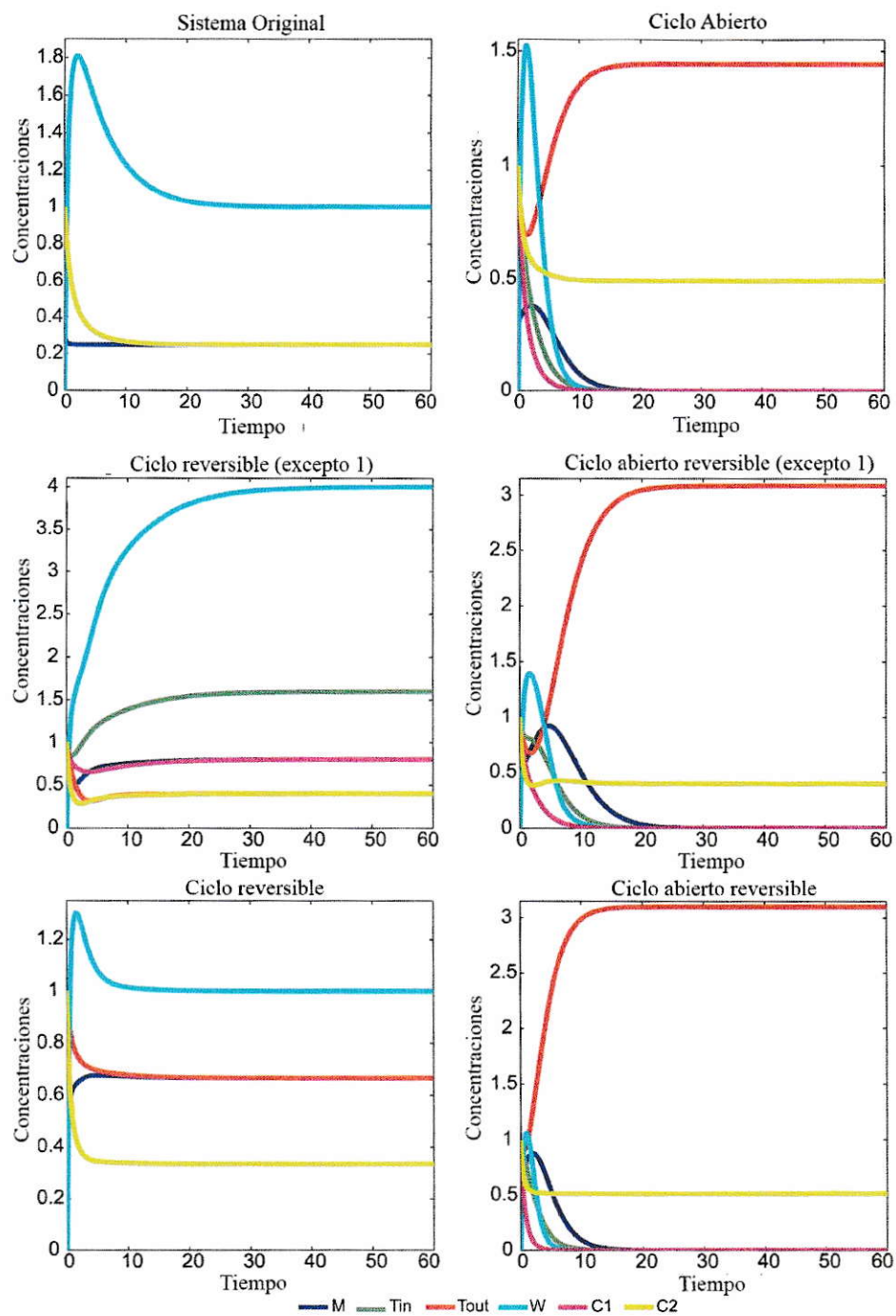
## Otras modificaciones

En esta sección eliminamos o agregamos distintos conjuntos de reacciones que cambian el carácter del sistema presentado en la sección anterior. También se presentan simulaciones representativas y análisis algebraicos aunque sin mucho detalle metodológico (estos pueden revisarse en (Kreyssig y cols., 2012)). Realizaremos las siguientes 5 modificaciones además del caso que ya estudiamos anteriormente:

1. Sistema original
2. Ciclo abierto (sin retroalimentación), se elimina la transformación  $C_n$  en  $C_1$
3. Ciclo reversible (la reacciones se hacen reversibles)
4. Ciclo abierto reversible (2 y 3 juntos)
5. Ciclo reversible excepto por una reacción
6. Ciclo abierto reversible excepto por una reacción (2 y 5 juntos)

Los resultados de simulaciones representativas se presentan en la figura 30.

Se probaron distintas tasas de reacción para las reacciones reversibles y se encontró que no tenían un efecto apreciable en los resultados (siempre y cuando las reacciones se mantuviesen equilibradas hacia los productos). Es aparente que en todos los sistemas en los que se elimina el lazo (ciclos abiertos) el sistema decae, mientras que los sistemas que conservan el ciclo intacto son capaces de acomodar distintas situaciones, reforzando una vez más el punto de esta tesis. Además es posible decir que (dado un análisis en



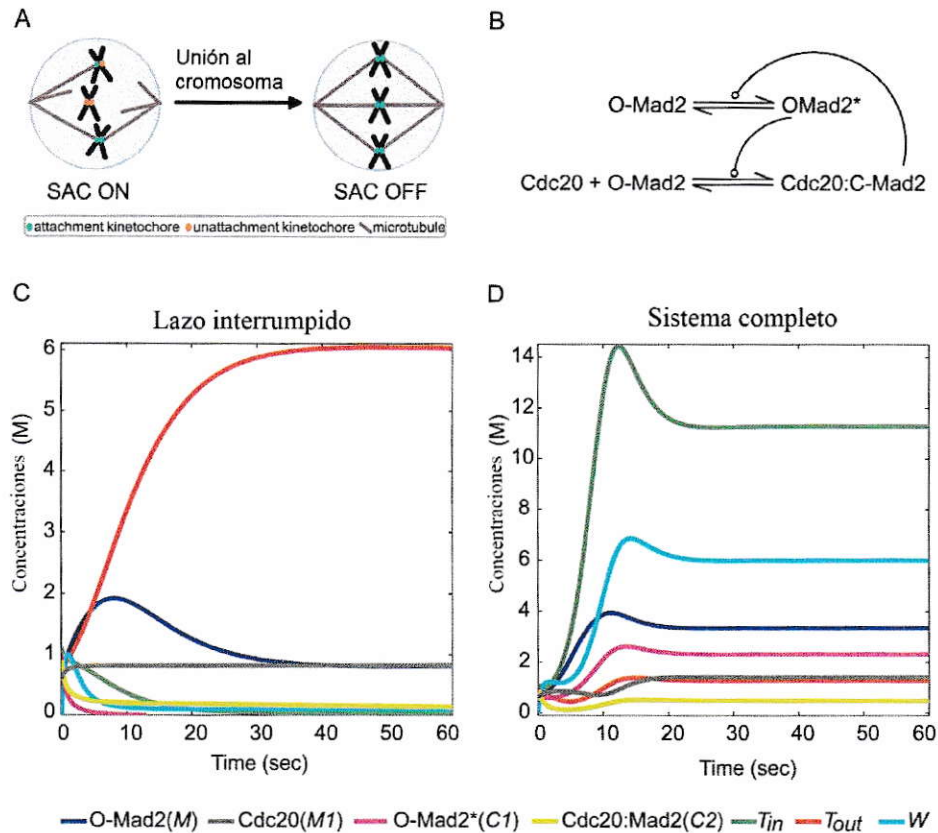
**Figura 30:** Simulaciones de sistemas estructuralmente similares. Todas las tasas de reacción valen 1, excepto las reacciones reversas, que valen 0.1. Los sistemas que conservan el lazo de catalizaciones son estables en el tiempo y se acomodan a distintas situaciones. Los sistemas en los que se abre el ciclo decaen inevitablemente.

términos de organizaciones) no existen puntos fijos para los sistemas de ciclo abierto. Esta conclusión es fácilmente obtenible del hecho de que los puntos fijos implican la existencia de organizaciones, como se mencionó en el capítulo 3. El resto de los casos sólo puede ser tratado numéricamente pero se encontró que para todas las simulaciones realizadas los resultados obtenidos estaban de acuerdo con la conclusión general de esta tesis: Los lazos de catalizaciones son sumamente relevantes para la estabilidad de los sistemas.

## **Un Lazo Biológico**

Finalmente, terminaremos este capítulo analizando la dinámica de un sistema biológico, lo cual nos permitirá ver cómo es que el lazo se implementa o realiza en los organismos vivos. En particular utilizaremos un modelo del punto de control del ensamblaje del huso mitótico (SAC por sus siglas en inglés) por un lado porque se cuenta con suficiente información como para hacer una simulación significativa, y por otro porque nos interesa que se entienda que la teoría que hemos desarrollado aquí no se limita a redes de reacción de moléculas bioquímicamente simples, sino que podemos empezar a entender fenómenos de distintas y variadas escalas con este enfoque basado en clausura.

Por brevedad y para mantenernos enfocados no haremos una descripción detallada de la simulación, sin embargo esta información y otros aspectos de este sistema pueden encontrarse en (Kreyssig y cols., 2012). El modelo se trata esencialmente del sistema ilustrado en la figura 31, el cual corresponde al sistema SAC que asegura la segregación adecuada de los cromosomas durante la mitosis, arrojando a la célula hasta que los cromosomas se encuentran correctamente unidos al huso mitótico (figura 31 A). Esencialmente una ubiquitina ligasa es activada por Cdc20. Una vez se alcanzan niveles suficientes de activación de esta ligasa, se gatilla la segregación de los cromosomas por la ubiquitinación de las securinas y ciclinas B. Sin embargo el sistema está en constante degradación pues difun-



**Figura 31:** Simulaciones del sistema SAC. **A** Diagrama de la función biológica del sistema SAC. **B** Modelo de reacciones del sistema SAC. **C** Simulación en la que se ha interrumpido el lazo. **D** Simulación del sistema completo. Puede verse como el lazo es necesario para que el sistema se mantenga y el sistema SAC permita el ensamblaje del huso mitótico.

de a través del citoplasma. El esquema de las interacciones se ilustra en la figura 31 B. La intuición indica que sin un lazo de realimentación positiva, el sistema se diluirá hasta desaparecer.

Puede verse que los resultados de las simulaciones confirman nuestra teoría de que el lazo es necesario para la existencia de un punto fijo estable no trivial. Como puede verse en la figura 31 C, el sistema sin lazo decae mientras que en la figura 31 D el sistema, al contener el lazo, encuentra un punto fijo.

A lo largo de este capítulo confío en haber hecho visible lo importante que es este esquema o motivo de catalizaciones. Si bien se trata de *toy-problems*, permiten manipu-

laciones que aclaran y enfatizan como es que el lazo de catalizaciones provee puntos de equilibrio estables. Si bien la mayoría de las simulaciones se realizaron en dimensión 2, o al menos el lazo contenía sólo dos especies, los resultados son generalizables a más dimensiones y el número 2, más que una cábala o restricción de nuestro resultado corresponde a la dimensión más alta que permitía visualizar los distintos elementos (redes de reacción, campos vectoriales, cursos temporales) cómodamente. también me gustaría hacer hincapie en que el *cLST* es válido para un uso bastante amplio de la palabra catálisis, como se vio en el último ejemplo. Esto significa que el *cLST* trasciende el dominio de las moléculas simples y podría tener implicancias a muchas y distintas escalas biológicas.

# Discusión, Perspectiva y Proyecciones

En este trabajo persigo un objetivo ambicioso: partir de la noción de clausura y obtener un resultado sobre la estabilidad de las redes metabólicas. Creo que este objetivo fue alcanzado, ya que el *cLST* es un teorema de estabilidad para lazos de catalizaciones, estructuras fuertemente ligadas a la clausura metabólica. Además, colateralmente, hemos introducido un conjunto interesante de técnicas que se pueden utilizar para estudiar la clausura. A continuación precisaré más estas ideas y contrastaré los resultados obtenidos en esta tesis con el estado del arte en otras teorías de la biología. Esto me permitirá situar en contexto la utilidad y alcance de los resultados obtenidos y desde allí plantear direcciones futuras para la línea de trabajo desarrollada aquí.

## Recapitulación de resultados

En términos resumidos y concretos en esta tesis se presenta un resultado principal (el *cLST*) y tres resultados secundarios relacionados que también son importantes:

- Definición de los cRAO
- Existencia de un lazo de catalizaciones
- El teorema de hegemonía del lazo

Finalmente, y aunque no se prueba, se presentan evidencias sobre la importancia de la teoría de la variedad central como herramienta para analizar cambios estructurales y su impacto sobre los estados estables de los cRAO. Creo que con las herramientas que he usado en esta tesis he abierto un camino para hablar *técnicamente* de la noción de clausura

además de haber probado un resultado importante que puede tener consecuencias para futuras investigaciones, el cual podría resumirse en:

Clausura Metabólica  $\Rightarrow$  Estabilidad

Este es el resultado central y durante este capítulo lo analizaré *in extenso*, pero antes es necesario también contextualizar los resultados secundarios de esta tesis.

## **Definición de los cRAO**

Un aporte de este trabajo es haber intentado -y parcialmente logrado- hacer una síntesis entre las distintas teorías y modelos que permita avanzar en el campo. En particular puede:

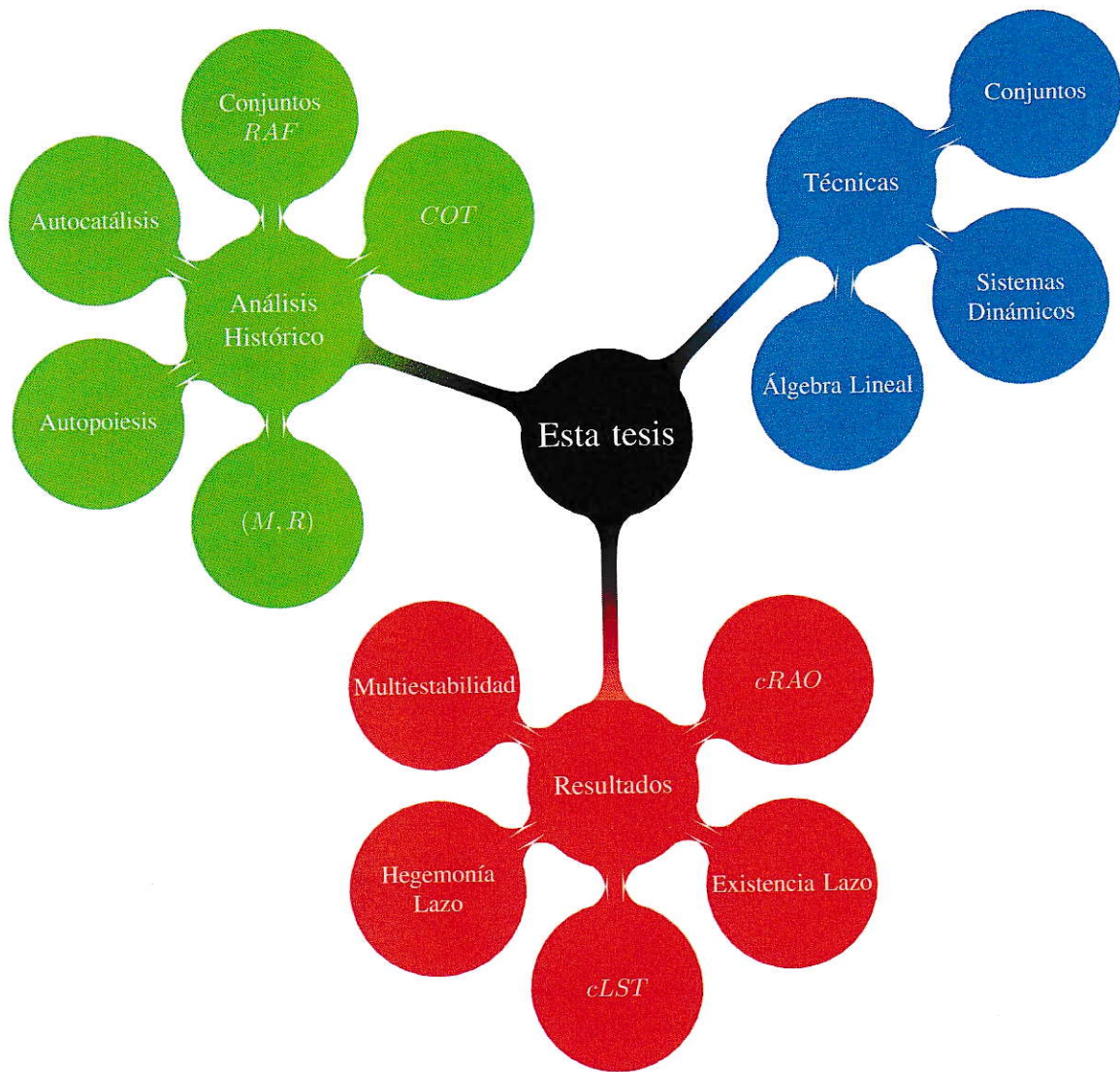
- Definir un nuevo tipo de sistemas (cRAO)
- Usar las herramientas de Conjuntos-RAF y de COT para avanzar analíticamente

Este resultado puede parecer simple pero se inserta en una tradición bastante única en este campo, iniciada por el laboratorio de Biología del Conocer, que es la de *tomar en cuenta* el trabajo de todos. Como dije en la introducción, el campo de la biología teórica se caracteriza por una autarquía intelectual donde cada grupo de ideas parece existir en un vacío conceptual carente de historia y puentes con otras teorías. Como evidencia de esta falla comunicacional podemos ver la falta de citas cruzadas entre autores con ideas tan afines como Feinberg, Heinrich y Rapoport, Sjordik y Steel, etc... Por otro lado, con alegría podemos observar al menos un par de trabajos que, (especulamos) motivados por nuestras publicaciones anteriores, han empezado a generar cooperaciones: Hordijk y cols. (2011); Kreyssig y cols. (2012)

## **Ubicuidad de los Lazos de Catalizaciones**

La existencia de al menos un lazo de catalizaciones en un sistema cRAO es un crucial detalle técnico que tiene la enorme virtud de abrir un camino para avanzar ya que este





**Figura 32:** Mapa conceptual de esta tesis. Esta tesis usó herramientas matemáticas simples, más un análisis de la historia del noción de *clausura metabólica* para deducir un resultado principal (*cLST*) y varios resultados secundarios que también son importantes.

resultado muestra que en la clase de sistemas con clausura tipo COT y RAF **debe** existir al menos un lazo de catalizadores y no un mero lazo de reacciones cíclicas. Dado que un lazo de catalizaciones está implícito en los sistemas  $(M, R)$  ( $\beta$  se trata de eso) y, como demostraron Letelier y cols. (2003), los sistemas Autopoiéticos son un sub-conjunto de los sistemas  $(M, R)$ , estos patrones de conectividad (los lazos) son comunes y efectivamente ubicuos en las teorías de lo vivo. Por lo tanto, nuestro teorema sobre la existencia de un lazo de catalización implica una característica universal de los modelos de seres vivos. Es interesante subrayar que en la definición clásica de sistemas autopoiéticos no hay nada que presagie la ocurrencia de este tipo de lazo.

### **Los Lazos *controlan* todo**

Un resultado importante, que en esta tesis no fue explorado a profundidad, es el descubrimiento de que en un sistema cRAO, un sistema que por lo tanto cuenta con al menos un lazo de catalizaciones, *cualquier* reacción del sistema tiene su catalizador controlado por un catalizador de un lazo. Este resultado es interesante por dos puntos:

- Muestra que los lazos (de catalizaciones) ejercen control más allá de las reacciones del lazo. Perseguir esta idea puede ser muy fecunda ya que mostraría como la red metabólica completa actúa como un todo integrado
- Vuelve a indicar que, para que se dé un esquema de clausura, es necesario que algunos catalizadores tengan dualidad de funciones (*moonlighting effect*) (Letelier y cols., 2006; Piedrafita y cols., 2012; Cornish-Bowden y Cardenas, 2007).

Estos dos teoremas, en conjunción con el *cLST* resultan poderosamente sugerentes pues parecen indicar a una de las fuentes primarias de estabilidad en los organismos vivos ya que se hace patente que los lazos están presentes en una gran clase de modelos de los organismos y su dinámica (aparentemente estabilizadora) unge a todo el sistema.



## Estado estacionario

El *cLST* es válido para los estados estacionarios, cuya existencia he utilizado como premisa. La condición de existencia en términos técnicos se traduce a que se respete la siguiente ecuación:

$$\prod_{i=1}^N a_i = \prod_{i=1}^N b_i \quad (114)$$

Es notable que esta ecuación exige un balance o equilibrio entre la entrada y salida de masa al lazo, sin embargo no es un balance de masa en el sentido tradicional sino que la ecuación hace referencia a la media geométrica de los flujos de entrada y salida. Este curioso hecho, junto con consideraciones relativas a la plausibilidad de satisfacción de la condición deben ser investigadas más a fondo.

## Variedad central

El *cLST* implica estabilidad para los lazos aislados pero que sucede con este resultado si la configuración de reacciones difiere en algunos elementos, es decir, la micro estructura local de la red cambia? Esta es una pregunta natural a la hora de entender la dinámica de los lazos insertos en sistemas más grandes. Estos cambios hacen que el corazón formal del modelo (la matriz estequiométrica y el Jacobiano) cambie y por lo tanto hay que re-analizar cuales son los estados estables. Lo interesante es usar técnicas matemáticas sofisticadas para ver cómo cambian los estados estables cuando se producen estos cambios pequeños en sistema bajo estudio. En esta dirección, la teoría de la variedad central nos permite analizar el comportamiento del sistema bajo la introducción de no-linealidades. Si bien esta avenida no fue visitada a profundidad en este trabajo, nos sirve como ilustración de como se puede seguir expandiendo los resultados expuestos en esta tesis.

## Importancia del *cLST*

La gran novedad del *cLST* es que reúne bajo un mismo techo las nociones de clausura, de causas eficientes y de estabilidad en el sentido de sistemas dinámicos. Este es el aspecto por el cual creemos que el *cLST* tendrá un potente impacto en la comunidad de investigadores interesados en las teorías sobre los sistemas vivos, pues permite hablar en un lenguaje unificado de elementos que existían anteriormente en universos disjuntos. En efecto, una lectura cuantitativa del *cLST* parece indicar que los lazos de catalizaciones no son meras mejoras, sino que generan estados estables. De esta manera nos acercamos al objetivo inicialmente establecido de poder explicar algo fenomenológico de los organismos vivos a partir de primeros principios.

Siendo el *cLST* es el principal resultado de esta tesis es obligatorio preguntarse: ¿Resistirá el embate de otros investigadores y será útil para avanzar? A continuación haré un sobrevuelo de las conexiones entre *cLST* y otras ideas para poder decidir si mi entusiasmo es justificado o no. Este análisis es necesario para poder generar un entendimiento tanto de los aspectos matemáticos como biológicos tras el *cLST*. Primero debemos dejar en claro que, comparado con muchos de los trabajos actuales, las técnicas usadas para demostrar el *cLST* son simples y son las mismas usadas por Christine Reder en su re-formulación del MCA (Reder, 1988). También son las técnicas usadas en lo que se entiende por ingeniería metabólica. Quiero enfatizar aquí que las ideas matemáticas usadas no representan la cúspide de ideas como bases de Grobner, esas ideas quedan abiertas para que otros las entiendan y las usen. Por lo tanto nuestra innovación no es primariamente matemática, pero conceptual en el dominio de la biología.

La gran diferencia en la demostración del *cLST* con **toda** la literatura en biología teórica es que nosotros pusimos en el mismo pie a los productos y los catalizadores, es decir, no hacemos distinciones entre tipos de moléculas. Esta idea era perfectamente tratable en

1988 (reformulación de MCA por Reder) y en 1973 (MCA), la razón por la cual no se hizo fue por no considerar a las enzimas como pertenecientes al sistema metabólico, probablemente porque el enfoque de ese tiempo era distinto y las nociones sistémicas y de clausura, aunque presentes en el *Zeigeist* de hace 40 años, no estaban maduras aún.

Obviamente que hay muchas relaciones del *cLST* con una multitud de resultados previos. No por nada el problema que ataca el *cLST* es un problema reconocido como tal por muchas personas desde hace décadas. Así, no nos debe sorprender que haya superposición conceptual con ideas del CRNT o de Systems Biology. A continuación haré una revisión de las similitudes y diferencias entre el trabajo de estos grupos y nuestro resultado de estabilidad.

## **Relación con resultados previos**

### **Estabilidad en otras teorías de clausura metabólica**

Como he repetido muchas veces en esta tesis, existen dos narrativas o tradiciones para hablar de estabilidad en redes de reacciones. Una tradición esta basada en lo estrictamente técnico y cuya aproximación parte de la consideración de flujos y matrices estequiométricas (CRNT, SNA, MCA, Systems Biology, Ingeniería metabólica). La otra tradición, más entretenida o pintoresca pero menos precisa técnicamente, es la que trata de entender a los sistemas vivos con principios generales de la organización del metabolismo (Cibernética, Sistemas  $(M, R)$ , Autopoiesis, Redes Auto catalíticas, Conjuntos RAF, COT, Chemotón, Hiperciclos, etc...). Respecto a esta última narrativa o *Weltanschauung* nuestro trabajo se posiciona de manera muy particular.

**Sistemas  $(M,R)$ .** Por ejemplo, en el esquema de Rosen no existe el tiempo, la estabilidad es perfecta, constante y algebraica. Una vez es alcanzado el esquema de invarianza organizacional, es decir, aparece en operación el misterioso operador  $\beta$ , el sistema  $(M, R)$

es como un cristal de estabilidad. Nuestro enfoque es muy diferente, pero contiene un elemento central de la teoría de los sistemas  $(M, R)$ : la noción de causa eficiente capturada por la presencia de las enzimas y su clausura. El *cLST* provee un paso esencial para esclarecer esta teoría al hacer más explícita la siguiente asociación:

$$\beta \Rightarrow \text{Estabilidad}$$

El *cLST* establece esta correspondencia pues el avatar primero de  $\beta$  es la idea de lazo bajo causas eficientes. Esto no está mal. Pero, aunque lo intentamos, nuestro análisis no ilumina mucho sobre que debe ser  $\beta$ . Por supuesto que este misterioso operador es irreducible y no puede esperarse establecer una relación tipo 1:1 con la presencia de lazos, pero si creo que los lazos son una alegoría más que válida para empezar a desenmarañar la naturaleza de  $\beta$ . Otra consecuencia de los resultados de este trabajo en la teoría de los sistemas  $(M, R)$  es clarificar qué se debe entender por *invarianza organizacional*: se puede entender que cada sistema  $(M, R)$  debe tener siempre un lazo de catalizaciones que le confiere estabilidad.

**Sistemas Autopoiéticos.** Los sistemas autopoiéticos, que de acuerdo a Letelier y cols. (2003) son un sub-conjunto de los sistemas  $(M, R)$ , tampoco tienen un tiempo explícito y tampoco ecuaciones de movimientos (en el sentido de un sistema dinámico subyacente). Son objetos algebraicos y a través de argumentos cualitativos se asegura que, debido a su autonomía, deben poseer algún grado de estabilidad, pero que pueden existir varios tipos y modos de obtener y mantener esta estabilidad. Esto queda subsumido en la frase: *cambio de estructura manteniendo la organización*. Por lo tanto el *cLST* tiene dos puntos de contacto con las ideas de Maturana y Varela: Primero, muestra el origen de la estabilidad y, en segundo lugar, gracias a las ideas de la variedad central sugiere un camino para entender el cambio estructural con invarianza de la organización. Este es un camino muy promisorio que se debe investigar a futuro. Especialmente en situaciones distintas del metabolismo,

el formalismo expresado por el *cLST* se refiere a *motifs* específicos de relaciones entre procesos donde se pueda hablar de: causas eficientes, producción y destrucción de componentes. En nuestro caso el *motif* es el lazo de catalizaciones.

**Conjuntos RAF & COT.** Esta tesis sólo se puede entender debido a la existencia de los trabajos sobre conjuntos RAF y COT. Ambas ideas contribuyeron a clarificar como se debe hablar de clausura y metabolismo. No por nada, previo al *cLST*, tenemos un resultado que define un nuevo tipo de sistemas (cRAO) como los sistemas que tienen las características de los sistemas vivos. Ambas teorías, a su manera, ya tienen teoremas de estabilidad (o existencia de redes estables). Mi contribución es hacer este aspecto más claro y poner el asunto en el andamiaje de una técnica aceptada por todos (sistemas dinámicos).

Observemos que en este caso ocurrió una situación curiosa. Una versión preliminar de estas ideas fue publicada en la conferencia europea de vida artificial (Contreras y cols., 2011), a consecuencia de eso las ideas de lazos de catalizaciones llegaron al grupo de Dittrich y se generó ya una publicación conjunta que trata sobre como los *lazos o ciclos* generan estabilidad (Kreyssig y cols., 2012). Claro que en esa ocasión se estudiaron los lazos como fenómeno general en biología y no se pudo concretar una demostración. El *cLST* se pudo demostrar de forma general y clara ya que parte del hecho de la clausura.

## **Resultados de estabilidad en el linaje de CRNT y SNA**

La narrativa de la ingeniería química/metabólica tiene desde mediados de los 70s<sup>1</sup> potentes teoremas de estabilidad y de relación entre flujos que desgraciadamente no se enseñan en pregrado ni en biología o en química. Estos teoremas, especialmente el *ZDT*, son el modelo en los cuales un investigador debiese inspirarse para tratar de entender el metabolismo como un todo integrado<sup>2</sup>.

---

<sup>1</sup>hace ya casi 40 años!

<sup>2</sup>Al hacer esta tesis, y considerando la Autopoiesis como idea nació en la Facultad de Ciencias, me di cuenta de lo útil que sería enseñar algunas de estas ideas en algunos cursos de pregrado

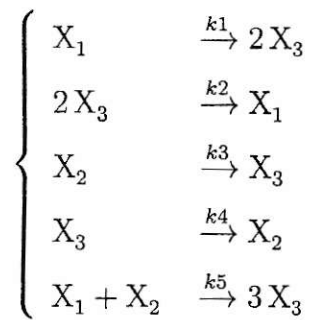
## **El *ZDT* y sus Derivados**

Desafortunadamente la *disposición escénica* del *ZDT* de Feinberg es inadecuada para el estudio de las redes metabólicas. La idea de centrarse en complejos en vez de moléculas, y el pedir que los complejos tengan sucesivas reacciones entre ellos hace que el *ZDT* se aplique más a mezclas químicas que a intrincadas redes metabólicas. Además, como ya expliqué, el metabolismo está muy lejos de cumplir con la exigencia de tener deficiencia igual a cero. El *cLST* es más simple que el *ZDT* ya que la conectividad implícita en el *cLST* es de un solo tipo (es decir el *lazo* de la figura 16). En cambio el *ZDT* se refiere a muchas situaciones potencialmente muy distintas que sólo comparten el hecho de tener deficiencia igual a zero. Por eso la demostración del *cLST* es más fácil que la del *ZDT*, la situación modelada es más acotada y, según el teorema de la existencia del lazo, más universal en las redes metabólicas reales.

## **Los Teoremas de Estabilidad de Clarke**

Bruce Clarke, hacia 1975, configuró una teoría distinta a la CRNT para estudiar la estabilidad de redes de reacciones. Este enfoque llamado (Stoichiometric Network Analysis o SNA) es más general que el enfoque del *ZDT* (Clarke, 1988). En SNA se habla de la matriz estequiométrica, de los flujos de cada reacción y de los estados estacionarios. Hay varios resultados sobre estabilidad, que siempre se relacionan con micro-ejemplos muy específicos de conectividad. Por ejemplo, usando ideas de la SNA en el llamado mecanismo de Langmuir-Hinshelwood que se encuentra a continuación se ha llegado a demostrar la existencia de estados estacionarios:

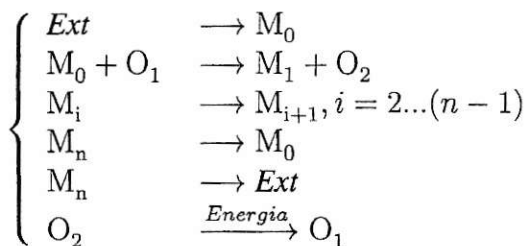




Pero, sin ánimo de desmerecer este trabajo, la aplicabilidad de sus resultados parece limitarse a sistemas pequeños. En fondo las técnicas del SNA son similares a las nuestras (sistemas dinámicos y uso inteligente de la matriz Jacobiana) y en casos específicos se pueden demostrar situaciones de estabilidad. Obviamente que en el caso de la CRNT o del SNA no hay noción de catálisis, en ese sentido son teorías muy distintas a nuestro modelo.

### **Teoremas de Estabilidad, multiestabilidad e inyectividad posteriores al 2000**

Hasta el año 2000 la tasa de publicación en temas de CRNT o SNA era muy limitada pero, posteriormente, los cultivadores del *Systems Biology* comenzaron a descubrir a Feinberg, Clarke y otros, comenzando a aplicar estos teoremas/resultados directamente al metabolismo. Es difícil, pero se ha comenzado a hacer. En (Brogioli, 2013) se puede ver una aplicación moderna de las ideas de Feinberg; la idea es encontrar metabolismos (siempre son pequeños en esta teoría) donde no sólo exista un único estado estable, sino que una variedad, es decir, un sub-conjunto de un espacio topológico que es localmente parecido a  $\mathbb{R}^n$ ). Lo que se prueba en este trabajo es que existen situaciones *marginamente estables*, es decir, que si se modifican algunos parámetros (por razones externas) el sistema cambia de un estado a otro pero permaneciendo en una nueva posición de estado estacionario. Por lo tanto si se siguen produciendo esas variaciones externas de parámetros el sistema exhibe una trayectoria que es una deriva entre estados estacionarios. Pero esta no es una



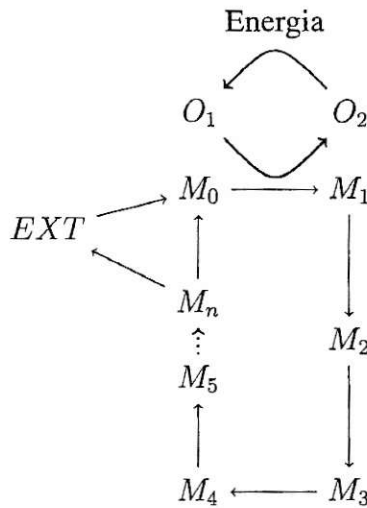
**Figura 33:** Conjunto de reacciones cíclicas, acopladas a una reacción impulsora. Este Metabolismo tiene un estado estable para cualquier valor de sus parámetros cinéticos.

propiedad general, sólo la exhiben algunas redes de reacciones.

### Un teorema de Estabilidad en un Lazo muy particular

En el campo de la teoría de las redes metabólicas (y de la ingeniería metabólica) hay una verdadera fiebre por encontrar **EL** principio que *hace andar* a las redes. Esta búsqueda lleva décadas, al menos 40 años, y aún no se encuentra una solución satisfactoria. Han existido muchas tentativas que se han centrado en dos líneas de pensamiento: a) la línea de la teoría fundamental y b) el enfoque de la ingeniería (Feinberg, Clarke, Kacser). Obviamente este último ha sido como un hongo atómico... Con resultados cada vez más técnicos y más interesantes. Todos quieren probar una generalización del *ZDT* para reacciones metabólicas. Para ver cuan atingentes son nuestros resultados mencionaré aquí un trabajo reciente (Reznik y cols., 2013) que justamente se centra en estudiar la estabilidad de redes de reacciones del siguiente tipo:

Este es, como el nuestro, un metabolismo cíclico muy simple... Donde no existen reacciones catalizadas. Los autores están, como nosotros, interesados en obtener algún resultado de estabilidad. La matriz estequiométrica de este sistema (las filas representan los metabolitos) es:



**Figura 34:** Sistema cíclico de reacciones analizado por *SKM*. Lazo de reacciones con clausura y con una reacción energética que impulsa el lazo

$$S = \begin{bmatrix} 1 & -1 & 0 & \dots & 1 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & -1 & 0 & \dots & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & -1 \end{bmatrix} \quad (115)$$

Para avanzar los autores usan una técnica, no mencionada en esta tesis, *Structural Kinetic Modeling* que trata de obtener resultados de estabilidad sin fijarse en las variables cinéticas (Steuer y cols., 2006; Grimbs y cols., 2007). Para ello se usa un procedimiento de normalización, que no explicaré, pero que permite expresar la matriz  $J$  (es decir el Jacobiano del sistema de ecuaciones diferenciales) como:

$$J = \Lambda \cdot \Theta \quad (116)$$

Donde la matriz  $\Lambda$  es una renormalización de la matriz  $S$ , pero –y aquí viene lo más

interesante del enfoque de SKM– la matriz  $\Theta$  contiene las *elasticidades relativas*: el elemento  $\Theta_{i,j}$  es la sensibilidad de la velocidad normalizada de la reacción  $i$  respecto de la concentración normalizada del metabolito  $j$ . SKM es un enfoque que se centra en entender  $J$  entendiendo  $\Theta$ , la cual se escribe para el sistema presentado de ejemplo como:

$$\Theta = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ \theta_1 & 0 & 0 & \dots & 0 & \theta_{n+2} \\ 0 & \theta_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & \theta_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \theta_{n+1} & 1 \\ 0 & 0 & 0 & \dots & 0 & t\theta_{n+3} \end{bmatrix} \quad (117)$$

Como se puede ver  $\Theta$  se parece, pero no es idéntica a nuestro Jacobiano. Interesantemente en este trabajo los autores elaboran un potente teorema que:

- Muestra que el *toy metabolism* analizado **para cualquier valor de constantes cinéticas, de concentraciones iniciales de metabolitos o de mecanismos de reacción** siempre tiene un equilibrio estable.
- Requiere del uso de una matemática absolutamente no trivial y hasta deben deducir un teorema (*The Stubborn Roots Theorem*) bastante abracadabrante de álgebra lineal. Este es un ejemplo de como problemas en biología iluminan nuevos desarrollos en matemáticas.
- Finalmente los autores afirman que, aunque ellos no la usan, la noción de clausura metabólica (*closure*) puede jugar un papel importante para entender la estabilidad de los *metabolic motifs*.

# Conclusión

En definitiva, el aporte fundamental de este trabajo consiste en un conjunto de teoremas y definiciones que permiten:

- Contar con una descripción formal de los sistemas vivos
- Formalizar la idea de clausura metabólica
- y por lo tanto utilizarla para derivar resultados técnicos
- Demostrar la existencia de patrones universales en los organismos vivos
- Indicar la importancia de estos patrones para la existencia de los organismos

Como se pudo ver en el capítulo anterior, el *cLST* tiene muchos antecedentes, desde el *ZDT* de 1974 hasta los resultados de Brogioli (2011) o Reznik y cols. (2013). Esto no nos debe extrañar: todo el ecosistema de la biología teórica está buscando resultados sobre la estabilidad de vías metabólicas. Además, Kitano ha dicho que la idea que se debe perseguir es la noción de *robustness* que es, desde nuestra perspectiva, la multiestabilidad vista con el cristal de técnicas como la teoría de la variedad central. Interesantemente, debido al esfuerzo de Kitano y otros, existe un ramillete de resultados que son intrigantes y que anuncian la llegada de resultados más universales sobre estabilidad, multiestabilidad y robustez (Kitano, 2007).

Aunque la dirección está clara, todavía contamos con pocos resultados generales. Por ejemplo el *ZDT* sólo se aplica en condiciones muy especiales, y los sistemas marginalmente estables sólo se encuentran en algunas configuraciones de reacciones. En este sentido el *cLST* es un representante más en un linaje de teoremas parecidos, y a la vez un disruptor de este linaje por el hecho de considerar a los catalizadores en el esquema de

reacciones. Este cambio de enfoque o salto conceptual nos permite abarcar una clase mucho más general de sistemas. En el *cLST* se han incorporado a los catalizadores por una necesidad teórica fundacional, introducir la idea de clausura, no por una necesidad técnica. Interessantemente esta incorporación trae de la mano (o hace aparecer) un tipo bastante específico de estabilidad que depende de una característica global de la conectividad. Creo que aquí está la diferencia mayor entre mi análisis y todas las tentativas previas. En todos los resultados de estabilidad, comenzando con el *ZDT*, se analizan ejemplos muy concretos de pocas reacciones y donde la estabilidad proviene de la existencia de lazos con un número muy pequeño de componentes. Por ejemplo, el trabajo de Brogioli y cols. analiza 6 redes de a lo más 8 reacciones cada una. Nuestro enfoque usa una propiedad global: que exista un lazo de catalizaciones, para obtener un resultado. En este aspecto tampoco somos los primeros ya que Reznik y cols. (2013) también usan un lazo de largo arbitrario, pero sin considerar a los catalizadores.

La gran diferencia entre mi resultado y el resto reside en el punto de partida. Al abrazar, como idea rectora, la noción de clausura nos comprometemos con un enfoque específico que permite prestar atención a una propiedad global considerando siempre las causas eficientes. Así, en cierta manera la clausura (o el argumento de la clausura) actúa como un atajo para obtener resultados generales de estabilidad. Las ideas del *cLST* son un primer paso para estudiar, usando un enfoque técnico y no meras intuiciones cualitativas, cómo es que la estabilidad, la multiestabilidad y la robustez se deducen de la clausura. Es posible que comprendiendo estas ideas podamos, en un futuro quizás lejano, entender las redes metabólicas a tal punto que incluso las podamos producir con propiedades diseñables.

Finalmente, con gran satisfacción me gustaría decir que no solamente he cumplido con los objetivos de esta tesis, sino que además hemos establecido un programa que nos permite formalizar intuiciones, derivar propiedades sistémicas y con ello profundizar nuestra comprensión de los organismos vivos.

# Bibliografía

- Arita, M. 2004. The metabolic world of escherichia coli is not small. *Proceedings of the National Academy of Sciences of the United States of America*, 101(6):1543–1547.
- Battley, E. H. 1993. Calculation of entropy change accompanying growth of escherichia coli k-12 on succinic acid. *Biotechnology and bioengineering*, 41(4):422–428.
- Benvenuti, L. y Farina, L. 2004. Eigenvalue regions for positive systems. *Systems & Control Letters*, 51(3–4):325 – 330.
- Brogioli, D. SEP 29 2011. Marginal stability in chemical systems and its relevance in the origin of life. *Physical Review E*, 84(3, Part 1).
- Brogioli, D. 2013. Looking for chemical reaction networks exhibiting a drift along a manifold of marginally stable states. *Journal of Theoretical Biology*, 318:110–123.
- Cartwright, J. H., García-Ruiz, J. M., Novella, M. L., y Otálora, F. 2002. Formation of chemical gardens. *Journal of Colloid and Interface Science*, 256(2):351 – 359.
- Castellan, G. 1987. Fisicoquímica addison-wesley, reading, mass., 1964. Addison Wesley Longman, México.
- Centler, F. y Dittrich, P. 2007. Chemical organizations in atmospheric photochemistries — a new method to analyze chemical reaction networks. *Planetary and Space Science*, 55(4):413 – 428.
- Centler, F., Kaleta, C., Speroni di Fenizio, P., y Dittrich, P. 2010. A parallel algorithm to compute chemical organizations in biological networks. *Bioinformatics*, 26(14):1788 – 11789.
- Clarke, B. 1988. Stoichiometric Network Analysis. *Cell Biophysics*, 12:237–253.
- Coatman, R., Thomas, N., y Double, D. 1980. Studies of the growth of “silicate gardens” and related phenomena. *Journal of Materials Science*, 15(8):2017–2026.
- Contreras, D., Pereira, U., Hernández, V., Reynaert, B., y Letelier, J. C. A loop conjecture for metabolic closure. In *Advances in Artificial Life, ECAL 2011: Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems*. MIT Press, 2011.
- Cornish-Bowden, A. y Cardenas, M. L. Oct 2007. Organizational invariance in  $(M, R)$ -systems. *Chemistry & Biodiversity*, 4(10):2396–2406.
- de La Mettrie, J. y Thomson, A. 1996. La mettrie: Machine man and other writings. Cambridge University Press.
- Devlin, T. M. 2004. Bioquímica: libro de texto con aplicaciones clínicas. Reverté, Barcelona, España.
- Dittrich, P. y di Fenizio, P. S. 2007. Chemical organisation theory. *Bulletin of Mathematical Biology*, 69(4):1199–1231.

- Eigen, M. y Schuster, P. 1978. The hypercycle. a principle of natural self organization. part c: The realistic hypercycle. *Naturwissenschaften*, 65(7):341–369.
- Feinberg, M. 1987. Chemical reaction network structure and the stability of complex isothermal reactors—i. the deficiency zero and deficiency one theorems. *Chemical Engineering Science*, 41 (10):2229–2268.
- Fell, D. A. September 1992. Metabolic control analysis: a survey of its theoretical and experimental development. *The Biochemical Journal*, 286 ( Pt 2):313–330.
- Ganti, T. 2003. Chemoton theory: Theory of living systems. Springer-Verlag GmbH.
- García-Ruiz, J. M., Melero-García, E., y Hyde, S. T. 2009. Morphogenesis of self-assembled nanocrystalline materials of barium carbonate and silica. *Science*, 323(5912):362–365.
- Gillespie, D. T. 1977. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340 – 2361.
- Grimbs, S., Selbig, J., Bulik, S., Holzhuetter, H.-G., y Steuer, R. 2007. The stability and robustness of metabolic states: identifying stabilizing sites in metabolic networks. *Molecular Systems Biology*, 3.
- Gulberg, C. M. y Waage, P. 1864. Studies concerning affinity. *Forhandlinger: Videnskabs-Selskabet i Christiana*, 35:1044+.
- Hearon, J. Z. 1953. The kinetics of linear systems with special reference to periodic reactions. *The bulletin of mathematical biophysics*, 15(2):121–141.
- Heinrich, R. y Rapoport, T. A. 1974. A linear steady-state treatment of enzymatic chains. general properties, control and effector strength. *European Journal of Biochemistry*, 42:89–95.
- Higgins, J. 1963. Analysis of sequential reactions. *Annals of the New York Academy of Sciences*, 108(1):305–321.
- Hordijk, W. y Steel, M. 2004. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227(4):451–461.
- Hordijk, W., Kauffman, S. A., y Steel, M. MAY 2011. Required Levels of Catalysis for Emergence of Autocatalytic Sets in Models of Chemical Reaction Systems. *International Journal of molecular Sciences*, 12(5):3085–3101.
- Hyver, C. 1972. Impossibility of existence of undamped oscillations in linear chemical systems. *Journal of Theoretical Biology*, 36(1):133 – 138.
- Jaramillo, S., Honorato-Zimmer, R., Pereira, U., Contreras, D., Reynaert, B., Hernández, V., Soto-Andrade, J., Cárdenas, M., Cornish-Bowden, A., y Letelier, J. 2010. (M, R) Systems and RAF sets: Common ideas, tools and projections. *Proceedings of the Twelfth International Conference on the Synthesis and Simulation of Living Systems*, pages 94 – 100.



- Kacser, H. y Burns, J. 1973. The control of flux. *Symposia of the Society for Experimental Biology*, 27:65–104.
- Kajander, E. O. y Çiftçioglu, N. 1998. Nanobacteria: An alternative mechanism for pathogenic intra- and extracellular calcification and stone formation. *Proceedings of the National Academy of Sciences*, 95(14):8274–8279.
- Kauffman, S. 1993. *The Origins of Order*. Oxford University Press, New York.
- Kauffman, S. A. March 1986. Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119 (1):1–24.
- Kitano, H. 2001. *Foundations of systems biology*. MIT Press.
- Kitano, H. 2007. Towards a theory of biological robustness. *Molecular Systems Biology*, 3:137.
- Kreyssig, P., Escuela, G., Reynaert, B., Veloz, T., Ibrahim, B., y Dittrich, P. OCT 11 2012. Cycles and the Qualitative Evolution of Chemical Systems. *PLoS ONE*, 7(10).
- Langton, C. y Santa Fe Institute (Santa Fe, N. 1989. *Artificial life i: The proceedings of an interdisciplinary workshop on the synthesis and simulation of living systems, held september, 1987, in los alamos, new mexico*. Addison-Wesley Publishing Company.
- Leduc, S. 1912. *La biologie synthetique*. A. Poinat.
- Letelier, J. C., Marín, G., y Mpodozis, J. 2003. Autopoietic and  $(M, R)$  systems. *Journal of Theoretical Biology*, 222(2):261 – 272.
- Letelier, J.-C., Kuboyama, T., Yasuda, H., Cárdenas, M.-L., y Cornish-Bowden, A. 2005. A selfreferential equation,  $f(f)=f$ , obtained using the theory of  $(M, R)$ -systems: overview and applications. *Algebraic Biology*, pages 115–126.
- Letelier, J. C., Soto-Andrade, J., Guíñez Abarzúa, F., Cornish-Bowden, A., y Cárdenas, M. L. 2006. Organizational invariance and metabolic closure: analysis in terms of  $(M, R)$  systems. *Journal of Theoretical Biology*, 238:949–961.
- Letelier, J. C., Cárdenas, M. L., y Cornish-Bowden, A. 2011. From *L Homme Machine* to metabolic closure: steps towards understanding life. *Journal of Theoretical Biology*, 286:100–113.
- Lifson, S. 1997. On the crucial stages in the origin of animate matter. *Journal of Molecular Evolution*, 44(1):1–8.
- Luisi, P. L. 2003. Autopoiesis: a review and a reappraisal. *Naturwissenschaften*, 90(2):49–59.
- Luisi, P. L. 2006. *The emergence of life: From chemical origins to synthetic biology*. Cambridge University Press.
- Luisi, P. L. 2012. Comments on Prebiotic Ecology, Supramolecular Selection and Autopoiesis. *Origins of life and Evolution of Biospheres*, 42(5):451–452.

- Matsumaru, N., Centler, F., di Fenizio, P. S., y Dittrich, P. 2005. Chemical organization theory as a theoretical base for chemical computing. *International Journal of Unconventional Computing*, 3:285 – 309.
- Matsumaru, N., Centler, F., di Fenizio, P. S., y Dittrich, P. 2006. Chemical organization theory applied to virus dynamics. *Information Technology*, 48:154 – 160.
- Maturana, H. y Varela, F. 1980. *Autopoiesis and cognition: The realization of the living*. D. Reidel Publishing Company.
- Maturana, H. y Varela, F. 1973. *De máquinas y seres vivos: Una teoría sobre la organización biológica*. Editorial Universitaria, Santiago de Chile.
- McMullin, B. y Program, S. E. 1997. *Computational autopoiesis: The original algorithm*. Santa Fe Institute.
- McMullin, B. 2004. 30 years of computational autopoiesis: A review. *Artificial Life*, 10(3):277–295.
- Mitkowski, W. 2008. Dynamical properties of Metzler systems. *Bulletin of the polish Academy of Sciences-Technical Sciences*, 56(4):309–312.
- Peter, S. y Dittrich, P. 2011. On the relation between organizations and limit sets in chemical reaction systems. *Advances in Complex Systems*, 14(01):77 – 96.
- Peter, S., Veloz, T., y Dittrich, P. Feasibility of organizations - a refinement of chemical organization theory with application to p systems. In *Proceedings of the 11th international Conference on Membrane Computing*, pages 325 – 337, Berlin, Heidelberg, 2010. Springer-Verlag.
- Piedrafitra, G., Cornish-Bowden, A., Moran, F., y Montero, F. May 2012. Size matters: influence of stochasticity on the self-maintenance of a simple model of metabolic closure. *Journal Theoretical Biology*, 300:143–151.
- Raoult, D., Drancourt, M., Azza, S., Nappéz, C., Guieu, R., Rolain, J.-M., Fourquet, P., Campagna, B., La Scola, B., Mege, J.-L., Mansuelle, P., Lechevalier, E., Berland, Y., Gorvel, J.-P., y Renesto, P. 02 2008. Nanobacteria are mineralo fetuin complexes. *PLoS Pathogens*, 4(2):e41.
- Rashevsky, N. 1954. Topology and life: In search of general mathematical principles in biology and sociology. *The bulletin of mathematical biophysics*, 16(4):317–348.
- Rashevsky, N. 1960. *Mathematical biophysics: physico-mathematical foundations of biology*. Dover Publications.
- Razeto-Barry, P. 2012. Autopoiesis 40 years later. a review and a reformulation. *Origins of Life and Evolution of the Biosphere*, 42(6):543–567.
- Reder, C. 1988. Metabolic control theory: A structural approach. *Journal of Theoretical Biology*, 135:175 – 201.

- Reka, A. y Barabási, . June 2002. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97.
- Reznik, E., Watson, A., y Chaudhary, O. 2013. The stubborn roots of metabolic cycles. *Journal of the Royal Society, Interface*, 10(83).
- Rietman, E. A., Colt, J. Z., Tuszynski, J. A., y others, . 2011. Interactomes, manufacturomes and relational biology: analogies between systems biology and manufacturing systems. *Theoretical Biology and Medical Modelling*, 8(1):19.
- Rosen, R. 1991. Life itself: A comprehensive inquiry into the nature, origin, and fabrication of life. Columbia University Press.
- Rosen, R. 1958a. A relational theory of biological systems. *The Bulletin of Mathematical Biophysics*, 20(3):245–260.
- Rosen, R. 1958b. The representation of biological systems from the standpoint of the theory of categories. *Bulletin of Mathematical Biology*, 20(4):317–341.
- Rosen, R. 1972. Foundations of mathematical biology. Academic Press.
- Rosen, R. 2006. Autobiographical reminiscences of robert rosen. *Axiomathes*, 16(1-2):1–23.
- Rozenberg, S. A., G. 1980. The Mathematical Theory of L Systems. Academic Press, New York.
- Savageau, M. A. 1976. Biochemical Systems Analysis: A Study of Function and Design in Molecular Biology. Addison-Wesley Publishing Co., Reading, MA.
- Schnell, S. y Mendoza, C. 1997. Closed form solution for time-dependent enzyme kinetics. *Journal of Theoretical Biology*, 187(2):207 – 212.
- Schrödinger, E. 1944. What is life?: The physical aspect of the living cell. The University Press.
- Schumacher, P. 2011. The autopoiesis of architecture: A new framework for architecture. Wiley.
- Spencer-Brown, G. 1972. Laws of form. Julian Press.
- Steuer, R., Gross, T., Selbig, J., y Blasius, B. 2006. Structural kinetic modeling of metabolic networks. *Proceedings of the National Academy of Sciences of the United States of America*, 103(32):11868–11873.
- Teubner, G. 1988. Autopoietic law: A new approach to law and society. Walter de Gruyter.
- Torres, N. V., Mateo, F., Meléndez-Hevia, E., y Kacser, H. 1986. Kinetics of metabolic pathways. a system in vitro to study the control of flux. *Biochemical Journal*, 234(1):169–174.
- Trevors, J. T. 2011. Origin of microbial life: Nano- and molecular events, thermodynamics/entropy, quantum mechanisms and genetic instructions. *Journal of Microbiological Methods*, 84(3):492–495.
- Urbano, P. y Urbano, F. 05 2007. Nanobacteria: Facts or fancies? *PLoS Pathogens*, 3(5):e55.

- Varela, F. G., Maturana, H. R., y Uribe, R. 1974. Autopoiesis: the organization of living systems, its characterization and a model. *Curr Mod Biol*, 5(4):187–196.
- Varela, F. 1979. Principles of biological autonomy. North Holland.
- Varela, F. y Bourgine, P. 1992. Toward a practice of autonomous systems: Proceedings of the first european conference on artificial life. M.I.T. P.
- Varela, F. y Maturana, H. 1972. Mechanism and biological explanation. *Philosophy of Science*, 39 (3):378 – 382.
- Veloz, T., Reynaert, B., Rojas, D., y Dittrich, P. A decomposition theorem in chemical organizations. In *Advances in Artificial Life, ECAL 2011: Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems*. MIT Press, 2011.
- Von Neumann, J. y Burks, A. 1966. Theory of self-reproducing automata. University of Illinois Press.
- Wagner, A. y Fell, D. A. 2001. The small world inside large metabolic networks. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1478):1803–1810.
- Watson, J. D. y Crick, F. H. C. 1953. Genetical implications of the structure of deoxyribonucleic acid. *Nature*, 171(4361):964–967.
- Wiener, N. 1948. Cybernetics: or, control and communications in the animal and the machine. Hermann.
- Zhou, Y., Pearson, J. E., y Auerbach, A. 2005.  $\phi$ -value analysis of a linear, sequential reaction mechanism: Theory and application to ion channel gating. *Biophysical Journal*, 89(6):3680 – 3685.
- Zopf, G., Von Foerster, H., y of Naval Research, U. S. O. 1962. Principles of self-organization: transactions. Symposium Publications, Pergamon Press.

# Apéndice

Este apéndice detalla los trabajos en los que se fue gestando el resultado principal y las ideas circundantes de esta tesis. El primer manuscrito, **(M,R) Systems and RAF Sets: Common Ideas, Tools and Projections** muestra como los conjuntos *RAF* tienen herramientas analíticas interesantes para estudiar los sistemas  $(M, R)$  y los sistemas autopiéuticos fue publicado y presentado en el congreso ALIFE XII en 2010. Los siguientes dos trabajos, **A Decomposition Theorem in Chemical Organizations** y **A Loop Conjecture for Metabolic Closure** establecen la estrecha relación entre *COT* y los sistemas autopiéuticos y los conjuntos *RAF*, además de contener las ideas germinales que condujeron al *cLST*. Estos fueron presentados en el congreso ECAL 2011 celebrado en París en memoria de Francisco Varela. ALIFE y ECAL reúnen anualmente a toda la comunidad de la biología teórica. La escritura de estos trabajos motivó colaboraciones internacionales que culminaron con mi visita al laboratorio fundador de *COT*, Peter Dittrich, durante varias semanas del 2011, experiencia sumamente motivante y edificante. Durante esta visita se escribió el último trabajo que presento en este apéndice, **Cycles and the Qualitative Evolution of Chemical Systems**, el cual fue publicado en la revista *PLoS ONE*. Esta colaboración se ha mantenido activa, al punto que no sólo yo he visitado al grupo de Análisis de Biosistemas de la Universidad de Jena, si no que es posible que durante a fines de 2013 uno de sus investigadores principales, Florian Centler visite el laboratorio de Biología del Conocer. Además hay un manuscrito en preparación al *Journal of Theoretical Biology* donde se presentará el *cLST*. Por último deseo mencionar que como consecuencia de toda esta actividad se ha ido revelando un espacio común en el que biólogos y matemáticos de esta facultad podemos trabajar juntos.

# $(M,R)$ Systems and RAF Sets: Common Ideas, Tools and Projections

S. Jaramillo<sup>1</sup>, R. Honorato-Zimmer<sup>1</sup>, U. Pereira<sup>1</sup>, D. Contreras<sup>1</sup>, B. Reynaert<sup>1</sup>, V. Hernández<sup>1</sup>,  
J. Soto-Andrade<sup>1</sup>, M.L. Cárdenas<sup>2</sup>, A. Cornish-Bowden<sup>2</sup> and J.C. Letelier<sup>1</sup>

<sup>1</sup> Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile.

<sup>2</sup> Bioénergétique et Ingénierie des Protéines, CNRS, Marseille, France.

letelier@uchile.cl

## Abstract

There are deep underlying similarities between Rosen's  $(M,R)$  systems as a definition of life and the *RAF sets* (Reflexive Autocatalytic systems generated by a Food source) introduced by Hordijk and Steel as a way of analyzing autocatalytic sets of reactions. Using RAF concepts we have systematically explored the set of possible small idealized metabolic networks, searching for instances of  $(M,R)$  systems. This exhaustive search has shown that the central requirement of Rosen's framework, unicity of  $\Phi$ , becomes harder and harder to obtain as the network grows in size. In addition, we give an expression for operators  $f$ ,  $\Phi$  and  $\beta$  in terms of RAF sets.

## Introduction

Metabolic closure is easy to introduce *informally* but rather difficult to define. Although it is crucial for understanding living organization it was neglected until late in the 20th century. The rebirth of the scientific study of biological organization can be traced back to the 30-year period from 1958 to 1987, which saw the publication of several distinct perspectives on closure, including  $(M,R)$  systems (Rosen, 1958), the chemoton (Gánti, 1975), hypercycles (Eigen and Schuster, 1977), autopoiesis (Maturana and Varela, 1980), autocatalytic sets (Kauffman, 1986), and the first Artificial Life conference in Los Alamos in 1987 (organized by Christopher Langton). There was, however, an almost complete lack of cross-fertilization between the different schools of thought, with each theory developed with almost no reference to any of the others (Letelier et al., 2006; Cornish-Bowden et al., 2007; Cárdenas et al., 2010). The most extreme case of isolation is represented by Robert Rosen (1934-1998), who introduced the concept of  $(M,R)$  systems early in his career to represent biological metabolic networks. His isolation was aggravated by the intricate nature of his writings, in which biological ideas were mixed with abstract mathematics. Furthermore, he expressed his mathematical ideas in non-standard notations and without any effort to help the reader by giving examples or offering many needed clarifications.

In recent years, we have undertaken a systematic attempt to understand and explain the core notions of Rosen's the-

ory (Letelier et al., 2006). We have (a) clarified the relationship between  $(M,R)$  systems and autopoiesis (Letelier et al., 2003); (b) reframed Rosen's original formulation in terms of biochemical networks, with the introduction of the notion of "organizational invariance" for understanding Rosen's elusive mathematical operators (such as his  $\beta$ ); (c) made a clear distinction between  $(M,R)$  systems in general and  $(M,R)$  systems with organizational invariance, a notion that is only implicit in Rosen's writing (he confusingly called these "replicative"  $(M,R)$  systems); (d) given mathematical and biological examples of simple idealized systems that can be understood within Rosen's intellectual framework; (e) clarified how these notions can be used to explore the origin of living systems and how they should be used in the context of what has come to be called "systems biology". Finally, we have also shown how our formulation of  $(M,R)$  systems can shed light on the problem of the computability of living systems (Cárdenas et al., 2010). This short summary is intended simply to underline how fruitful Rosen's view of metabolic closure has become, and to explain why we feel that the boundaries of our knowledge can be pushed to qualitatively new grounds by continuing the exploration of his ideas.

The systematic absence of examples (whether mathematical or biological) from Rosen's work has always been problematical, especially of simple examples that can serve as heuristic devices for enhancing theoretical research. In this paper we address the two points outlined above by pointing out the close relationship between  $(M,R)$  systems and a recent theory of living organization based on what have been called RAF sets. We show how many examples of simple  $(M,R)$  systems can be found by a computer algorithm constructed on the model of RAF sets. We discuss how the technical tools originating in RAF sets can be used to enhance the research of  $(M,R)$  systems, and specifically we address the problem of the nature and unicity of Rosen's  $\Phi$  in the context of RAF sets.

## (M,R) systems

Rosen’s original formulation of  $(M,R)$  systems (Rosen, 1958), relied on a view of metabolism as a graph, and on a very abstract view of enzymes as functions (in the mathematical sense). The metaphor of metabolism as a graph, new in 1958, has subsequently been adopted by many people, without attribution to Rosen. The view of enzymes as functions has not attracted a wide following as Rosen’s formulation seems unnecessarily abstract, without bringing practical or theoretical benefits. He used this approach in order to be able to use *category theory* for framing his important intuition about metabolic closure. Although this demanding mathematical approach has some advantages, as described in our previous work, we shall not use it here as the fundamental ideas exposed by Rosen can be explained using set theory, and thereby become accessible to mainstream biologists.

Our analysis of  $(M,R)$  systems, together with our examples, shows that the crucial aspect to understand organizational invariance is to understand the nature of the equation

$$\Phi(b) = f$$

Here  $\Phi$  represents the aspect of biological organization that relates how catalysts are produced by the system. This equation seems to imply that a living system is organized in such a way that knowing  $b$  (right-hand side of biochemical equations) should be enough to unambiguously assign the catalysts (represented by  $f$ ) to the reactions in the network.

Rosen, moreover, requires that there be only way to carry out this assignment, i.e., that there is only one mapping  $\Phi$  such that  $\Phi(b) = f$ , a demanding assumption indeed. In other words, that we can reverse the procedure that gives  $f$  back from  $\Phi$ . The reverse procedure is Rosen’s  $\beta$ , so that

$$\beta(f) = \Phi$$

Mathematically,  $\beta$  is just the inverse of the “evaluation at  $b$ ” operator that evaluates every function at  $b$ . Biologically,  $\beta$  represents the mechanisms that specify how the process of creating catalysts is maintained over time, i.e., *organizational invariance*.

To clarify these notions, we created a small metabolic network where they can be embodied in actual molecules that implement the functions  $\Phi$  and  $\beta$  (Letelier et al., 2006).

### RAF sets

We now give a brief introduction to the work of Hordijk and Steel (2004), who constructed a formal framework to study autocatalytic systems. Their main aim appears to have been to expand Kauffman’s formalism about autocatalytic sets (Kauffman, 1993), to respond the criticisms that arose out of Kauffman’s assumptions. At the same time, their analysis developed interesting algorithms that handle this expanded

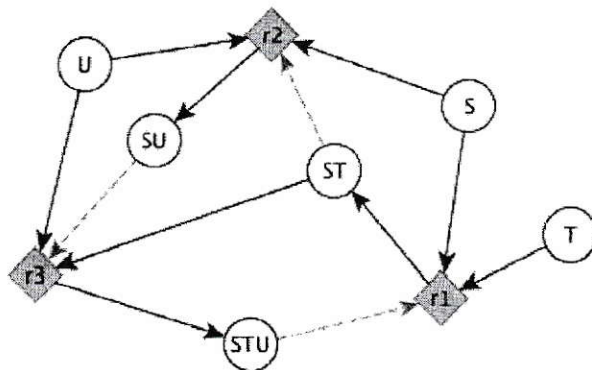


Figure 1:  $(M,R)$  system described by a *catalytic reaction graph*. Gray squares represent reactions and circles denote metabolites and enzymes. The black arrows represent chemical transformations while gray dashed arrows indicate catalyzations. This small network also contains a RAF set generated by the food set  $\{S, T, U\}$ .

framework. As a result, they have produced a powerful approach that can be used to analyze a wide variety of systems, and here we shall describe how it applies to  $(M,R)$  systems. Their formalism depends on the following two sets:  $X$ , the set of molecules involved in metabolism as metabolites, catalysts or external input material (termed *food* in the formalism), and  $\mathcal{R}$ , the set of reactions that defines the metabolic network.

Each reaction  $r$  is represented as a tuple  $(A, B)$ , where  $A, B \subset X$ ,  $A \cap B = \emptyset$ ,  $A$  are the reactants and  $B$  the products of reaction  $r$ . This formalism is similar to Rosen’s treatment of enzymes as transformations between two sets of molecules.

Further, to formalize the notion of catalysis, a specific set  $C$  (called the set of “catalyzations” by Hordijk and Steel), is introduced. Each catalyzation  $c$  is a tuple  $(x, r)$ , where  $x \in X$  is the catalyst and  $r \in \mathcal{R}$  is the reaction catalyzed by  $x$ . The similarity with Rosen (1958) is evident, as any given catalyzation  $c = (x, r)$  can be rewritten as  $c = (x, r) = (x, (A, B)) = (A, x, B)$ , making transparent the fact that molecule  $x$  catalyzes the reaction  $A \rightarrow B$ .

With the set of catalyzations defined, Mossel and Steel (2005) introduced a function  $\gamma$  that helps to simplify formulae in later sections:

$$\gamma_C(A, r) = \begin{cases} 1 & \text{if } \exists x \in A : (x, r) \in C, \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Additionally, a specific subset of  $X$  containing every molecule that is used but not produced by the metabolism is denoted  $F$  and it represents the food molecules.

Thus a *catalytic reaction system* over a food source  $F$  is composed by a triplet  $\mathcal{L} = (X, \mathcal{R}, C)$  that defines the universe of molecules ( $X$ ), the reactions occurring among these

molecules ( $\mathcal{R}$ ) and the identity of the catalyst involved in each reaction ( $C$ ) (see Figure 1). The following additional functions are defined:  $\rho(r) = A$  and  $\pi(r) = B$ , which return the reactants and the products of any given reaction  $r$ , respectively. With the help of these elementary functions the same notion can be extended to a set of reactions  $\mathcal{R}'$  as  $\rho(\mathcal{R}') = \bigcup_{r \in \mathcal{R}'} \rho(r)$ , where  $\mathcal{R}' \subseteq \mathcal{R}$ . This definition captures the conglomerate of molecules that participate as reactants for a set of reactions. A similar definition holds for  $\pi(\mathcal{R}')$ , the products of a subset of reactions. With these ideas, we can define the *closure* of a subset  $X' \subseteq X$  relative to  $\mathcal{R}' \subseteq \mathcal{R}$  ( $cl_{\mathcal{R}'}(X')$ ) as the set of reachable molecules that can be synthesized by starting from  $X'$  and applying all the reactions in  $\mathcal{R}'$  until no new molecule types appear. Then, a non-empty reaction subset  $\mathcal{R}'$  of  $\mathcal{R}$  is a *reflexively autocatalytic network* over  $F$  if  $\rho(\mathcal{R}') \subseteq cl_{\mathcal{R}'}(F)$  and for each  $r \in \mathcal{R}'$ ,  $\gamma(\rho(\mathcal{R}') \cup \pi(\mathcal{R}'), r) = 1$ . In other words every catalyst must be produced by a reaction in the same system or be part of the food set. This definition allows many reflexively autocatalytic networks in a *catalytic reaction system*. The network is  $F$ -generated if every reactant is either produced by the system or incorporated as a food item (i.e. formally  $\rho(\mathcal{R}) \subseteq F \cup \pi(\mathcal{R})$ ). A network that is *reflexively autocatalytic* and  $F$ -generated is called a *RAF set* (see Figure 1).

RAF sets can be understood informally as an interdependent set of biochemical reactions where all of the metabolites are produced by the collection of reactions  $\mathcal{R}'$ . The advantage of this formalism is that it is precise enough to be coded in well defined algorithms that check whether a given reaction subset  $\mathcal{R}' \subseteq \mathcal{R}$  is a RAF set over some food set  $F$ . We have implemented these algorithms, and we have created a simple framework in Lisp and Python, allowing us to carry out qualitative and quantitative analyses of  $(M,R)$  systems in terms of RAF formalism. Before discussing this, however, we need to show the extent to which RAF sets and  $(M,R)$  systems are equivalent.

### RAF sets and $(M,R)$ systems

Are  $(M,R)$  systems RAF sets? The original definition of an  $(M,R)$  system (Rosen, 1958) explicitly requires every catalyst ( $M$  in his original symbols) must be produced by the metabolism ( $R$  sub-systems are responsible for this task). This condition shows that  $(M,R)$  systems must be *reflexively autocatalytic* (RA) sets. Although, this does not necessarily imply that a RA set is an  $(M,R)$  system, because metabolic closure requires that no catalyst is given in the food set. In other words, a RA set is not in general an  $(M,R)$  system, but it may become one if all the catalysts in  $C$  are produced by the system and are not part of the food set  $F$ .

As  $(M,R)$  systems must be open to the flow of matter in order to satisfy thermodynamic requirements, their molecules derive ultimately from a food source, and they are, obviously,  $F$ -generated in the terminology of RAF sets. So  $(M,R)$  systems without organizational invariance are a sub-

set of RAF sets, as are  $(M,R)$  systems with organizational invariance. The latter must, however, have additional features (in the context of RAF) to explain the unusual properties of operators  $\Phi$  and  $\beta$ .

### Algorithmic search for simple metabolic $(M,R)$ systems

In this section we explore the probability of occurrence of an  $(M,R)$  system with a unique assignment of catalysts. For this purpose we characterized all the possible graphs describing a system consisting of a number  $\#F$  of initial molecules and  $\#\mathcal{R}$  synthesis reactions between any two molecules in the system. More specifically, we analyzed systems that conformed with the requirement of being  $(M,R)$  systems, that is, we did not allow any catalyst to be food, nor a reactant nor a product in the reaction it catalyzed.

Attention must be paid to avoid having two apparently distinct reaction networks exhibiting the same topological structure. The mathematical term for this is *graph isomorphism* (see Figure 3). Two graphs are said to be isomorphic when they can be transformed into each other by a simple relabeling of their vertices. Isomorphic metabolisms can be grouped under an equivalence class.

Thus, for a given pair  $(\#F, \#\mathcal{R})$  we enumerate the number of all possible different equivalence classes of reaction networks. Next, for each one of these reaction networks, we generated the set of all possible assignments for the catalysts complying with the restrictions stated previously. But again, by the argument of relabeling, the set of assignments can be

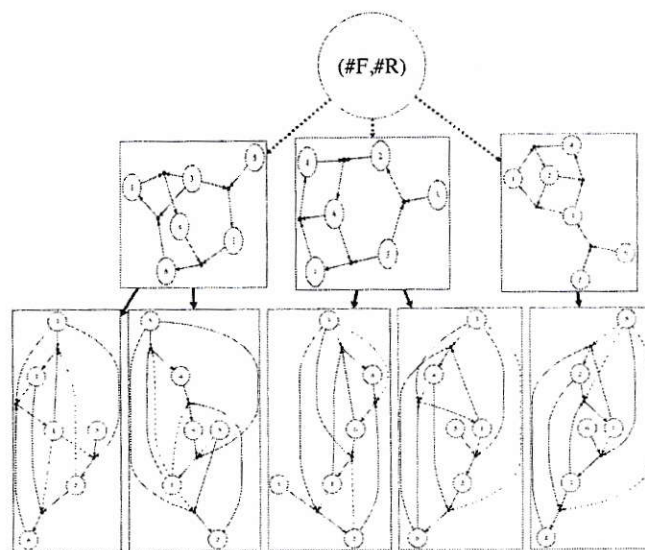


Figure 2: Diagram representing an example for the procedure to compute results from table 1. In the first step, the equivalence classes (3 in this example) are estimated for a given  $(\#F, \#\mathcal{R})$ ; in the second step, all possible catalyst assignments for each equivalence class are calculated.



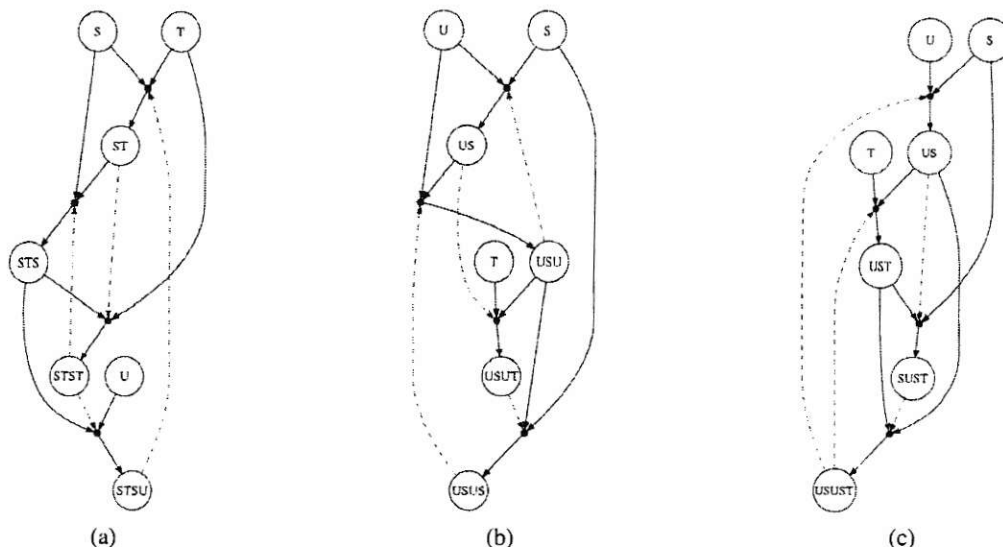


Figure 3: Three automatically generated RAF sets illustrating equivalence class and multiple catalyst assignments. Systems (a) and (b) have the same topological structure, i.e. there is an isomorphism from one to the other. Although this might not be obvious at first sight, a simple procedure of node relabeling transforms the reaction pathway in (a) to the one in (b). In spite of that, the systems differ in their catalyst assignments, i.e., even with the additional rules imposed by  $(M,R)$  systems, it is possible to make different choices when assigning the catalysts. System (c) has the same number of elements in the food set and the same number of reactions, but it belongs to another equivalence class.

also divided into equivalence classes (see Figure 2). Table 1 shows for  $(\#F, \#R)$  the number of metabolic equivalence classes and the interquartile range<sup>1</sup> of the number of assignments. It can be seen that the number of possible assignments grows steeply with the number of reactions, so that it becomes more and more difficult to have a unique  $\Phi(b) = f$  (Letelier et al., 2006).

There are some cases in which the range includes the critical value 1, which implies organizational invariance. Although, if we increase the number of food elements and leave the number of reactions unchanged, the generated reaction networks become shallower, and so we can consider the complexity of the network to be reduced and therefore the degrees of freedom of the assignment process are also reduced. In principle we could separate the trivial cases from those in which the unicity of the assignment reflects organizational invariance.

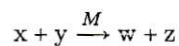
### Rosen's triad in RAF formalism

The RAF formalism is not only useful for exploring the landscape of possible  $(M,R)$  systems, but it can also help to clarify some core concepts of  $(M,R)$  systems, namely Rosen's triad:  $f$ ,  $\Phi$  and  $\beta$ .

To explore the potential of the RAF formalism, we analyze the old problem in the theory of  $(M,R)$  systems of how

<sup>1</sup>This refers to the range in which data falls after removing lower and upper 25%, thus giving a notion of the amplitude of the mean values

to treat molecules as functions. Consider the following biochemical reaction:



According to Rosen, this is the manifestation of the following function:

$$\begin{aligned} M &\in \text{Map}(X \times Y, W \times Z) \\ M &: X \times Y \rightarrow W \times Z \\ (x, y) &\rightarrow (w, z) \end{aligned}$$

The input elements are derived from the cartesian set  $X \times Y$  that contains all the molecular types that, because of their structural similarities, can be used by the enzyme  $M$  as substrates. Our RAF-derived formalism extends the domain of function  $M$  to the whole set of molecules as follows:  $M$  is a function that, when given a set of molecules with the reactants, e.g.  $(\dots, x, \dots, y, \dots)$ , returns a set containing elements  $w$  and  $z$ . But if the original input set lacks elements  $x$  or  $y$ , we have  $M(\text{input set}) = \emptyset$ . Interestingly, with this formalism any molecule in the network ( $x \in X$ ) can be treated as a function operating on any subset ( $X' \subseteq X$ ) as follows:

$$x(X') = \pi(r_x) \text{ provided that } \rho(r_x) \subseteq X'$$

where  $r_x$  stands for the reaction that  $x$  catalyzes. If  $x$  catalyzes more than one reaction<sup>2</sup>, then the above definition can

<sup>2</sup>This multifunctionality seems to be necessary for  $(M,R)$  systems (Letelier et al., 2006).

Number of food molecules	Number of reactions					
	3		4		5	
2	4	2-2	19	12-24	136	144-216
3	10	1-4	72	12-31	685	216-324
4	8	1-6	75	1-36	933	204-432
5	2	1-1	37	1-34	577	1-432
6	1	1-1	11	1-1	212	1-1

Table 1: Number of metabolic equivalence classes and the interquartile range of the number of their possible assignments. The number of equivalence classes increases dramatically with the number of reactions.

be generalized to:

$$x(X') = \{x_i : x_i \in \pi(r) \mid (x, r) \in C \wedge \rho(r) \subseteq X'\} \quad (2)$$

Note that defining  $x$  only requires the set of reactions each molecule catalyzes, not the whole reaction network. This means that every molecule-as-a-function definition depends only on local information.

In our earlier work, the following small metabolism was used as a testbed for exploring concepts related to  $(M, R)$  systems.



Then, treating every molecule as a function we have:

$$\begin{aligned} SU(S, T) &= \{ST\} \\ STU(S, U, T) &= \{SU\} \\ U(S, T, U, ST, STU) &= \emptyset \\ &\dots \end{aligned}$$

The last equation means that molecule U cannot transform the given mixture, because U is not a catalyst in the given metabolism. That said, we shall now analyze how concepts like  $f$ ,  $\Phi$  and  $\beta$  can be expressed with these ideas.

### Metabolism: $f$

One of the basic equations in Rosen's model is  $f(a) = b$ , in which  $a$  represents the input materials (food set) needed by the organism to produce the complete set of metabolites and enzymes ( $b$ ), i.e., every molecule reachable by the metabolism. Therefore, the function  $f$  is related to the notion of closure ( $cl_{\mathcal{R}}(X')$ ). To be able to define  $f$  in our terms, let us define function  $expand$ .

$$expand_X(X') = X' \cup \bigcup_{x_i \in X} x_i(X') \quad (6)$$

Moreover, let us define how a molecule set ( $X'$ ) can be applied to another molecule set ( $Y'$ ).

$$\overrightarrow{X'}(Y') = \begin{cases} Y' & \text{if } expand_{X'}(Y') = Y', \\ \overrightarrow{X'}(expand_{X'}(Y')) & \text{otherwise} \end{cases} \quad (7)$$

Thus, we use a molecular set as a function (distinguished from regular molecular set by a "semi-arrow") by repeatedly applying  $expand$  until no further additions occur. With these two last definitions, for any given catalytic reaction system  $L = (X, \mathcal{R}, C)$ ,  $f(a)$  can be defined as:

$$f(a) = \overrightarrow{catalysts(C)}(a) = b \quad (8)$$

where  $catalysts$  is a function that returns every catalyst in the given catalyzation set  $C$  ( $catalysts(C) = \{x : (x, r) \in C\}$ ). The function  $catalysts$  is not required, as non-catalyst molecules do not modify the result. But it is used here as Rosen's formalism considers only catalysts as the core components of the metabolism.

### Replacement: $\Phi$

The formulation of  $\Phi$  under RAF sets is more elaborate as we need to generate a function that using  $b$  as an input returns function  $f$ . The basic idea is to create mathematical objects that somehow keep track of which catalysts are produced and how these are created as a result of the metabolism. To begin we introduce operator  $Op$ . This operator returns the subset of molecules  $X'' \subseteq X'$  that can act as catalysts upon the molecules in  $X'$  (the given molecule set).

$$Op(X') = \{x \in X' : x(X') \neq \emptyset\}$$

Then, for any given catalytic reaction system  $L = (X, \mathcal{R}, C)$  over a food source  $F$ ,  $\Phi(b)$  will be defined as

$$\Phi(b) = \overrightarrow{Op}(cl_{\mathcal{R}}(b) \cup F) = f' \quad (9)$$

where  $cl_{\mathcal{R}}(b)$  is the closure of  $b$  relative to the reaction set  $\mathcal{R}$  as defined above. Therefore,  $\Phi$  returns the catalyst set that are reachable from  $b$  as a function ( $f'$ ), because the "semi-arrow" over the expression transforms the resulting set into a function. Thus,  $f'$  is operationally equivalent to function  $f$ .

## Organizational invariance: $\beta$

Finally, it remains to define  $\beta$ , which should take the metabolism  $f$  as input and give us the replacement system  $\Phi$ . The function  $\beta$  receives a hypothetical metabolism  $f'$  in the form of a function, thus our first step will be to find which catalysts can be related to that function  $f'$ . For that purpose, let us define the function  $\nu$  that given a molecular set  $b$  and a function  $f'$ , returns every reaction catalyzed by molecules in  $b$ , which produces part of the result of  $f'$  applied to  $F$ .

$$\nu(b, f', F) = \{r : \gamma(b \cup F, r) = 1\}$$

By using a new function  $\mu$ , we filter out those reactions that cannot take place given the molecule set of interest ( $b \cup F$ ).

$$\mu(b, f', F) = \{r \in \nu(b, f', F) : \rho(r) \subseteq b \cup F\} \quad (10)$$

This equation gives the reactions that are related to  $f'$ , therefore  $\beta$  can be defined. For simplicity we shall define it as applied to a molecular set  $b$ .

$$\beta(f')(b) = \overline{Op(Cl_{\mu(b, f', F)}(b) \cup F)} \quad (11)$$

This formula is similar to that of  $\Phi$ , the main difference being that it uses function  $\mu$  to obtain  $\mathcal{R}$  instead of using  $\mathcal{R}$  directly. In this way  $\beta$  returns a function that, used in an  $(M, R)$  system, would relate unequivocally to  $\Phi$ .

## Conclusion

A formidable challenge for using  $(M, R)$  systems as a framework for modeling biological systems has been the lack of operational definitions for the important functions  $f$ ,  $\Phi$  and  $\beta$ . Here we have presented various definitions for those functions that can be used for any catalytic reaction system.

An important unresolved matter is to make explicit how Rosen's equations can be fulfilled using concepts and definitions imported from RAF sets. Suppose that a given molecule set  $X$  and reaction set  $R$  compose an  $(M, R)$  system, how can that be proved using RAF-derived functions? First, let us distinguish a particular subset  $a$  of  $X$ , which contains every molecule that is not a product or a catalyst for any reaction. Then, we can write:

$$f(a) = b$$

This signifies "let the molecular system evolve until no further novelty can be produced". Now, we should expect that using the produced molecules as function will have the same effect as using  $f$ . In our terms, that means:

$$\Phi(b)(a) = b$$

This has the important consequence that  $f$  becomes equivalent (operationally) to  $\Phi(b)$  in this molecular system.

$\beta$ , as introduced here, does not explain Rosen's basic result ( $\beta(f) = \Phi$ , which means that  $\Phi$  is uniquely determined by  $f$ ). The definition of  $\beta$  and all associated formulae cannot explain Rosen's result, they merely serve as formal language that could help us to operate on modern metabolic data using Rosen's viewpoint.

Since the beginning of the 21st century there has been a resurgence of interest in the work of Robert Rosen, but it is not easy to understand and it is not apparent how to advance in a theory full of powerful but often obscure ideas (Letelier et al., 2006). Many attempts have been made to find the route to be followed in developing the theory (Wolkenhauer and Hofmeyr, 2007). Here we apply another formalism (RAF sets) that could be useful for clarifying the nature and properties of the operators  $f$ ,  $\Phi$  and  $\beta$ .

Finally, we have the caveat that living systems are not mere "soups of letters", and their complex properties are due to more than some combinatorics among molecules. It is apparent that to advance in our understanding of living organisms, it will be necessary to include further considerations into our current theory. These could be geometrical, thermodynamical, topological, or even merely historical, that is, relative to how life has come into existence, and later evolved here on Earth.

The RAF formalism may usher in an era in which the theory of  $(M, R)$  systems will demand reasoning tools that begin to resemble category theory more and more... Rosen would be amused!

## References

- Cárdenas, M. L., Letelier, J. C., Gutierrez, C., Cornish-Bowden, A., and Soto-Andrade, J. (2010). Closure to efficient causation, computability and artificial life. *J. Theor. Biol.*, 263:79–92.
- Cornish-Bowden, A., Cárdenas, M. L., Letelier, J. C., and Soto-Andrade, J. (2007). Beyond reductionism: metabolic circularity as a guiding vision for a real biology of systems. *Proteomics*, 7:839–845.
- Eigen, M. and Schuster, P. (1977). A principle of natural self-organization. *Naturwissenschaften*, 64(11):541–565.
- Gánti, T. (1975). Organization of chemical reactions into dividing and metabolizing units: the chemotons. *Bio Systems*, 7(1):15–21.
- Hordijk, W. and Steel, M. (2004). Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *J. Theor. Biol.*, 227:451–461.
- Kauffman, S. (1993). *The Origins of Order*. Oxford University Press, New York.
- Kauffman, S. A. (1986). Autocatalytic sets of proteins. *J. Theor. Biol.*, 119(1):1–24.
- Letelier, J. C., Marín, G., and Mpodozis, J. (2003). Autopoietic and  $(M, R)$  systems. *J. Theor. Biol.*, 222:261–272.

- Letelier, J. C., Soto-Andrade, J., Guñez Abarzúa, F., Cornish-Bowden, A., and Cárdenas, M. L. (2006). Organizational invariance and metabolic closure: analysis in terms of  $(M,R)$  systems. *J. Theor. Biol.*, 238:949–961.
- Maturana, H. and Varela, F. (1980). *Autopoiesis and Cognition: the Realisation of the Living*. D. Reidel Publishing Company, Dordrecht, The Netherlands.
- Mossel, E. and Steel, M. (2005). Random biochemical networks: the probability of self-sustaining autocatalysis. *J. Theor. Biol.*, 233(3):327–336.
- Rosen, R. (1958). A relational theory of biological systems. *Bull. Math. Biophys.*, 20:245–260.
- Wolkenhauer, O. and Hofmeyr, J.-H. (2007). An abstract cell model that describes the self-organization of cell function in living systems. *J. Theor. Biol.*, 246:461–476.

# A decomposition theorem in chemical organizations

Tomas Veloz<sup>1</sup>, Bryan Reynaert<sup>2</sup>, David Rojas-Camaggi<sup>2</sup> and Peter Dittrich<sup>3</sup>

<sup>1</sup>Department of Psychology, University of British Columbia

<sup>2</sup>Department of Biology, Universidad de Chile

<sup>3</sup>Friedrich-Schiller-University Jena, Institute of Computer Science, Bio Systems Analysis Group, D-07743 Jena  
tomas.veloz@ubc.ca

## Abstract

The Chemical Organization Theory (COT) is an abstract reaction network model that has a deep connection to autopoiesis as they share the same central topic: Organization. The main characteristic of autopoietic systems is that they preserve their own organization; this constitutes their identity. In terms of COT, organizations are special reaction networks which are closed and self-maintaining. Organizations compose the majority of stable behaviours of a reaction network (Peter and Dittrich, 2011), in particular every fixed point can be mapped to an organization (Dittrich and Di Fenizio, 2007). Obtaining the set of organizations of a network is a central objective in CCT, but it is usually a complex computational task. This work intends to reveal the underlying mathematical structure of organizations. We state a theorem of decomposition for organizations to understand the difficulties of verifying if a set of molecular species is an organization. This suggests a step towards the development of more efficient algorithms and the classification of reaction networks in terms of how complex it is to obtain its set of organizations. We also discuss the consequences of this theorem in relation to autopoietic systems.

## Introduction

During a 30-years period, from the 1950's to the 1980's, the field of biological systems and their generalized properties saw the birth of multiple theories (Eigen and Schuster, 1977; Kauffman, 1969; Maturana and Varela, 1973; Wiener, 1948; von Bertalanffy, 1968; Rosen, 1958). A wealth of formalisms were laid out, which focused on different perspectives on the fundamental properties of living systems, but as it was to be expected, there have been deep similarities between most of these theories (Hordijk and Steel, 2004; Jaramillo et al., 2010; Letelier et al., 2003).

Since their conception, most of these theories have been consigned to the theoretical domain having little incidence in applied sciences, with the possible exception of what is currently known as *systems biology*.

This situation may be because the process of translation between the language employed in these theories and the language commonly used in biology is not trivial (Cornish-Bowden et al., 2007). The chemical organization theory, inspired by Fontana and Buss (1994), provides an interesting departure from this tendency as it provides a language

which is not only clear and well-defined but also correlates directly to the biomolecular domain. Due to its mathematical foundations theorems can be formally proven and developed (Benkő et al., 2009; Peter and Dittrich, 2011; Peter et al., 2010). Also, COT is a powerful tool to analyze the asymptotic behaviour of reaction networks that other analytic or simulation methods cannot cope with. In particular, the chemical organization theory has been applied to biochemical domains (Centler et al., 2008b; Kaleta et al., 2006; Matsumaru et al., 2006), atmospheric photochemistries (Centler and Dittrich, 2007), and as a tool for the study of P-systems (Peter et al., 2010). It also has been proposed as a theoretical framework to design chemical computers (Matsumaru et al., 2007), and recently, for the study of social systems (Dittrich and Winter, 2008).

Thus, COT is very well suited to study autopoietic systems as both theories focus on the problem of self-maintaining organizations. At first it may seem inappropriate that a theory developed around artificial chemistries may be used to study autopoietic systems, but it should be noted that autopoietic systems are not obliged to a molecular structure or realization, that just happens to be the case of living organisms. Furthermore, the "protobio" (Varela et al., 1974) was both an early attempt to simulate autopoietic systems and an artificial chemistry. Therefore, any advance in COT might be transported directly to the theory of autopoietic systems, independent of the domain in which they are actually realized.

In this paper, we first introduce COT and its relation to autopoietic systems. Then, we present a decomposition theorem from COT and finally analyze its consequences for the long-term time behavior of biological systems.

## Autopoiesis and Chemical Organization Theory

Autopoiesis was developed as a theory for living systems by Maturana and Varela (1973). The central idea is that a living organism is a machine, constituted as a unit in space, which maintains its organization through its operation. Moreover, a living organism performs a set of pro-

cesses which generate the components necessary to realize these processes. Thus, the notion of organization as a network of interacting components which stably maintains itself in time is of most importance in this theory. Hence, a theory which concerns itself with such a concept may relate closely to autopoiesis.

The COT, which was introduced in Dittrich and Di Fenizio (2007) in the context of algebraic chemistries, is a mathematical theory, that by using the structures of sets and matrices, is able to formalize chemical reaction systems at a topological and dynamical level. In this theory, an organization is a reaction network which has the potential of being self-maintaining and thus matches very closely to the definition given by Maturana and Varela. Moreover, as “an autopoietic system is an homeostatic machine which has its organization as the variable it maintains constant”, organizations must be stable in time. The COT explores these considerations and has already had important results in this regard. In particular, in this work we present a decomposition theorem for organizations. In order to present our main result, we must first introduce the basics of COT.

## Chemical Organization Theory

### Basic Definitions

At the most basic level of this theory, we deal with two types of objects: molecular species (from now on species) and reactions. The species are the elements of a species set  $\mathcal{M} = \{m_1, \dots, m_n\}$ , and each reaction  $R$  is modeled by a pair  $R = (A, B) \in \mathcal{P}_M(\mathcal{M}) \times \mathcal{P}_M(\mathcal{M})$ , where  $\mathcal{P}_M(\mathcal{M})$  denotes the set of all the multisets formed by elements in  $\mathcal{M}$ . A multiset is defined by a pair  $(X, \eta_X)$ , where  $X$  is a set and the function  $\eta_X : X \rightarrow \mathbb{N}_0$  states the number of occurrences  $\eta_X(x)$  (multiplicity) of  $x$  in the multiset. In order to be consistent with the usual notation of chemical reactions, we will write the multiset  $(X, \eta_X)$  by  $\sum_{x \in X} \eta_X(x)x$ . Moreover, we will refer to the reaction  $R = (A, B)$  by  $R = A \rightarrow B$ , where  $A = (\mathcal{M}, \eta_A)$  and  $B = (\mathcal{M}, \eta_B)$ .

From now on, let  $\mathcal{R} = \{R_1, \dots, R_k\}$ , where  $R_i = A_i \rightarrow B_i$ , with  $A_i = a^{i1}m_1 + \dots + a^{in}m_n$  and  $B_i = b^{i1}m_1 + \dots + b^{in}m_n$ , for  $i = 1, \dots, k$  and  $j = 1, \dots, n$ .  $a^{ij}$  corresponds to the stoichiometric coefficient of  $m_j$  in reaction  $R_i$ , that is, the multiplicity  $\eta_{A_i}(m_j)$  of molecule  $m_j$  in  $A_i$ ;  $b^{ij}$  is defined in a similar way. Now we can define an Algebraic Chemistry, which captures the notion of system, as follows:

**Definition 1** An Algebraic Chemistry is a pair  $\langle \mathcal{M}, \mathcal{R} \rangle$ .

A species  $m \in \mathcal{M}$  is said to be *present* in a multiset  $(X, \eta_X) \in \mathcal{P}_M(\mathcal{M})$  if and only if its multiplicity  $\eta_X(m)$ , is greater than zero. The *reactants* and *products* of a reaction  $R = A \rightarrow B$  are the species present in  $A$  and in  $B$  respectively. The reaction  $R$  can be fired by a set  $X \subseteq \mathcal{M}$  if and only if all species present in  $A$  are in  $X$ .

From now on let  $X \subseteq \mathcal{M}$ . Note that there exists a maximal set of reactions  $\mathcal{R}_X \subseteq \mathcal{R}$  which can be fired by  $X$ .  $\mathcal{R}_X$

is composed by the reactions  $R_i = A_i \rightarrow B_i$  such that, if  $m$  is present in  $A_i$ , then  $m \in X$ . We call  $\mathcal{R}_X$  the *possible reactions* set of  $X$ .

In order to deal with the dynamical aspects of any system, it is desirable that the system maintains its identity. This leads to the question of whether the system, left to react for an arbitrary amount of time, will generate species which where originally absent. Note that in a general chemical setting, in which no species will be used up completely, all the reactions that can be fired will fire at some positive rate; therefore, it suffices to check if the set of possible reactions for the system produces any novel species. If it does not, we say that the set of species is closed. The following definition states this formally:

**Definition 2** We say  $X$  is closed if and only if for all  $R = A \rightarrow B \in \mathcal{R}_X$ ,  $m$  is present in  $B$  implies  $m \in X$ . Let  $G_{CL}(X)$  be the closure of  $X$ , then it is the smallest closed set containing  $X$ .

**Remark** The closure of a set has been proven to be unique (Dittrich and Di Fenizio, 2007).

Thus, any given set of species will react growing in qualitative novelty until it reaches its closure, but it is unclear whether the set will be stable in time, considering that during reactions species are consumed and their concentration could drop to zero. This motivates the study of dynamical properties of sets of species.

### Dynamical Aspects

The *stoichiometric matrix*  $\mathbf{S} = (s_{ij})$  associated with  $\langle \mathcal{M}, \mathcal{R} \rangle$  is a  $n \times k$  matrix, where  $s_{ij}$  is the stoichiometric coefficient of species  $m_i$  in the reaction  $R_j$  ( $s_{ij}$  is negative if species  $m_i$  is consumed by reaction  $R_j$ ). Indeed,  $s_{ij} = b^{ji} - a^{ji}$ . The stoichiometric matrix is at the core of current systems biology (Schuster et al., 1999; Schilling and Palsson, 1998) and its properties have been extensively studied (Kacser and Burns, 1973; Heinrich and Rapoport, 1974). Let the *flux vector*  $\mathbf{v} = (v_1, \dots, v_k)$  be a non negative vector such that the application of  $\mathbf{v}$  on the stoichiometric matrix  $\mathbf{S}$  represents a reaction process, i.e. for  $i = 1, \dots, k$ , the rate of the reaction  $R_i$  in the system is given by  $v_i$ . We define the *production rate vector* by  $\mathbf{f} = \mathbf{S}\mathbf{v}$ . Thus, for  $i = 1, \dots, n$ , we have that  $f_i$  is the rate of production of the species  $m_i$  in the reaction process determined by  $\mathbf{v}$ .

We can describe the dynamics of the species concentrations  $\mathbf{x} = (x_1, \dots, x_n)$  by the system of ODEs

$$\dot{\mathbf{x}} = \mathbf{S}\mathbf{v}(\mathbf{x}, \mathbf{k}), \quad (1)$$

where according to mass-action kinetics

$$v_i = k_i \prod_{j=1}^n x_j^{a_{ij}}$$

for  $i = 1, \dots, k$ , is the *flux*, and  $\mathbf{k} = (k_1, \dots, k_k)$  is a strictly positive vector denoting reaction rate constants. We call ODE (1) a *chemical reaction system*.

In order to relate the statical domain with the dynamical domain, we introduce the idea of abstractions and instances:

**Definition 3** The *abstraction* of state  $\mathbf{x}$  is the set  $\phi(\mathbf{x})$  with

$$\begin{aligned} \phi: \mathbb{R}_{\geq 0}^n &\mapsto \mathcal{P}(\mathcal{M}) \\ \mathbf{x} &\mapsto \phi(\mathbf{x}) \equiv \{m_i \in \mathcal{M} : x_i > \epsilon\}, \end{aligned} \quad (2)$$

where  $\mathbb{R}_{\geq 0}^n$  denotes the set of non-negative real numbers, and  $\epsilon$  is a concentration threshold. Moreover, given a set of species  $X \subseteq \mathcal{M}$ , a state  $\mathbf{x}$  is an *instance* of  $X$  if and only if its abstraction equals  $X$ .

## Chemical Organizations

The following definition is at the core of chemical organization theory:

**Definition 4** A subset of species  $X \subseteq \mathcal{M}$  is an *organization* if and only if

1.  $X$  is *closed* and
2.  $X$  is *self-maintaining*, i.e. there is a strictly positive flux vector  $\mathbf{v}$  so that

$$\mathbf{S}_X \mathbf{v} \geq \mathbf{0}$$

where  $\mathbf{S}_X$  is the stoichiometric matrix associated to the Algebraic Chemistry  $\langle X, \mathcal{R}_X \rangle$ .

Organizations are sets of species which cannot produce new species by their possible reaction set. Also, it is possible that during the operation of an organization, the concentration of none of the species decreases; thus, an organization either maintains itself in time or grows in terms of the concentration of its species. This definition shares fundamental properties with that of autopoietic systems to the extent that all autopoietic systems are organizations. Note that not all organizations are autopoietic systems as an organization that keeps growing is not homeostatic and will eventually rupture its container. This motivates the study of the fixed points and other attractors of the chemical reaction systems.

The following theorem relates fixed points and organizations<sup>1</sup>.

**Theorem 1** If  $\mathbf{x}$  is a fixed-point of the ODE (1), i.e.  $\mathbf{Sv}(\mathbf{x}, \mathbf{k}) = \mathbf{0}$ , then the abstraction  $\phi(\mathbf{x})$  is an organization.

Fixed points are related to the dynamic stability of chemical systems. Moreover, since fixed points determine most of the characteristics of the dynamic systems they belong to (Strogatz, 2000), Theorem 1 provides a link between the

long-term behavior of a chemical reaction system and its underlying reaction network. This allows the study of the system's dynamics by the chemical organization theory. Furthermore in (Peter and Dittrich, 2011), Theorem 1 is extended to other stable asymptotic behaviours such as periodic orbits and limit cycles. In addition, the necessary conditions for the existence of adequate flux vectors are explored in (Peter et al., 2010). Note that a fixed point in this context does not refer to thermodynamic equilibrium but to the maintenance of the size of the system in terms of the number of its components. The question about stability refers to the conservation of the structure or organization of the processes in a given timescale as the system is also subject to an evolutionary dynamic which can lead to change or desintegration. Now that we have introduced the idea of organization and shown some relevant aspects, we will focus on our main result; the decomposition theorem.

## Species Role in a Network

The idea behind the role of a species is that it can be classified in relation to a set of species by how it behaves in the set of possible reactions.

## Reactivity and Catalysts

**Definition 5** Let  $m \in X$ , then

- $m$  is *non-reactive* w.r.t  $X$  if and only if for all reactions  $R = A \rightarrow B \in \mathcal{R}_X$ ,  $m$  is not present in  $A$  nor in  $B$ .
- $m$  is a *catalyst* w.r.t  $X$  if and only if for some reaction  $R' = A' \rightarrow B' \in \mathcal{R}_X$ ,  $m$  is present in  $A$  and for all reactions  $R = A \rightarrow B \in \mathcal{R}_X$ ,  $\mathcal{A}(A, m) = \mathcal{A}(B, m)$ .
- $m$  is *reactive* w.r.t  $X$  if and only if for some reaction  $R' = A' \rightarrow B' \in \mathcal{R}_X$ ,  $m$  is present either in  $A'$  or in  $B'$  and for some reaction  $R = A \rightarrow B \in \mathcal{R}_X$ ,  $\mathcal{A}(A, m) \neq \mathcal{A}(B, m)$ .

We say that  $Y \subseteq X$  is a *non-reactive*, *catalytic* or *reactive* set of  $X$ , if for all  $m \in Y$ ,  $m$  is non-reactive, a catalyst or reactive w.r.t  $X$  respectively.

The following lemma is straightforward

**Lemma 1** There is a unique maximal non-reactive, catalytic and reactive set of  $X$ .

**Definition 6** The maximal non-reactive, catalytic and reactive sets of  $X$  are called the *non-reactive*, *catalytic* and *reactive* sets of  $X$  respectively.

## Overproduction

**Definition 7** Consider the Algebraic Chemistry  $\langle \mathcal{M}, \mathcal{R} \rangle$  and a non-negative flux vector  $\mathbf{v}$  such that  $(\mathbf{Sv})_i = \mathbf{f}_i \geq 0$  for  $i = 1, \dots, n$ . If  $\mathbf{f}_j > 0$  for some  $j = 1, \dots, n$ , we say that  $m_j$  is an *overproduced* species in  $\langle \mathcal{M}, \mathcal{R} \rangle$ .

<sup>1</sup>Proof can be found in (Dittrich and Di Fenizio, 2007).

Overproduced species have a positive production rate for certain flux vectors which do not lead to the consumption of any other species. We remark that on the one hand, the definition of overproduced species does not demand that the system is self-maintaining because the flux vector is only required to be non-negative, but on the other hand, overproduced species definition not only requires the non-negative production of all the species, but also the positive production of at least one species. Thus, overproduced species are the species that can be *indefinitely* produced by some *reaction pathway*. Note that although this seems to violate the law of mass conservation, real systems require a constant input of mass or energy, and thus, it is usual when simulating or analyzing chemical networks to include an outer source of mass which does not decrease when consumed by a reaction. The relevance of these species is that they can actually be overproduced without consuming any of the inner species of the system; hence, they embody the notion of input. The following lemma is straightforward.

**Lemma 2** *Let an overproduced species  $m \in X$  in  $\langle X, \mathcal{R}_X \rangle$ . If  $X \subset Y$ , then  $m$  is overproduced in  $\langle Y, \mathcal{R}_Y \rangle$ .*

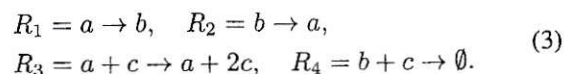
**Corollary 1** *If  $X$  is a set of overproduced species in  $\langle \mathcal{M}, \mathcal{R} \rangle$ , then its closure  $G_{CL}(X)$  is also overproduced.*

**Lemma 3** *There exists a unique maximal set  $F$  of overproduced species in  $\langle X, \mathcal{R}_X \rangle$ .*

**Proof** If there are no overproduced species in  $X$  then the maximal overproduced set is the empty set. Otherwise the set containing all the overproduced species in  $\langle X, \mathcal{R}_X \rangle$  leads to a maximal overproduced set. Now we are going to prove that the maximal overproduced set is unique. Suppose that there are two maximal overproduced sets  $F_1, F_2 \in X$  and  $F_1 \neq F_2$ , let  $\mathbf{v}_1, \mathbf{v}_2$  the flux vectors required to verify the overproduced property of  $F_1$  and  $F_2$  w.r.t  $\langle X, \mathcal{R}_X \rangle$  respectively. Trivially,  $\mathbf{v}_1 + \mathbf{v}_2$  verifies the overproduced property of  $F_1 \cup F_2$  w.r.t  $\langle X, \mathcal{R}_X \rangle$ , and  $F_i \subset F_1 \cup F_2$  for  $i = 1, 2$ . As the inclusion is strict we have a contradiction.

**Definition 8** *The maximal set of overproduced species  $F$  with respect to  $X$  is called the overproduced set of  $X$ .*

**Remark** Consider the situation of adding a species  $m$  to an organization  $O$ . The fact that  $m$  is overproduced in  $O' = O \cup \{m\}$  does not guarantee that  $O'$  is an organization. For example, consider the set of species  $O' = \{a, b, c\}$  and the set of reactions



We have that  $O = \{a, b\}$  is an organization,  $c$  is overproduced in  $O'$ , but  $O'$  is not an organization.

Then, at first sight the overproduced species could be seen as a useless definition concerning the self-maintainance of a reaction network because the overproduced species ( $c$  in the

example) can catalyze the consumption of species that cannot be recovered in the network ( $b$  through reaction  $R_4$  in the example). However, the identification of the overproduced molecules of a set  $X$  simplifies the verification of the self-maintaining condition of any set that contains  $X$ . Indeed in the example above, we have that  $c$  is overproduced in  $\{a, c\}$ , thus we can avoid the calculation of the production of the species  $c$  when verifying the self-maintainance of  $O'$ .

## Roles and Organizations

From now on let  $N, E, F$  the non-reactive, catalyst and overproduced set of  $X$  respectively.

**Definition 9**  *$X - (F \cup E \cup N)$  is the potential active cycle (PAC) of  $X$  w.r.t  $F$ .*

**Remark** For any given flux vector which verifies the self-maintainance of  $X$ , the PAC has a production rate equal to zero. But PAC should not be confused with the set of species with a production rate equal to zero. Indeed, the non-reactive and catalytic species have a production rate zero, but they do not belong to the PAC. The following lemma states that no species can be only produced or only consumed in the PAC of an organization:

**Lemma 4** *Let  $C$  be the PAC of  $X$ . If  $X$  is an organization, then for every  $m \in C$  we have that  $m$  is consumed by some reaction  $R \in \mathcal{R}_X$  and produced by other reaction  $R' \in \mathcal{R}_X$ .*

**Proof** Let  $m \in C$ , then  $m$  cannot be non-reactive either catalyst. As  $m$  has production zero, then  $m$  is a reactive species w.r.t  $X$ . Then, there must exist a reaction  $R = A \rightarrow B \in \mathcal{R}_X$  s.t  $A(A, m) \neq A(B, m)$ . If  $A(A, m) > A(B, m)$ , as  $X$  is an organization, there has to exist some reaction  $R' = A' \rightarrow B'$  s.t  $A(A', m) < A(B', m)$ . On the other hand, if  $A(A, m) < A(B, m)$ , as  $m$  is not overproduced (because  $m \in C$ ), there has to exist some reaction  $R' = A' \rightarrow B'$  s.t  $A(A', m) > A(B', m)$ .

## PAC and Dependent Connectivity

We are going to define a special notion of connectivity which will allow us to separate the PAC of a set  $X$  in a number of partially non-overlapping sub-PACs, such that the self-maintainance of  $X$  can be studied from the self-maintainance of the sub-PACs. From now on, let  $X$  be a closed set<sup>2</sup>.

The following definition of connectivity appears in (Centler et al., 2008a):

**Definition 10** *Two species  $m_i$  and  $m_j$  in  $X$  are directly connected in  $\langle X, \mathcal{R}_X \rangle$  if and only if there exist a reaction  $R = A \rightarrow B \in \mathcal{R}_X$  such that  $\{m_i, m_j\} \subseteq A \cup B$ .*

<sup>2</sup>Verifying the closed property, and obtain the closure of a set of species is trivial compared with verifying its self-maintainance (Centler et al., 2008a)



**Definition 11** Two species  $m_i$  and  $m_j$  in  $X$  are connected in  $\langle X, \mathcal{R}_X \rangle$  if and only if there exist a sequence of species  $m_0, \dots, m_p \in X$  such that  $m_i = m_0$ ,  $m_k$  and  $m_{k+1}$  are directly connected in  $\langle X, \mathcal{R}_X \rangle$  for all  $k = 0, \dots, p-1$  and  $m_p = m_j$ .

We present a more restricted notion of connectivity than definition 11. This restriction only connects species that are non-independent when verifying self-maintainance.

**Definition 12** Two species  $m_i$  and  $m_j$  in  $X$  are dependently connected in  $\langle X, \mathcal{R}_X \rangle$  if and only if there exists a sequence of species  $m_0, \dots, m_p \in X - (E \cup F)$  such that  $m_i = m_0$ ,  $m_k$  and  $m_{k+1}$  are directly connected in  $\langle X, \mathcal{R}_X \rangle$  for all  $k = 0, \dots, p-1$  and  $m_p = m_j$ .

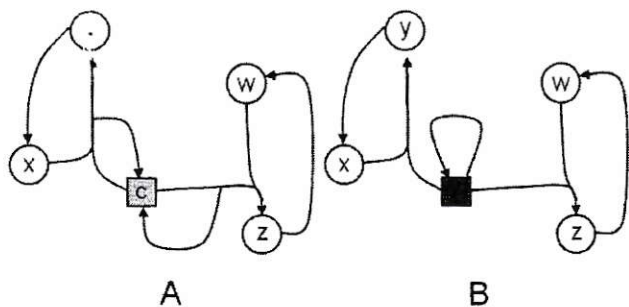


Figure 1: Following definition 11, both networks A and B are fully connected. Note that in A, the self-maintainance of  $C_1 = \{x, y, c\}$  and  $C_2 = \{z, w, c\}$  are independent. The same situation occurs in B with sets  $C'_1 = \{x, y, o\}$  and  $C'_2 = \{z, w, o\}$ . *Dependent connection* allows to connect all the species in  $C_1$  without connecting them to species in  $C_2$  and viceversa because  $C_1$  and  $C_2$  are connected through a catalyst. Analogously  $C'_1$  and  $C'_2$  are not dependently connected because they are connected through an overproduced species.

**Lemma 5** Let  $m, \bar{m} \in X$ .  $m$  is dependently connected in  $\langle X, \mathcal{R}_X \rangle$  to  $\bar{m}$  if and only if  $\bar{m}$  is dependently connected in  $\langle X, \mathcal{R}_X \rangle$  to  $m$ .

To continue, we need to mention that a computer science formalism called Petri Nets (Murata, 1989), has been considered as an interesting source of insights for the biochemical pathways research (Reddy et al., 1993, 1996). Petri Nets arose from the necessity to formalize concurrent processes. We will incorporate some fundamental topological parameters of Petri Nets to our analysis: the set of input transitions of a place and the set of input places of a transition.

**Definition 13** Let  $m \in \mathcal{M}$ . We define

$$\text{Act}(m, \mathcal{R}) = \{R = A \rightarrow B \in \mathcal{R} \mid m \text{ is present in } A\}.$$

We say  $\text{Act}(m, \mathcal{R})$  is the activable set of reactions of  $m$  in  $\mathcal{R}$ .

**Definition 14** Let  $R = A \rightarrow B \in \mathcal{R}$ . We define

$$\text{Req}(R) = \{m \mid m \text{ is present in } A\}.$$

We say  $\text{Req}(R)$  is the required set of species of  $R$ . Furthermore, for a set of reactions  $S \subseteq \mathcal{R}$  we define

$$\text{Req}(S) = \bigcup_{R \in S} \text{Req}(R).$$

The set of input places in Petri Nets corresponds to the set  $\text{Req}(\cdot)$ , and the set of input transitions corresponds to  $\text{Act}(\cdot, \mathcal{R})$ .

**Definition 15** We define  $\text{Causal}^*(m, \mathcal{R}_X)$  as the set of dependently connected species in  $\langle X, \mathcal{R}_X \rangle$  to  $m$ , and  $\text{Causal}(m, \mathcal{R}_X) = \text{Req}(\text{Act}(\text{Causal}^*(m, \mathcal{R}_X)))$ .

The following lemmas are derived straightforward from lemma 5 and definition 15:

**Lemma 6** Let  $m, \bar{m} \in \mathcal{M}$ .  $\bar{m} \in \text{Causal}^*(m, \mathcal{R}_X)$  if and only if

$$\text{Causal}^*(m, \mathcal{R}_X) = \text{Causal}^*(\bar{m}, \mathcal{R}_X).$$

**Lemma 7** Let  $R \in \mathcal{R}_X$ ,  $m, \bar{m} \in X - (E \cup F)$  s.t  $m \notin \text{Causal}(\bar{m}, \mathcal{R}_X)$ . If  $R \in \text{Act}(\text{Causal}(m, \mathcal{R}_X) \cap \text{Causal}(\bar{m}, \mathcal{R}_X), \mathcal{R}_X)$  then  $R \in \mathcal{R}_{E \cup F}$ .

$\text{Causal}^*(\cdot, \cdot)$  provides a way to split a set  $X$  of species in dependent connected subsets. It is necessary to identify a the catalytic set  $E$  and the overproduced set  $F$  to generate such separation. The more elements are in  $E \cup F$ , the more chance of recognize the independent causal connected sets we have.

**Lemma 8** Let  $D$  be the PAC of  $X$ . Then

$$D = \bigcup_{m \in D} \text{Causal}^*(m, \mathcal{R}_X)$$

**Proof** Note that  $D \subseteq \bigcup_{m \in D} \text{Causal}^*(m, \mathcal{R}_X)$ . Let  $m \in$

$\bigcup_{m' \in D} \text{Causal}^*(m', \mathcal{R}_X)$ , then for some species  $m' \in D$  we have  $m$  is dependently connected to  $m'$ , then  $m'$  is also dependently connected to  $m$ . This means  $m$  is a reactive, non overproduced, and non-catalytic species. Then  $m \in D$ .

From now on we let  $D \subseteq X$  be the potential active cycle of  $X$ .

**Definition 16** Any set  $D' \subseteq D$  s.t  $D = \bigcup_{m \in D'} \text{Causal}^*(m, \mathcal{R}_X)$  is called a base of  $D$ . Any  $m \in D'$  minimal cardinality base of  $D$  is called a minimal base of  $D$ .

**Lemma 9** Let  $D', D''$  be two minimal bases of  $D$ . Then every species in  $D'$  is dependently connected to one and only one species of  $D''$ .

**Proof** Let  $m \in D'$  and suppose that there is no species in  $D''$  dependently connected to  $m$ . By corollary 6 we have  $Causal^*(m, \mathcal{R}_X)$  is not contained in  $\bigcup_{m' \in D''} Causal^*(m, \mathcal{R}_X)$ . Then there has to be at least one species dependently connected to  $m$  in  $D''$ . Now suppose there is more than one species dependently connected to  $m$  in  $D''$ . Let  $m_1, m_2 \in D''$  such species. As  $m_1$  and  $m_2$  are dependently connected to  $m$ , then  $m_1$  and  $m_2$  are dependently connected. By corollary 6 we have  $Causal^*(m_1, \mathcal{R}_X) = Causal^*(m_2, \mathcal{R}_X)$ . Then  $D''$  is not a minimal base of  $D$ .

A minimal base of  $D$  is a set which generates all the non dependent sub-PACs of the  $D$ . We are going to prove that the self-maintenance of a potential active cycle can be obtained from the self-maintenance of its non dependent sub-PACs.

**Lemma 10** Let  $D'$  a minimal base of  $D$ . Then

$$Act(D, \mathcal{R}_X) = \bigcup_{m \in D'} Act(Causal^*(m, \mathcal{R}_X), \mathcal{R}_X).$$

**Proof** Note that

$$\bigcup_{m \in D'} Act(Causal^*(m, \mathcal{R}_X), \mathcal{R}_X) \subseteq Act(D, \mathcal{R}_X).$$

Let  $R \in Act(D, \mathcal{R}_X)$  then for some  $m \in D$  we have  $R \in (m, \mathcal{R}_X)$ . From definition 16 we have that there is  $m' \in D'$  s.t  $m \in Causal^*(m', \mathcal{R}_X)$ . Then by corollary 6 we have

$$\begin{aligned} R &\in Act(Causal^*(m', \mathcal{R}_X), \mathcal{R}_X) \\ &\subseteq \bigcup_{m \in D'} Act(Causal^*(m, \mathcal{R}_X), \mathcal{R}_X). \end{aligned}$$

### Decomposition Theorem for Organizations

**Theorem 2** Let  $D' = \{\bar{m}_1, \dots, \bar{m}_d\}$  a minimal base of  $D$ . For  $i = 1, \dots, d$  let

$$\begin{aligned} D_i &= Causal(\bar{m}_i, \mathcal{R}_X), \\ F_i &= Causal(\bar{m}_i, \mathcal{R}_X) \cap F. \end{aligned}$$

Let  $\emptyset \rightarrow Y = \{\emptyset \rightarrow y / y \in Y\}$ .  $X$  is self-maintaining if and only if for all  $i = 1, \dots, d$  we have that  $D_i$  is self-maintaining in the subnetwork  $\langle D_i, \mathcal{R}_{D_i} \cup \emptyset \rightarrow F_i \rangle$ .

**Proof**  $\Rightarrow$ : Let  $\bar{F} = \bigcup_{i=1}^d F_i$ . Note that  $X$  is self-maintaining in  $\langle X, \mathcal{R}_X \rangle$  if and only if  $X$  is self-maintaining in  $\langle X, \mathcal{R}_X \cup \emptyset \rightarrow \bar{F} \rangle$ . Let  $\mathbf{v}$  be a vector which verifies the self-maintenance of  $X$  in  $\langle X, \mathcal{R}_X \rangle$ . Let  $Act(D_i, \mathcal{R}_X) = \{R_{\alpha_1}, \dots, R_{\alpha_j}\}$ , then  $\bar{\mathbf{v}}$  lead to a non-negative production on all the species of  $Causal^*(\bar{m}_i, \mathcal{R}_X)$  where

$$\bar{\mathbf{v}}_i = \begin{cases} \mathbf{v}_i & \text{if } i = \alpha_j \text{ for some } j, \\ 0 & \text{else} \end{cases}$$

As the rest of species belong to  $\bar{F}$ , to reach their non-negative production we use the reactions in  $\emptyset \rightarrow \bar{F}$ .

$\Leftarrow$ : Let  $\mathbf{v}_1, \dots, \mathbf{v}_d$  the flux vectors which verifies the self-maintenance of  $\langle D_i, \mathcal{R}_{D_i} \cup \emptyset \rightarrow F_i \rangle$ ,  $i = 1, \dots, d$  and  $\mathbf{v}^F$  the flux vector which verifies the potential overproduction of  $F$  w.r.t  $X$ . Then there exist a non-negative number  $\beta$  s.t  $\beta \mathbf{v}^F + \sum_{i=1}^d \bar{\mathbf{v}}_i$  verifies the self-maintenance of  $X$ , where  $\bar{\mathbf{v}}_i$  is the flux  $\mathbf{v}_i$  represented as a flux vector for  $\mathcal{R}_X$ , i.e. completed with zeros in the coordinates representing reactions that are not in  $\mathcal{R}_{D_i}$ .

**Corollary 2** Let  $D'$  a minimal base of  $D$ . Then  $X$  can be non-overlapping decomposed (partitioned) as

$$X = N \cup E \cup F \cup D_1^* \cup \dots \cup D_d^*. \quad (4)$$

With  $D_i^* = Causal^*(m_i, \mathcal{R}_X)$ , and  $m_i$  the  $i$ -th element of  $D'$ . Moreover  $X$  is self-maintaining if and only if  $E \cup F \cup D_i^*$  is self-maintaining for  $i = 1, \dots, d$ .

### Stability, COT and Autopoiesis

Living organisms are systems far from thermodynamic equilibrium, therefore the question about stability refers to the conservation of the organization of the organism's processes. This issue requires some attention, as the processes an autopoietic system are homeostatic in essence, but they are always potentially subject to dramatic changes. A clear example of this is the cell cycle, which is driven by cyclic, i.e. non stationary, processes like the cyclins proteins family expression patterns. In relation to this, the autopoiesis theory describes living organisms as processes that produce the components that give rise to those processes, where some attributes are preserved and others may change (Varela et al., 1974) allowing a structural drift (Maturana and Mpodozis, 2000). Therefore, in a way, autopoietic systems are not obliged to exhibit stability in the long run.

The basis of structural drift in biochemical networks is metabolic regulation. This is managed by the modulation of enzymes expression and also by means of a co-catalysis phenomenon in which coenzymes and regulators interact with the enzyme structure-function relation. This metabolic regulation determines the cell's developmental *direction* between a wide range of possible organizations. In order to illustrate this idea, imagine a system *following* an attractor when suddenly the concentration of a given regulator reaches a level that triggers a deep change in the structure of the network. Now, the attractor mentioned above is absent and the system takes a different pathway in which another phase space shift can occur. If this process becomes cyclic, the entire loop can be described as a limit cycle, but a decomposition of phase space in a set of *contextual* mini phase spaces that emerge in different regulatory scenarios could contribute both to the understanding of biological operations as to the algorithms research in biology inspired AC simulations. Therefore, in the COT, the changes in structure can be represented as changes in the phase space, leading to a dependence of phase space with the state of enzymes and regulators.

Some future goals in biologically inspired AC are the study of cyclic behaviour and autopoiesis theory's structural coupling. Preliminary studies concerning artificial autopoietic systems have been done in (Peter et al., 2010), by making use of the P-systems formalism. A P-system is formed by a set reaction networks, each one enclosed by a membrane. The reaction networks can interact diffusing particles through membranes. In particular, it is shown how a bistable cyclic process can be reached among two different unstable reaction networks, by exchanging in low rates their subproducts, i.e. forming together an organization.

## Conclusions

In this paper we have shown how the chemical organization theory connects deeply to notions of autopoietic systems. Moreover, every autopoietic system is an organization, and thus, theorems derived for organizations are valid for autopoietic systems.

We introduce the notion of the role played by a species in a subnetwork of a reaction network. The different possible roles that a species can play in a subnetwork (non-reactive, catalyst, overproduced, active cyclic) give information about the structure of the subnetwork (lemmas 1, 3 and 4). We also introduce the notion of dependent connectivity in a reaction network (definition 12), which is useful to split a reaction network into the minimal parts required to verify the self-maintenance (Theorem 2). This theorem helps to simplify the organization verification not only by decomposing the set, but also when we keep track of the decomposition of the set to verify the self-maintenance of sets that contain it.

The fact that an organization can be subdivided into self-maintaining subnetworks which are mostly independent one from another (and not necessarily closed), is both striking and noteworthy. The decomposition theorem stated in this work shows that the long-term behaviour of an organization can depend on sets of species whose states are weakly coupled. This result opens new paths of analysis for a broad set of fields, from metabolic dynamics to ecological networks. This result also relates directly to a debated subject which is the composition of autopoietic systems by other autopoietic systems.

At this point a comment on the domain of applicability of this theorem is convenient. On some cases, like reactive flow systems where there is a spontaneous decay of every species, the decomposition is trivial, i.e., the system cannot be further subdivided. This is because each species of an organization would be overproducible (to counteract the decay). So, this paper addresses those systems where some species take part in reactions but do not decay spontaneously. In the case of living systems, every molecule decays spontaneously or, equivalently, dilutes when the system grows. This seems to make an argument towards the unapplicability of the decomposition theorem to living systems. However, the fact that every molecule decays in a

living system is only relevant at the right timescale. For a smaller timescale some molecules do not decay. Therefore, it is important to examine living systems at different timescales. Choosing a very long time scale, basically no system would continue to exist so that autopoiesis would not become visible. While with a smaller timescale more and more elements (molecules) would become stable and would not decay spontaneously. Here is where the decomposition can potentially be applied.

This decomposition theorem is suggested as a starting point for the complexity analysis of the organization verification problem as well as for the classification of reaction networks in terms of how complicated it is to compute their organizational structure. Although this is a relevant result for current systems biology as it simplifies analysis and simulations, and for artificial life, at it may guide the construction of artificial self-sustaining systems, it also demands revisiting questions on biological systems concerning their modularity, autonomy, and even the notion of their identity.

## References

- Benkő, G., Centler, F., Dittrich, P., Stadler, B. M. R., and Stadler, P. F. (2009). A topological approach to chemical organizations. *Artif. Life*, 15:71–88.
- Centler, F. and Dittrich, P. (2007). Chemical organizations in atmospheric photochemistries: A new method to analyze chemical reaction networks. *Planet. Space Sci.*, 55:413–428.
- Centler, F., Kaleta, C., Di Fenizio, P. S., and Dittrich, P. (2008a). Chemical organizations in a toy model of the political system. *Advances in Complex Systems*, 11(4):609–627.
- Centler, F., Kaleta, C., Di Fenizio, P. S., and Dittrich, P. (2008b). Computing chemical organizations in biological networks. *Bioinformatics*, 24:1611–1618.
- Cornish-Bowden, A., Cárdenas, M. L., Letelier, J. C., and Soto-Andrade, J. (2007). Beyond reductionism: Metabolic circularity as a guiding vision for a real biology of systems. *PROTEOMICS*, 7:839–845.
- Dittrich, P. and Di Fenizio, P. S. (2007). Chemical organization theory. *Bulletin of Mathematical Biology*, 69:1199–1231.
- Dittrich, P. and Winter, L. (2008). Computing chemical organizations in biological networks. *Bioinformatics*, 24(14):1611–1618.
- Eigen, M. and Schuster, P. (1977). A principle of natural self-organization. *Naturwissenschaften*, 64:541–565.
- Fontana, W. and Buss, L. (1994). The arrival of the fittest: Toward a theory of biological organization. *Bulletin of Mathematical Biology*, 56:1–64.
- Heinrich, R. and Rapoport, T. A. (1974). A linear steady-state treatment of enzymatic chains. *European Journal of Biochemistry*, 42:89–95.
- Hordijk, W. and Steel, M. (2004). Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227:451–461.

- Jaramillo, S., Honorato-Zimmer, R., Pereira, U., Contreras, D., Reynaert, B., Hernandez, V., Soto-Andrade, J., Crdenas, M., Cornish-Bowden, A., and Letelier, J. C. (2010). (m,r) systems and raf sets: Common ideas, tools and projections. In Fellermann, H., Drr, M., Hanczyc, M. M., Laursen, L. L., Maurer, S., Merkle, D., Monnard, P. A., Stoy, K., and Rasmussen, S., editors, *Artificial Life XII*, pages 94–100. MIT Press, Cambridge, MA.
- Kacser, H. and Burns, J. A. (1973). The control of flux. *Symposia of the Society for Experimental Biology*, 27:65–104.
- Kaleta, C., Centler, F., and Dittrich, P. (2006). Analyzing molecular reaction networks from pathways to chemical organizations. *Molecular Biotechnology*, 34:117–123.
- Kauffman, S. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, 22:437–467.
- Letelier, J. C., Marin, G., and Mpodozis, J. (2003). Autopoietic and (M,R) systems. *Journal of Theoretical Biology*, 222:261–272.
- Matsumaru, N., Centler, F., di Fenizio, P., and Dittrich, P. (2006). Chemical organization theory applied to virus dynamics. *Information Technology*, 48:154–160.
- Matsumaru, N., Centler, F., di Fenizio, P., and Dittrich, P. (2007). Chemical organization theory as a theoretical base for chemical computing. *International Journal on Unconventional Computing*, 3:285–309.
- Maturana, H. and Mpodozis, J. (2000). The origin of species by means of natural drift. *Revista chilena de historia natural*, 73(2).
- Maturana, H. and Varela, F. (1973). *De Máquinas y Seres Vivos: Una teoría sobre la organización biológica*. Editorial Universitaria.
- Murata, I. (1989). Petri nets and its applications. *Proceedings of the IEEE*, 77:541–580.
- Peter, S. and Dittrich, P. (2011). On the relation between organizations and limit sets in chemical reaction systems. *Advances in Complex Systems*, 14:77–96.
- Peter, S., Veloz, T., and Dittrich, P. (2010). Feasibility of organizations - a refinement of chemical organization theory. In: *M. Gheorghe, T. Hinze, G. Paun (Eds.), Proc. of the Eleventh International Conference on Membrane Computing*, pages 36–382.
- Reddy, V. N., Liebman, M. N., and Mavrovouniotis, M. L. (1996). Qualitative analysis of biochemical reaction systems. *Comput Biol. Med.*, 26:9–24.
- Reddy, V. N., Mavrovouniotis, M. L., and Liebman, M. N. (1993). Petri net representations in metabolic pathways. *ISMB-93 Proceedings*, pages 328–336.
- Rosen, R. (1958). A relational theory of biological systems. *Bulletin of Mathematical Biophysics*, 20:317–341.
- Schilling, C. H. and Palsson, B. O. (1998). The underlying pathway structure of biochemical reaction networks. *Proceedings of the National Academy of Sciences*, 95:4193–4198.
- Schuster, S., Dandekar, T., and Fell, D. A. (1999). Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnology*, 17:53–60.
- Strogatz, S. (2000). *Nonlinear Dynamics and Chaos*. Westview Press, Cambridge.
- Varela, F., Maturana, H., and Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *Biosystems*, 5:187–196.
- von Bertalanffy, L. (1968). *General System Theory: Foundations, Development, Applications*. George Braziller, New York.
- Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. Hermann and Cie Editeurs, Paris.

# A Loop Conjecture for Metabolic Closure

D.A. Contreras<sup>1,‡</sup>, U. Pereira<sup>1,2,‡</sup>, V. Hernández<sup>2,‡</sup>, B. Reynaert<sup>2,‡</sup>, J.C. Letelier<sup>2,‡</sup>

<sup>1</sup> Departamento de Física, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile.

<sup>2</sup> Departamento de Biología, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile.

letelier@uchile.cl

## Abstract

Although in the last few decades a variety of theoretical tools have been developed to better understand living organisms, their impact on experimental research has been rather limited. A common element between these theories is the idea of metabolic closure, i.e., the systems that produce all their metabolites and catalysts. In spite of an increasing consensus on the relevance of closure, a formal and operative definition has remained elusive. In this paper we revisit RAF sets and chemical organization theory and show how these two theories overlap and could help bring forth real world results. We also state a theorem ensuring the presence of a cycle of interdependent catalysts for RAF sets and conjecture that these cycles give stability to the network. This conjecture is illustrated and supported by computer simulations. Unavoidably, our viewpoint introduces the notion of fluxes and thus a temporal dimension to the purely algebraic model of RAF sets. The results of this work show that the incorporation of closure, topological and dynamical tools altogether is a promising path for a deeper understanding of living systems.

## Introduction

In the last thirty years there have been many efforts directed to develop theories to understand biological systems in terms of metabolic closure or, equivalently, systems that produce and maintain themselves. Two crucial models that definitively put metabolic closure at the very center of biological organization are: Autopoiesis, formulated by Maturana and Varela (Maturana and Varela, 1980), and Rosen's ( $M,R$ ) Systems (Rosen, 1958). But these two theoretical studies and similar theories (like the Chemoton or Autocatalytic sets), although very clarifying in basic aspects, have not yet produced technical results that illuminate the daily life of bench biologists involved in experimental research.

In the past (Jaramillo et al., 2010) we have emphasized that a little known formalism called RAF sets (Hordijk and Steel, 2004) is a particularly suited technical tool to understand closure in general and autocatalytic sets in particular. Here we study the relation between RAF sets and the chemical organization theory (COT), which is a theory that

adds to the dynamical aspects by introducing the notion of metabolic fluxes to the purely algebraic vision of RAF sets, an idea deeply embedded in basic metabolic engineering. This is accomplished by expressing the kinetic behaviour of the components (molecules) in terms of a stoichiometric matrix, which then leads directly to the concepts of rates and fluxes, introducing the temporal dimension. This approach can be used to expand the original RAF sets theory, which we consider to be highly valuable for biology, but unfortunately too algebraic to be of use, in particular lacking a way to describe the time behaviour of the systems, which is of most importance in the direction of a more realistic biological context.

Here we will show how notions from chemical kinetics can be fused with RAF sets to search for closure in metabolic networks. Although the results presented here seems, initially, as a mere technicalities without theoretical relevance, they open new research paths as we adjoin highly theoretical notions (RAF set and the metabolic closure) with an accepted used tool to understand metabolism in steady state. In particular we show the logical relation between COT and RAF sets.

## RAF sets and COT in a Nutshell

We now give a brief introduction to the work of Hordijk and Steel (2004), who came up with a formal framework to study a autocatalytic systems. Their main aim appears to have been to develop algorithms with which autocatalytic systems in Kaufmann's sense (1993) could be described and found computationally. They have produced a powerful approach that can be used to analyze a wide variety of systems. Their formalism is based on the following two important sets:  $X$  is the set of molecules involved in metabolism (i.e. metabolites, catalysts or external input material, termed *food set* in the formalism), and  $\mathcal{R}$  is the set of reactions that define the metabolic network. Each reaction  $r$  is represented as a tuple  $(A, B)$ , where  $A, B \subset X$ ,  $A \cap B = \emptyset$ ,  $A$  are the reactants and  $B$  the products of reaction  $r$ .

Further, to formalize the notion of catalysis, a specific set  $C$  (called the set of "catalyzations" by Hordijk and Steel)

<sup>‡</sup>All authors contributed equally

is introduced. Each catalyzation  $c$  is a tuple  $(x, r)$ , where  $x \in X$  is the catalyst and  $r \in \mathcal{R}$  is the reaction catalyzed by  $x$ . Additionally, the subset of molecules that are used but are not produced by metabolism is called food and denoted by  $F$ . Thus, a *catalytic reaction system* over a food source  $F$  is composed by a triplet  $\mathcal{L} = (X, \mathcal{R}, C)$  which defines the universe of molecules ( $X$ ), the reactions occurring among these molecules ( $\mathcal{R}$ ) and the identity of the catalyst involved in each reaction ( $C$ ). Note that this already provides, although at a very simple level, a way to refer to a system, and distinguish the inner and outer components and the transformations that the components undergo.

The following additional functions are defined:  $\rho(r) = A$  and  $\pi(r) = B$ , which return the reactants and the products of any given reaction  $r$ , respectively, and the function  $\text{supp}(r) = \rho(r) \cup \pi(r)$ . With the help of these elementary functions, the same notion can be extended to a *set of reactions*  $\mathcal{R}'$  as  $\rho(\mathcal{R}') = \bigcup_{r \in \mathcal{R}'} \rho(r)$ , where  $\mathcal{R}' \subseteq \mathcal{R}$ . This definition captures the conglomerate of molecules that participate as reactants for a set of reactions. A similar definition holds for  $\pi(\mathcal{R}')$ , the products of a subset of reactions. With these ideas we can define the closure of a subset  $X' \subseteq X$  relative to  $\mathcal{R}' \subseteq \mathcal{R}$  ( $\text{cl}_{\mathcal{R}'}(X')$ ) as the set of reachable molecules that can be synthesized by starting from  $X'$  and iteratively applying all the reactions in  $\mathcal{R}'$ . Note that this definition is of most importance, as it follows that a set of molecules which is closed (i.e. it is equal to its closure) under a set of reactions will not generate any new molecule and thus, conserves its identity. This operation captures the central idea of metabolic closure, which is fundamental for achieving organizational invariance in autopoietic systems. A *catalytic reaction system* is *reflexively autocatalytic* if for each  $r \in \mathcal{R}$  there is an  $x \in \text{supp}(\mathcal{R})$  such that  $(x, r) \in C$ . In other words, every catalyst must be a reactant or product of a reaction in the same system. The system is  $F$ -generated if every reactant is either produced by the system or incorporated as a food item (i.e. formally  $\rho(\mathcal{R}) \subseteq F \cup \pi(\mathcal{R})$ ). A system that is *reflexively autocatalytic* and  $F$ -generated is called a *RAF set* (see figure 1).

RAF sets can be understood informally as an interdependent set of biochemical reactions where all of the metabolites, with the exception of the so-called food set, are produced by the collection of reactions  $\mathcal{R}$ . This self generation, a defining feature of autopoietic and  $(M,R)$  systems, is the core of metabolic closure. Thus, RAF sets, autopoietic and  $(M,R)$  systems overlap to a great extent; positioning RAF sets as an operative theory to metabolic closure. The advantage of RAF set formalism is that it is precise enough to be coded in well defined algorithms that exploit its intrinsic recursiveness. To check if a given collection of biochemical reactions is indeed a RAF set, Hordijk and Steel (2004) developed algorithms aimed to analyze the interdependence between a given catalyst and its production pathway.

The chemical organization theory, initially developed by

Dittrich and Di Fenizio (2007), deals with chemical reaction networks. In what is called *static analysis*, the part of this theory that is concerned with the topology of the system, molecules and reactions are defined in a very similar way as in RAF sets. Most notably, both theories share the definition of the closure operator. But while COT makes no explicit mention to catalysts and therefore distances itself from biological systems in which this concept is fundamental, it does incorporate tools to study the dynamical behaviour of chemical reaction networks, thus provides a connection between the structure of a system and the dynamical aspects of it. This is accomplished by first expressing the system in terms of the stoichiometric matrix and associated differential equations.

In COT, it is useful to recognize systems fulfilling certain properties, such as closure. For example, a system is self-sustained if it is able to generate every molecule that is used up. When this topological consideration is transported to the time domain, we can define mass-maintaining systems. A system is said to be mass-maintaining when:

1. All reactions that can be fired by the molecules in the system occur at some positive rate
2. Reactions whose reactants are missing from the system do not occur
3. There is a combination of reaction rates such that all molecules increase or maintain their concentration.

A system which is both closed and mass-maintaining is called an organization. Organizations are interesting as they resemble very closely autopoietic systems. Also, organizations are at the center of many theorems in COT. This theory and RAF sets deal with closure. While one makes no distinction between catalysts and metabolites, the other one lacks the notion of time, which are essential elements of living systems. In the next paragraph we will show an relation between these two theories.

## Kinetics in RAF sets

If a theory is to have impact on real biochemical world, it must deal with the notions of that domain, thus, to gain a full understanding of closure we must complement the purely algebraic nature of RAF sets with ideas taken from Metabolic Control Analysis (MCA), a field generated to understand and measure fluxes in biochemical systems which is of common use in the field of metabolic engineering.

Current MCA is a quantitative theory that does not consider closure, as catalysts (i.e. enzymes) are placed in the network, but the reactions generating them are not taken into account. By putting the quantitative aspects of MCA and applying them to RAF sets, side by side, we can gain insight in how to study closure quantitatively. All the theories of metabolic closure (Autopoiesis,  $(M,R)$  systems, Autocatalytic sets, etc) are essentially algebraic or conceptual

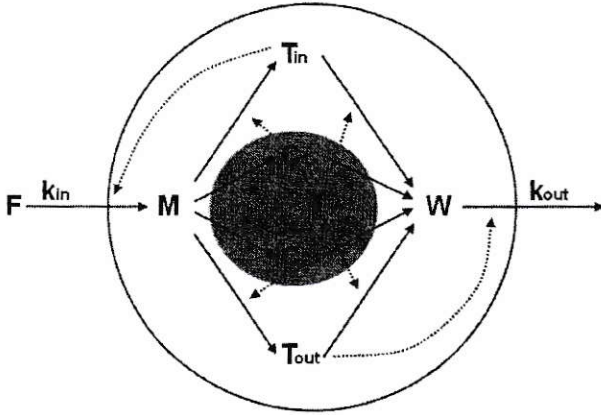


Figure 1: A simple example of a RAF set. Food elements  $F$  are incorporated into the system and generate metabolites  $M$ , which are transformed into two different catalysts sets: a)  $T_{in}$  which regulates the inflow of  $F$  and  $C_2$  which catalyzes its generation and b)  $T_{out}$  which regulates the outflow of waste metabolites  $W$  and  $C_1$  who catalyzes its formation. In addition,  $C_1$  and  $C_2$  catalyze the formation and destruction of the transporter catalysts ( $T_{out}$  and  $T_{in}$ , respectively), and also they mediate the generation and consumption of each other, forming the Reflexive Autocatalytic core of the system. Finally, growth is regulated by modulating the inflow of  $F$  and the outflow of  $W$ . We want to highlight the loop defined by metabolites  $M$  which turn into  $C_2$  who regulates the formation of  $C_1$  starting from  $M$ , a reaction regulated by  $C_2$ .

models centered on connectivity but not in dynamics. To go further in our understanding we must include the time course evolution of the concentrations inside the system.

Fortunately, the formalism of Reder (1988) that uses the stoichiometric matrix and the matrix  $D_x v$  to study rates, can be applied almost *verbatim* to analyze if a RAF set will grow or disappear. The great advantage of applying MCA formalisms is that we can quantitatively study how a system with metabolic closure can grow or disappear.

RAF sets are sets of coupled biochemical reactions with the attribute that the catalytic dependences between reactions and their catalysts are explicitly given. As said, RAF sets demand that almost all the molecules that conform a system can eventually be generated, directly or indirectly, from certain food materials and that all catalysts are produced by the system.

The *transformation* part of a RAF set can be represented by the formalism of the stoichiometric matrix, a well known tool extensively used in fields like MCA and Systems Biology in which every reaction is written as a column and every metabolite is referred to as a row. For example, the matrix  $N$  of the system described in figure 1 would be as following:

$$N = \begin{pmatrix} 1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

By using the column representation of reactions, it is convenient to define the addition of reactions as a standard addition operation of vectors. This operation expresses the occurrence of both reactions as a single net reaction.

Note that the *catalytic* part lies outside the stoichiometric matrix and cannot be deduced from it. But, in an idea that can be traced back at least to Reder (1988), the *catalytic* part can be represented by a matrix  $D_x v$  (also known as the Jacobian of the system) that contains all partial derivatives relating every reaction with every metabolite (or catalyst) in the system. Thus, the catalysts for a given reaction can be discovered by ranking the partial derivatives of the rate of this reaction respective to all metabolites (molecules) in the system. For example, the Jacobian matrix  $D_x v$  of the system described in figure 1 would be:

$$D_x v = \begin{pmatrix} \partial_M v_1 & 0 & 0 & 0 & \partial_{T_{in}} v_1 & 0 \\ \partial_M v_2 & 0 & \partial_{C_1} v_2 & \partial_{C_2} v_2 & 0 & 0 \\ \partial_M v_3 & 0 & \partial_{C_1} v_3 & \partial_{C_2} v_3 & 0 & 0 \\ \partial_M v_4 & 0 & \partial_{C_1} v_4 & 0 & \partial_{T_{in}} v_4 & 0 \\ \partial_M v_5 & 0 & 0 & \partial_{C_2} v_5 & 0 & \partial_{T_{out}} v_5 \\ 0 & \partial_W v_6 & \partial_{C_1} v_6 & \partial_{C_2} v_6 & 0 & 0 \\ 0 & \partial_W v_7 & \partial_{C_1} v_7 & \partial_{C_2} v_7 & 0 & 0 \\ 0 & \partial_W v_8 & 0 & \partial_{C_2} v_8 & \partial_{T_{in}} v_8 & 0 \\ 0 & \partial_W v_9 & 0 & \partial_{C_2} v_9 & 0 & 0 \end{pmatrix}$$

Every RAF set can then be described by two matrices;  $N$ , that shows the network connectivity, and  $D_x v$ , that quantifies catalizations.

The necessity of using the  $D_x v$  matrix to analyze RAF sets lies in the fact that autocatalysis is a phenomenon that does not depend *only* on connectivity. As it has been show recently by Plasson et al. (2010) and Piedrafita et al. (2010), the stability of an autocatalytic set depends on the relative rate of some reactions. Thus, two systems with identical connectivities but with different kinetics for some reactions can have vastly different behaviors.

As stated above, another theory concerned with formalizing biological organization is COT, a theoretical framework also centered in the idea of closure differing from RAF as the idea of catalyzation, perhaps the hallmark of RAF sets, is not considered. On the other hand COT brings an idea, the importance of fluxes in a network, that are not considered in RAF sets which is a purely algebraic approach to the description of biological organization. Thus an important question arises: can these two models be related? Can they support each other, in the sense of across fields fertilization? In the next section we clarify some relations between these two models.

An observation needed at the very beginning is that analysis using RAF sets and COT belong to two very different viewpoints as crucial elements in one theory are totally absent in the other. Thus as organizations (in the sense of COT) require that the overall flux across a relevant subset of reactions is maintained (thus avoiding the disappearance of crucial metabolites that, if absent, will produce network collapse). A mirror like situation can be stated with respect catalyzations, a cornerstone idea in RAF sets, and (surprisingly) an idea that is absent from COT. Thus we should expect that if a system is a RAF set it is not immediate that it is also an organization. Only in some especial conditions we should be able to find how these ideas can be concurrently applied.

## A hidden relation between RAF sets and COT

A further dissection of RAF sets shows that, although fluxes and reaction rates initially seem to be absent from this model, kinetic ideas do exist just below the surface. In effect we propose two lemmas and a theorem that will bring new light to the problem of comparing both approaches:

**Lemma 1** *If a catalytic reaction system  $\mathcal{L} = (X, \mathcal{R}, C)$  is  $F$ -generated, then for all metabolites  $x$  (including catalysts) produced by any reaction  $r \in \mathcal{R}$ ,  $x \in \text{supp}(\mathcal{R})$  there is a positive linear combination of reactions  $\bar{r}_x = \sum_i \alpha_i r_i$  such that the metabolite  $x$  belongs to the products of the reaction  $\bar{r}_x$ ,  $x \in \pi(r_0)$  and the reactants of  $\bar{r}_x$  belong to the Food set, i. e.,  $\rho(r_0) \subseteq F$ .*

**Proof:** Considering the algorithm used to find the closure of  $\mathcal{L}$  (Hordijk and Steel, 2004), let  $W = F$ . Then add

the products of reactions  $\mathcal{R}_0 = \{r \in \mathcal{R} | \rho(r) \subseteq W\}$  to  $W$ . Adding all reactions in  $\mathcal{R}_0$  gives a global net reaction  $\bar{r}_0$  that consumes metabolites from the Food set  $F$  only and produces each metabolite in  $W$ . If this process is repeated, considering  $W = F \cup \pi(\mathcal{R}_0)$ , it is possible to build the set  $\mathcal{R}_1$  that contains all reactions that have their reactants in  $W$ , but excluding the reactions from  $\mathcal{R}_0$ . Adding all reactions in  $\mathcal{R}_1$  we obtain a new reaction  $\bar{r}'_1$  that requires metabolites from  $W$  only and produces any metabolite in  $\pi(\mathcal{R}_1)$ . To obtain the fact that this last reaction uses only metabolites from the Food source, let  $\bar{r}_1 = \alpha \bar{r}_0 + \bar{r}'_1$ , where  $\alpha$  is the most negative stoichiometric coefficient of the reaction  $\bar{r}'_1$ . This procedure takes enough metabolites from  $F$  to generate  $\pi(\mathcal{R}_1)$ . If we repeat this algorithm until it is not possible to find new metabolites, we will have generated  $cl_{\mathcal{R}}(F) = W$ . If the system is  $F$ -generated, according to Hordijk and Steel (2004), we have that  $cl_{\mathcal{R}}(F) = F \cup \text{supp}(\mathcal{R})$ . We have shown that for every metabolite  $x \in cl_{\mathcal{R}}(F)$  a composite reaction exists which generates it consuming food items only, in fact it is one of the  $r_i$ .

**Lemma 2** *If a catalytic reaction system  $\mathcal{L} = (X, \mathcal{R}, C)$  is  $F$ -generated, there is a strict positive linear combination of reactions  $\hat{r} = \sum_i \alpha_i r_i$  with  $\alpha_i > 0$  such that all metabolites are products of this reaction, i. e.,  $\hat{r}$  is a strictly positive vector.*

**Proof:** From Lemma 1 it follows that for each metabolite  $m_j$  there is a positive linear combination of reactions  $\bar{r}_{m_j} = \sum_i \alpha_{i,j} r_i$  such that this metabolite is produced exclusively from the Food set. This linear combination  $\bar{r}_{m_j}$  is the resultant net reaction associated with the path of reactions  $j$  that generate each metabolite  $m_j$ . If for all metabolites we add their generating reactions  $\bar{r}_{sum} = \sum_i \alpha_i r_i$  with  $\alpha_i = \sum_j \alpha_{i,j}$ , we have from the Lemma 1 that  $\bar{r}_{sum}$  is strictly positive. We must note that not all reactions will be used. We refer to these reactions generally as  $r_s$ , having  $\alpha_s = 0$  in  $r_{sum}$  for those reactions. If we consider the sum of this reactions  $r_{not} = \sum_s r_s$ , it will consume metabolites. To maintain  $\hat{r}$  positive and still fire these non-essential reaction (this will be needed later), we set  $\hat{r} = \beta r_{sum} + r_{not}$  with  $\beta$  sufficiently large. Then we can construct a strictly positive linear combination  $\hat{r} = \beta r_{sum} + r_{not} = \sum_i \bar{\alpha}_i r_i$  with:

$$\bar{\alpha}_i = \begin{cases} 1 & \text{if } r_i \in \{r_s\} \\ \beta \alpha_i & \text{if } r_i \notin \{r_s\} \end{cases} \quad \text{note } \bar{\alpha}_i > 0$$

and all the metabolites are product of this reaction  $r_0$ .

These lemmas, framed completely in the language of RAF sets, could be interpreted as mere technical results about RAF sets. In essence they state that every metabolite can be generated from the food set and makes explicit the overall reaction producing each, non-food, metabolite. But every time we use a stoichiometric matrix  $N$  we are implying a given kinetics because of the necessary equation



relating  $N$  to the change of concentrations:  $N \cdot v = dX/dt$ , where  $v$  is the vector of rates. Thus the requirement, in COT, that  $(dX/dt \geq 0)$  can be phrased as a condition on the components of  $v$ . These lemmas show how some (not all) Organizations could be RAF sets, and it is interestingly that they are proved by using notions of linear algebra. Also note that the positive linear combination predicted by the lemmas explicitly shows how to combine individual reactions in any RAF set to attain mass-maintenance.

Once we have established this link we can a little bit further and search for deeper connections. The next theorem continues to exploit matrix  $N$  to sketch how some RA sets are F-generated using the stoichiometric matrix  $N$ .

**Theorem 1** *If a catalytic reaction system is F-generated, then there is a strictly positive rate vector  $v$ , such that  $N \cdot v = dX/dt$  is also strictly positive, where  $N$  is the stoichiometric matrix of the system.*

**Proof:** We note that the operation  $N \cdot v = dX/dt$  is equivalently mathematical to make a linear combination of reactions  $r = \sum_j \alpha_j r_j$  if we consider each reaction as a column and  $\alpha_j$  as the velocity of reaction  $r_j$ . From Lemma 2, if we take the reaction  $r_0$  and choose  $v_j = \bar{\alpha}_j$  (normalizing time units), therefore a  $v$  exists with components  $v_j > 0$  associated to a flux vector  $dX/dt$  that satisfies  $dX_i/dt > 0$ , equivalent to the column representation of  $r_0$  with all of their components also positive.

**Corollary 1** *If a catalytic reaction system is F-generated, then it is also an organization.*

**Proof:** An F-generated system is, by definition, *closed* and as theorem 2 shows, it also satisfies the property of *mass-maintenance*. Thus, it is an organization.

This theorem explains the existing relations between Organizations, F-generated sets, RA sets and RAF (see figure 2). Essentially, we have proved that all F-generated sets are *organizations* and a subset of them are also *reflexive autocatalytic*. This subset is the RAF sets. Theorem 1 is a simple one that has the virtue of illuminating how these two theoretical frameworks are related to each other.

This result is important because some new technical theorems have been obtained by Dittrich's group, for example, on how to detect organizations among real metabolic networks (Centler et al., 2010, 2008). Thus, our theorem shows that these new tools, developed to find organizations, could be also used to search for RAF sets.

In addition, we will make a definition to the sets that are organizations and RA at the same time.

**Definition 1** *If a catalytic reaction system is Reflexive Autocatalytic and an Organization, then it is a Reflexive Autocatalytic Organization, RAO.*

These sets are reaction systems that can be sustained, but not necessarily can be generated from a food set  $F$  exclusively. We have shown that all F-generated sets are orga-

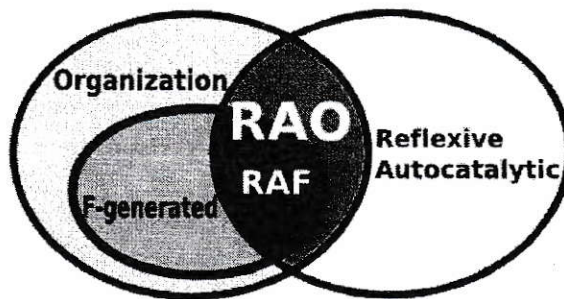


Figure 2: Venn's diagram depicting the logical relations between RA, RAF, and F-generated sets and *organizations* under COT's definition. All RAF sets are organizations, but whether all organizations are F-generated is an ambiguous matter.

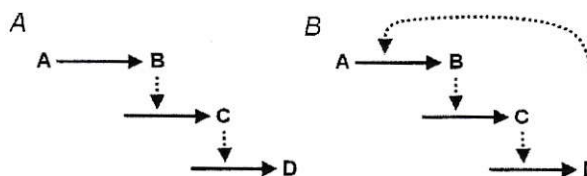


Figure 3: A: A chain of dependent catalyzations. B: A catalyzation loop

nizations, but the converse result (all organizations are F-generated sets) is more difficult to handle. We propose two different approaches: First, if one decides that the Food set corresponds only to the molecules generated from the empty set (in COT's phrasing of reactions), then it is clear that there are organizations which are not F-generated. On the other hand, for any organization it is always possible (due to the closure property) to find a suitable set (generally not unique)  $F$  such that the corresponding F-generated set is equal to the given organization. Thus, the extend to which organizations and F-generated sets overlap depends on which approach one takes to express COT systems in terms of RAF sets.

## The Loop Theorem

As most theories on biological organization are centered in the notion of closure (Letelier et al., 2011), RAF sets formalism give a succinct and useful description of closure. First, we shall consider a *chain* of catalyzations in which a product from one reaction catalyzes another reaction in the chain (figure 3A). If eventually a product catalyzes an earlier step (figure 3B), we have a catalyzation loop. As we shall see, in a RAO it is always possible to find such a catalyzation loop if the catalysts are not part of the food set. This condition seems natural for systems with metabolic closure.

Considering this definition we propose:

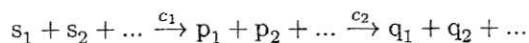
**Theorem 2** *If all catalysts in a RAO are generated by the system, then there is at least one catalyzation loop.*

**Proof:** In such a RAO every catalyst must be generated by a reaction, which in turn must be catalyzed too. In this sense, the production of every catalyst is directly dependent on another and indirectly dependent on a sequence of catalysts. The number of catalysts is finite therefore, at some point a catalyst must depend indirectly on itself, thus, forming at least one catalyzation loop. Note that not every catalyst is part of a loop as it is allowed that some catalysts may catalyze reactions which yield no catalysts as products, yet the system as a whole must have at least one catalyzation loop. Note that in case of direct autocatalysis, the loop is trivial. Also, whether there is more than one catalyzation loop is a question that must be addressed in each particular case.

An unsuspected consequence of the loop theorem is that some of the catalysts inside the catalyzation loop must have a dual catalytic role, that is enzymes that catalyze at least the creation of other two enzymes, if not happen the trivial case of all enzymes catalyze the creation of another one enzyme. This is interesting, as one modern re-interpretation of Rosen's results about how living systems avoid infinite regress is by having enzymes with dual functions (Letelier et al., 2006). Thus, this systemic result (i.e. existence of moonlighting enzymes) can be achieved by two different methods.

This theorem is a basic result that follows directly from the basic definitions of RAOs, but it shows an important property that needs to be underlined: the catalyzation loops (one or more) inside a RAO may be considered as its autocatalytic core and, functionally, there is a difference between the catalysts of the loop and the ones outside it.

We conjecture that the functional segregation hinted has important consequences. In effect, to confer stability to the core the catalysts outside it control the inflow and outflow of matter to and from the core. Thus, the net flow of matter inside it must be controlled, as a large flow would generate an exponential runaway and a small one would extinguish some core components, destroying its organization. Keeping this balance between in and outflow will be seen as homeostatic regulation. In summary, we conjecture that the catalyzation loop confers long term stability to the network. The analytical proof of this result seems difficult, but we did computer simulations in small (toy-like) systems and using mass-action kinetics, expressed for reactions:



By the formula:

$$\frac{d[p_k]}{dt} = k_1[c_1] \prod_i [s_i] - k_2[c_2] \prod_j [p_j]$$

Figure 4 shows one example for the temporal evolution of the concentrations of molecules for the RAF system of

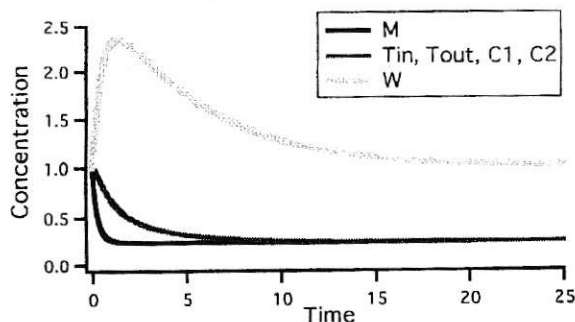


Figure 4: Temporal evolution of RAF toy system (see figure 1). The system reaches a steady state in which all concentrations are different from 0.

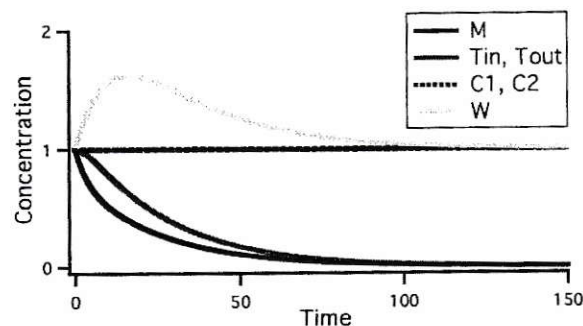


Figure 5: Temporal evolution of a non-RAF system. Although the concentrations of the catalysts  $C_1$  and  $C_2$  were fixed, the system decays until its components disappear.

figure 1. We can see that a steady state is achieved. In figure 5 we simulated a similar system but without the reactions that generate or destroy the catalysts  $C_1$  or  $C_2$ , removing the catalyzation loop and making the system a non-RAF set. In this last case the concentrations of many components decay to zero, stopping the network dynamics.

We also did the bifurcation analysis by varying rate constants  $k_{in}$  and  $k_{out}$  corresponding to the reactions  $r_1 : \phi \rightarrow M$  and  $r_9 : W \rightarrow \phi$  respectively. For the RAF set, almost every combination of parameters  $k_{in}$  and  $k_{out}$  leads to a steady state, except at the border where  $k_{in} = 0$  or  $k_{out} = 0$  and some regions close to these.

On the other hand, the non-RAF set has no stable points for most values of  $k_{in}$  and  $k_{out}$ . This puts in evidence the relevance of the autocatalytic core, so that the growth of one part of the system encompasses the rest and grows harmonically and coherently between the inflow and outflow of matter.

## Growth and Homeostasis in Autopoietic Systems

Any increase in the concentration of a loop catalyst will translate into an increased concentration of every other cata-

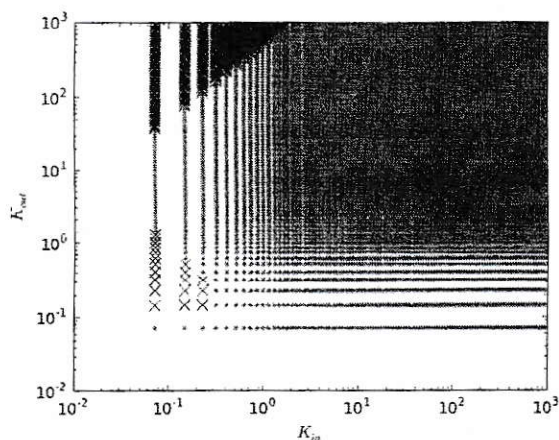


Figure 6: Bifurcation diagram of the RAF set in figure 1, for variables  $k_{in}$  and  $k_{out}$ . A similar diagram for a non-RAF system has no stable region. (cyan dots=stable, black Xs=unstable).

lyst in the loop, which would consequently lead to a further increment of the first one, exhibiting an apparent autocatalytic behavior. At the same time, this loop must be connected to side branches that lead to the production of catalysts that are not directly related to autocatalysis, but with the obtention and processing of the food sources that sustain it. An interesting type of branches are the ones that lead to the regulation of the enzymes that control the in and outfluxes, because they are supposed to regulate the whole metabolisms growth rate by coordinating these fluxes. This fact shows us the importance of the topology, because every enzyme must be whether part of a loop or a branch of it, thus, a change in an enzyme concentration which is part of the loop will have repercussions in the whole systems growth, as it also affects the enzymes that regulate the fluxes. Thus, the RAF sets may help to understand the dynamics of the homeostatic process. This is not mutually exclusive with the fact that an increase in the concentration of an enzyme outside the loop may have a direct repercussion on global growth.

The loop theorem has an important application for autopoietic systems, that can be defined as self-encapsulated RAF sets. For an autopoietic system (that according to the above theorem must contain at least one autocatalytic loop) to be stable in time, there must be a fine balance between the generation and destruction of molecules. But there must also be a balance respective to control its volume in order to keep the concentrations unaltered. Thus, the organization (*à la* Autopoiesis) of a RAF set must be under a precise homeostatic control, as growth must be promoted, but in the context of compensating for the volume increase without suffering the consequences of autocatalytic growth. Thus, in a first

approach, we must allow for a system to grow in terms of the net amount of molecules, but not in concentration. This implies that volume must be under active control and that allowing the system to grow would not be a contradiction with homeostatic principles.

## Discussion and Conclusion

As we have previously stated (Jaramillo et al., 2010) we conclude again that RAF sets formalism is particularly suited to study closure. Of course many aspects of metabolic closure escape this theory (the operator of organizational invariance  $\beta$  of  $(M,R)$  systems is a prime example), but this framework provides a solid starting point. The loop theorem proved here, which is a property shared by RAOs, Autopoietic and  $(M,R)$  systems is a good example of its power.

Another important point of the present study is to apply the analysis of COT to RAF sets. As it is usual in theoretical biology, the different frameworks generated to explain living organization exist in closed universes without dialogue between competing theories. Here, we partially break this isolation by showing how organizations *à la* COT contain all RAF sets, but not all RA sets. This inclusion, although obvious and expected from a theoretical viewpoint, is not easy to prove. We have developed demonstrations using arguments from linear algebra, instead of the set theory arguments favored in RAF. The most unexpected result is the uncovering of chemical kinetics arguments in RAF sets. In effect, RAF sets appear to be a purely algebraic entity, without considerations for time nor kinetics; but as soon as their stoichiometric matrix is expressed, the kinetic arguments of COT are made obvious. Thus our lemmas and theorems show deep relations between the pure algebraic formulation of RAF sets with the dynamics of organizations in COT. Perhaps this same reasoning could be also be applied to  $(M,R)$  systems. Taken together, the results shown here show the value of putting all the different notions of metabolic closure under a common analytical umbrella.

COT has already produced an interesting number of results on the dynamics of reaction networks, in particular regarding to the long-term temporal behaviour and stability of these systems (Dittrich and Di Fenizio, 2007). An interesting result from this theory, which complements the loop theorem presented here, is the decomposition theorem for organizations (Veloz et al., 2011). This theorem states that, under certain conditions, it is possible to split a system into subsystems whose dynamic behaviour are weakly coupled. Thus, an open question is to investigate how our loop theorem, which seems to indicate that systems cannot be segmented, is compatible with such uncoupling of subsystems. In effect a catalyzation loop might constitute the minimum decomposable unit.

We presented the conjecture that systems with at least one catalyzation loop are more stable than similar systems without such loops. This is a powerful result that will unavoid-

ably demand tools from MCA, the most elaborate theory about fluxes in biochemical networks, to be proven or refuted.

In summary, our efforts show that closure is a conceptual key to understand biological organization, as an example we have come close to use closure as an argument to prove one theorem (loop theorem), which we believe is a valuable conceptual step and a fertile direction for theoretical biology.

### Acknowledgements

We would like to thank the reviewers for their comments. They were very helpful to improve this manuscript. Funded by FONDECYT grant 1110247, CONICYT scholarship 22101236 and 22110804.

### In memoriam

This work reflects the school of thinking generated by Francisco Varela in Chile during the 70's. Francisco was a valued friend, a guide and an inspiration for new generations of biologists.

### References

- Centler, F., Kaleta, C., Speroni di Fenizio, P., and Dittrich, P. (2008). Computing chemical organizations in biological networks. *Bioinformatics*, 24(14):1611 – 1618.
- Centler, F., Kaleta, C., Speroni di Fenizio, P., and Dittrich, P. (2010). A parallel algorithm to compute chemical organizations in biological networks. *Bioinformatics*, 26(14):1788 – 1789.
- Dittrich, P. and Di Fenizio, P. S. (2007). Chemical organization theory. *Bulletin of Mathematical Biology*, 69:1199–1231.
- Hordijk, W. and Steel, M. (2004). Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227:451–461.
- Jaramillo, S., Reynaert, B., Hernández, V., Soto-Andrade, J., Cárdenas, M., Cornish-Bowden, A., and Letelier, J. (2010). (M,R) Systems and RAF sets: Common ideas, tools and projections. *Proceedings of the Twelfth International Conference on the Synthesis and Simulation of Living Systems*, pages 94 – 100.
- Kauffman, S. (1993). *The Origins of Order*. Oxford University Press, New York.
- Letelier, J. C., Cárdenas, M. L., and Cornish-Bowden, A. (2011). From *L'Homme Machine* to metabolic closure: steps towards understanding life. *Journal of Theoretical Biology* (in print).
- Letelier, J. C., Soto-Andrade, J., Guíñez Abarzúa, F., Cornish-Bowden, A., and Cárdenas, M. L. (2006). Organizational invariance and metabolic closure: analysis in terms of (M,R) systems. *Journal of Theoretical Biology*, 238:949–961.
- Maturana, H. and Varela, F. (1980). *Autopoiesis and Cognition: the Realisation of the Living*. D. Reidel Publishing Company, Dordrecht, The Netherlands.
- Piedrafitá, G., Montero, F., Morán, F., Cárdenas, M. L., and Cornish-Bowden, A. (2010). A simple self-maintaining metabolic system: Robustness, autocatalysis, bistability. *PLoS Comput Biol*, 6(8):e1000872.
- Plasson, R., Brandenburg, A., Jullien, L., and Bersini, H. (2010). Autocatalyses. *Proceedings of the Twelfth International Conference on the Synthesis and Simulation of Living Systems*, pages 4 – 11.
- Reder, C. (1988). Metabolic control theory: A structural approach. *Journal of Theoretical Biology*, 135(2):175 – 201.
- Rosen, R. (1958). The representation of biological systems from the standpoint of the theory of categories. *Bulletin of Mathematical Biology*, 20(4):317–341.
- Veloz, T., Reynaert, B., Rojas, D., and Dittrich, P. (2011). A decomposition theorem in chemical organizations. *Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems (this volume)*.

# Cycles and the Qualitative Evolution of Chemical Systems

Peter Kreyssig<sup>1</sup>, Gabi Escuela<sup>1</sup>, Bryan Reynaert<sup>2</sup>, Tomas Veloz<sup>3</sup>, Bashar Ibrahim<sup>1,4\*</sup>, Peter Dittrich<sup>1\*</sup>

<sup>1</sup>Bio Systems Analysis Group, Institute of Computer Science, Jena Centre for Bioinformatics and Friedrich Schiller University, Jena, Germany, <sup>2</sup>Faculty of Biological Sciences, University of Chile, Santiago, Chile, <sup>3</sup>Department of Psychology and Computer Science, University of British Columbia, Kelowna, Canada, <sup>4</sup>German Cancer Research Center, DKFZ-ZMBH Alliance, Heidelberg, Germany

## Abstract

Cycles are abundant in most kinds of networks, especially in biological ones. Here, we investigate their role in the evolution of a chemical reaction system from one self-sustaining composition of molecular species to another and their influence on the stability of these compositions. While it is accepted that, from a topological standpoint, they enhance network robustness, the consequence of cycles to the dynamics are not well understood. In a former study, we developed a necessary criterion for the existence of a fixed point, which is purely based on topological properties of the network. The structures of interest we identified were a generalization of closed autocatalytic sets, called chemical organizations. Here, we show that the existence of these chemical organizations and therefore steady states is linked to the existence of cycles. Importantly, we provide a criterion for a qualitative transition, namely a transition from one self-sustaining set of molecular species to another via the introduction of a cycle. Because results purely based on topology do not yield sufficient conditions for dynamic properties, e.g. stability, other tools must be employed, such as analysis via ordinary differential equations. Hence, we study a special case, namely a particular type of reflexive autocatalytic network. Applications for this can be found in nature, and we give a detailed account of the mitotic spindle assembly and spindle position checkpoints. From our analysis, we conclude that the positive feedback provided by these networks' cycles ensures the existence of a stable positive fixed point. Additionally, we use a genome-scale network model of the *Escherichia coli* sugar metabolism to illustrate our findings. In summary, our results suggest that the qualitative evolution of chemical systems requires the addition and elimination of cycles.

**Citation:** Kreyssig P, Escuela G, Reynaert B, Veloz T, Ibrahim B, et al. (2012) Cycles and the Qualitative Evolution of Chemical Systems. PLoS ONE 7(10): e45772. doi:10.1371/journal.pone.0045772

**Editor:** Claude Prigent, Institut de Génétique et Développement de Rennes, France

**Received:** March 5, 2012; **Accepted:** August 22, 2012; **Published:** October 11, 2012

**Copyright:** © 2012 Kreyssig et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the German Research Foundation priority programs InKoMBio (SSP 1395, Grant DI 852/10-1), Organic Computing (SPP 1183, Grant DI 852/4-3) and the European Commission NeuNeu project (248992). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: bashar.ibrahim@uni-jena.de (BI); peter.dittrich@uni-jena.de (PD)

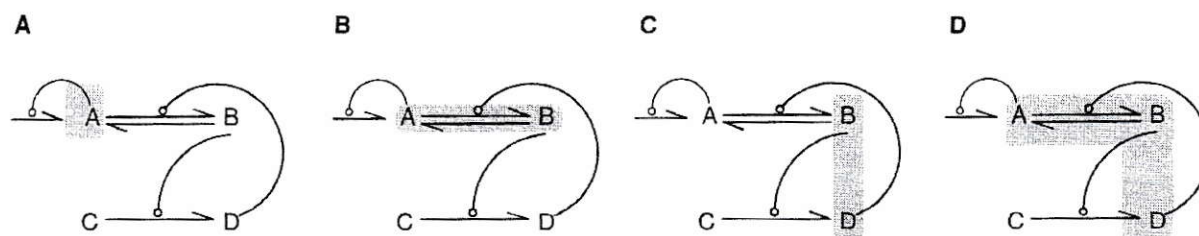
## Introduction

Many mechanisms are characterized by the presence of cycles, especially those that play central roles in biological systems. Specific functions such as regulation, memory and differentiation have been associated to cycles [1–4]. Furthermore, systems containing cycles exhibit robustness to environmental changes, which is a key feature for the evolution of biochemical networks [5,6]. Cycles vary in appearance from simple feedback loops to complex ones [7] or large cycles, these include some signaling cascades [8], and the Krebs and Calvin cycles [9,10]. Frequently, simple cycles are considered as network motifs [11] and therefore, they are analyzed as isolated modules neglecting the role of the molecular environment. In contrast, it has been proposed that the topological structure of a subgraph alone cannot determine its effects over the whole network [7,12]. Moreover, some authors have found interesting results on feedback loops and have concluded these as prerequisites for multistability in gene regulation and mixed networks [1,2], as well as in metabolic networks [4,13]. However, the existence of cycles as well as their necessity and contribution to stability in networks remains largely elusive. Additionally the question of how to analyze large systems,

in which classical approaches like differential equations fail, is open.

Chemical organization theory (COT) offers a lucid formalism with novel methods to analyze complex systems and their dynamics at specific states. It can be particularly useful for analyzing large scale models and can be applied in a broad range of disciplines such as political, social, biological and chemical systems. COT has been developed for the last 20 years [14–16] and right from the beginning it has focused on two central notions, namely the existence of closure and self-maintenance. These notions contribute to the further understanding and characterization of relevant systemic properties, since COT developed a relation between both topological and dynamical aspects of reaction networks [17].

On the other hand, reflexive autocatalytic food-generated (RAF) sets theory [18] is another formal framework, which contains some indications on the necessity of the existence of a cycle [13] for the stability of living systems. A RAF set is a set of molecules which can be produced from an initial set of food molecules in a reflexive autocatalytic way, i.e. every reaction is catalysed by at least one molecular species. RAF sets and COT share similarities as they are based on set theory and even coincide in some definitions, e.g.



**Figure 1. Example for the definitions of reaction networks and cycles.** There are four cycles in this example network, given by  $\mathcal{M} = \{A, B, C, D\}$  and  $\mathcal{R} = \{A \rightarrow 2A, A + D \rightarrow B + D, B \rightarrow A, C + B \rightarrow D + B\}$ , shown in shaded areas. Panels A to D show cycles  $\{A\}, \{A, B\}, \{B, D\}, \{A, B, D\}$  respectively. The arrow  $\rightarrow$  refers to a reaction while the arrow  $\dashrightarrow$  indicates the catalytic effect on a reaction. doi:10.1371/journal.pone.0045772.g001

closure. The main difference between these formalisms resides on how dynamics are approached, since RAF sets place catalysis as the only element concerned with kinetics and occurrence of reactions. The requirements that every reaction should be catalysed and all catalysts should come from within the reaction network is very strong. However it is sensible since many networks in biology fulfill this condition. In contrast, COT focuses on the connections between the stationary states and the topological and stoichiometric properties of the reaction network [13,17] while making less restrictive hypotheses, especially none on the structure of the reactions directly. This leads to a greater generality.

We give a definition of cycles in reaction networks based on a directed graph extracted from the reaction network (Definition 1). Using this definition, we prove the existence of a cycle to be a necessary condition of non-trivial stationary states in reaction networks. This is our main result, formulated as Lemma 1. It also yields a potential way to create or alter a non-zero stationary state of a system by adding a cycle to a closed, but not necessarily self-maintaining, set of molecules. Thereby the system is changed qualitatively since in the new fixed point a different composition of molecular species is present.

When employing this COT-based analysis for systems that contain cycles, we found that there are some interesting properties that can be easily deduced, including the relation between the hierarchical structure of the set of chemical organizations and the existence of cycles. Moreover, in this work we rediscover the abundance of cycles in biological systems. Considering the BioModels Database [19], we see that more than 60% of the models have more than one organization [20], and thus we can deduce from Lemma 0 that they contain cycles. This insight encourages us to believe that COT will be able to contribute to the further understanding of cycles in reaction networks. Hence we focus on the combination of COT with results about the existence of cycles (i.e. an overlay of COT and cycles).

We apply our main result to four examples of different nature. Firstly, we analyze the cycle found in a particular autocatalytic network, which is a generalization of the system discussed in [13], and investigate its stability properties further using COT, the Jacobian and numerical simulations. Secondly and thirdly, in order to provide some realistic examples for such autocatalytic networks, we discuss two real biological models of mitotic control mechanisms in detail. One of these has already been published [21] and one was newly constructed for this study. These examples support our ideas and illustrate how they can be used to study biological systems. Fourthly, we further investigate the consequences of Lemma 1 for a large network, namely the sugar metabolism model proposed in [22].

## Methods

We shortly summarize the needed definitions and results from COT in informal terms, still providing a mathematically precise version in the Appendix. A tool for the computation of chemical organizations is freely available on our website <http://www.biosys.uni-jena.de/Services.html>.

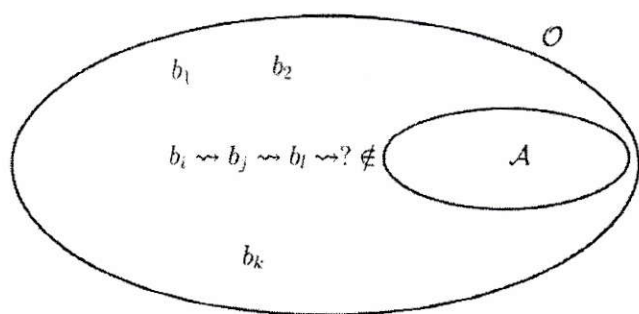
The pair  $(\mathcal{M}, \mathcal{R})$  is called *reaction network* and we call  $\mathcal{M}$  the set of *molecules* (or *species*) and  $\mathcal{R}$  the set of *reactions*. Each reaction  $(l, r) \in \mathcal{R}$  consists of a left hand side  $l$  and right hand side  $r$  and will be denoted also by  $l \rightarrow r$  in concordance with the chemical vocabulary. A reaction  $(l, r) \in \mathcal{R}$  occurs when the molecules in  $l$  are consumed to produce the molecules in  $r$  (what this actually means has to be made precise for every dynamics separately, e.g. the application of a reaction in stochastic dynamics is different from continuous dynamics).

Figure 1 depicts the network  $(\mathcal{M}, \mathcal{R})$  consisting of the set of species  $\mathcal{M} = \{A, B, C, D\}$  and the set of reactions  $\mathcal{R} = \{A \rightarrow 2A, A + D \rightarrow B + D, B \rightarrow A, C + B \rightarrow D + B\}$ .

Furthermore we define  $\text{supp}(l, r)$ , the *support*, and  $\text{prod}(l, r)$ , the *product*, of  $(l, r)$  to be the set of species occurring on the left hand side and on the right hand side, respectively. We emphasize that the support and product are sets, meaning that the *multiplicity* of each molecular type in the reaction does not matter. The sets  $\text{supp}(l, r)$  and  $\text{prod}(l, r)$  only collect the species that are present (i.e. have a multiplicity greater or equal than one) on the left and right hand side respectively. In our example this means that the product of the first reaction  $A \rightarrow 2A$  is the set  $\{A\}$  and not the multiset  $\{2A\}$ . Unfortunately the notion of a support with the same notation is also used in RAF sets theory [13,18] for all the molecules on both left hand side and right hand side. Within this article the usage of  $\text{supp}$  is consistent with COT.

Let  $\mathbf{A}$  be a subset of  $\mathcal{M}$ . We define  $\mathcal{R}_{\mathbf{A}}$  as the set of reactions applicable to  $\mathbf{A}$ , or in other words the reactions using only species from  $\mathbf{A}$ . Using this terminology we define a subset  $\mathbf{C}$  to be closed if by applying reactions from  $\mathcal{R}_{\mathbf{C}}$  we do not get molecules outside  $\mathbf{C}$ . In other words, in a reactor containing molecules from  $\mathbf{C}$ , we will never find molecules not already in  $\mathbf{C}$ . The set  $\{A\}$  in the example is therefore closed. If we only have molecules of type  $A$  there is no reaction producing new species (for the reaction  $A + D \rightarrow B + D$  to be applicable, for example, we would need molecules of type  $D$  as well). Also we call  $\mathbf{S}$  *semi-self-maintaining* if a reaction application consumes a species, there is a reaction producing this species. Combining these two definitions we arrive at the notion of a *semi-organization* which is a closed and semi-self-maintaining set.

Semi-organizations entail a topological (non-stoichiometric) notion of stability in reaction networks. On the one hand, the closure property ensures that a semi-organization will not produce novel species. On the other hand the semi-self-maintenance ensures that the species that are consumed within the network are



**Figure 2. Situation in the recursion argument of the proof of Lemma 1.**

doi:10.1371/journal.pone.0045772.g002

also produced. However, although semi-organization is a necessary condition for the occurrence of a fixed point, semi-organizations does not necessarily ensure permanence, i.e. consider the reaction network  $(\{A, B\}, \{A \rightarrow B, 2B \rightarrow A\})$  with the semi-organization  $\{A, B\}$ .

In order to capture more of the notion of fixed states in a reaction network, we need to consider the stoichiometry of the reactions, and verify if the network can maintain itself, i.e. produce all the species at a higher or equal rate to what they are consumed. The *stoichiometric matrix* represents the numerical relation between production and consumption of species in a reaction network, and it is denoted as  $N_A$ , where  $A \subseteq \mathcal{M}$ .  $N_A$  is calculated taking the difference between the right and left hand sides of each reaction in  $\mathcal{R}_A$ .

We say  $A$  is *self-maintaining* if and only if all the reactions from  $\mathcal{R}_A$  can occur at a certain strictly positive rate without decreasing the concentration of any species of  $A$ . A subset of  $\mathcal{M}$  is a *chemical organization* if it is closed and self-maintaining [14,16].

The term organization refers to a more stringent condition than semi-organization. Indeed, every organization is a semi-organization [16]. Moreover, organizations bridge the reaction network with its dynamical behavior. Given a reaction network and a kinetic law (e.g. mass action kinetics), it has been proven that the fixed points of the obtained ordinary differential equation (ODE) system, relate to the set of organizations as follows. For every fixed point the set of molecules with concentration higher than zero in that fixed point is an organization of the reaction network [16,17].

In general, an organization may contain a species that does not take place in any reaction within this organization. Hence, it is reasonable to introduce the notion of a *reactive set* of species. A set  $A$  of species is reactive if and only if for all  $A \in A$  there exists a reaction  $(l, r)$  in  $\mathcal{R}_A$  such that  $A \in \text{supp}(l, r) \cup \text{prod}(l, r)$ . Note that the dynamical properties of a set of species depends only on its largest reactive set, as the other species do not react and remain with constant concentration. Furthermore, the number of organizations in most reaction networks is considerably larger than the number of reactive organizations, and hence focusing on the set of reactive organizations instead of the whole set of organizations simplifies the understanding of the dynamical behavior of a reaction network [23]. From now on, we will consider only reactive sets of species.

## Results

Our definition of cycle can be explained using the usual definition for cycles in directed graphs. Given a reaction network we can deduce a directed graph in the following way. The nodes of that graph are the species of the reaction network, and there is an

edge from species  $A$  to species  $B$  if there is a reaction with  $A$  on its left hand side and  $B$  on its right hand side. We say that  $A$  is directly-causally connected to  $B$ . The cycles found in this directed graph are the ones we identify as cycles of the reaction network in our terminology (see Figure 1). This definition captures a variety of specific forms of cycles, yet yielding a structural result about reaction networks.

### Definition 1

Two species  $A, B \in \mathcal{M}$  are directly-causally connected, if there exist a reaction  $(l, r) \in \mathcal{R}$  with  $A \in \text{supp}(l, r)$  and  $B \in \text{prod}(l, r)$ , i.e.  $A$  is on the left hand side of a rule and  $B$  on its right hand side. We write  $AB$ . We say that the network contains a cycle if there is a subset of species  $\{M_1, \dots, M_k\} \subseteq \mathcal{M}$  such that  $M_i M_{i+1}$  for  $i = 1, \dots, k-1$  and  $M_k M_1$ . We write  $M_1 M_2 \dots M_k M_1$  and say that  $\{M_1, \dots, M_k\} \subseteq \mathcal{M}$  is a cycle.

We can now formulate our main theoretical result:

### Lemma 1

Given a reaction network and a reactive semi-self-maintaining set  $S$  that strictly contains a closed set  $C$ , i.e.  $C \subseteq S$ , then  $S - C$  contains a cycle.

**Proof.** Let  $B = S - C = \{B_1, \dots, B_k\}$  for some  $k \geq 1$  as  $C \subseteq S$ . We will prove first that at least one reaction in  $\mathcal{R}_S$  consumes species from  $B$ . Suppose that no species in  $B$  is consumed by any reaction in  $\mathcal{R}_S$ . As  $S$  is reactive, then each species  $B \in B$  must be produced by some reaction  $(l, r) \in \mathcal{R}_S$ . However, as no species in  $B$  is consumed, we have that either  $\text{supp}(l, r) = \emptyset$  or  $\text{supp}(l, r) \subseteq C$ . This contradicts the hypothesis that  $C$  is closed, since elements from  $C$  would produce an element which is not in  $C$ . Hence, at least one species, say  $B_i$ , is consumed within  $\mathcal{R}_S$ .

As  $S$  is semi-self-maintaining,  $B_i$  is produced by a reaction  $(l, r) \in \mathcal{R}_S$ . Moreover,  $\text{supp}(l, r)$  must contain at least one species  $B_j \in B$  since  $C$  is closed. There are two possible cases.

- First case:  $B_j = B_i$ . By definition  $B_i B_i$  and we have a (trivial) cycle.
- Second case:  $B_j \neq B_i$ . By definition we have  $B_j B_i$ .

Since the same argument applies for  $B_j$  as well, by a recursion argument we prove that there exists a cycle (see Figure 2 for a schematic representation).

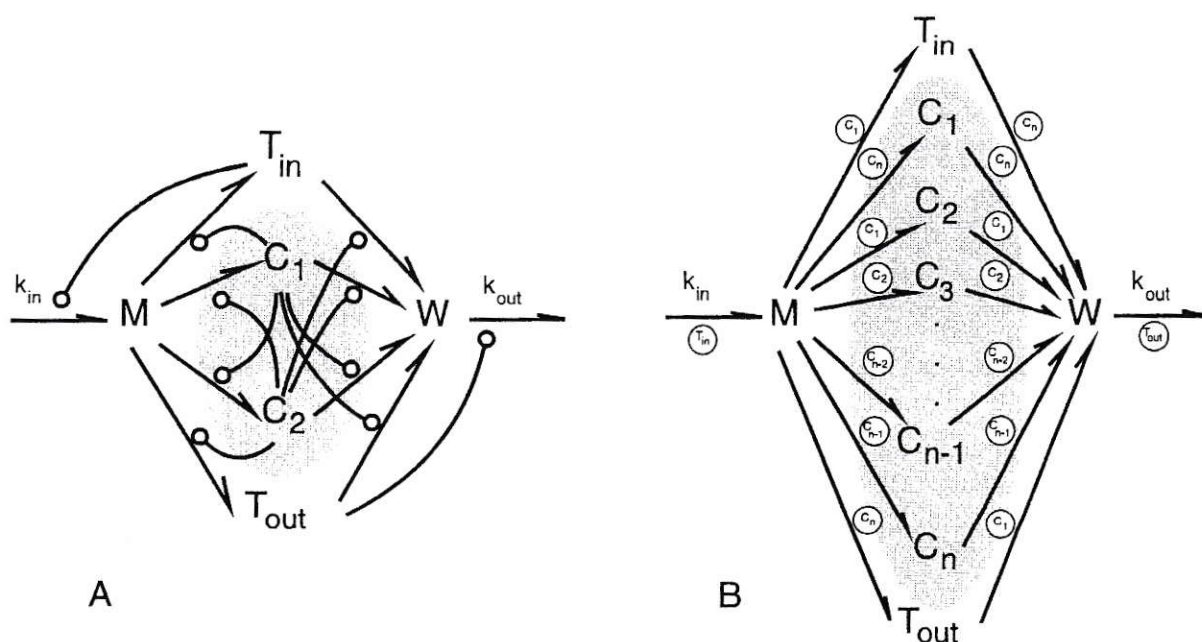
Lemma 1 connects the existence of reactive semi-self-maintaining sets with the existence of cycles. It states that every reactive semi-self-maintaining set that strictly contains a closed set must contain a cycle made from molecules not contained in the closed set. Why do we require that there is a closed set? Let us consider the example  $\mathcal{R} = \{\rightarrow A, A \rightarrow B, B \rightarrow\}$ . The set  $\{A, B\}$  is self-maintaining, but does not contain a cycle. Note that it does not strictly contain any closed subset, because the closure of the empty set is  $\{A, B\}$ . Since the closure of the empty set is contained in all closed sets we can say that a reactive semi-self-maintaining set whose molecules cannot be made from the inflow must contain a cycle.

Note that every organization is semi-self-maintaining and closed. Thus, equating in Lemma 0 the closed set  $C$  with an organization  $O_1$  and the reactive semi-self-maintaining set  $S$  with a reactive organization  $O_2$  we arrive at the following corollary.

### Corollary

Given a reaction network and two different reactive organizations  $O_1$  above  $O_2$ , i.e.  $O_2 \subseteq O_1$ , then  $O_1$  contains at least one more cycle than  $O_2$ .

From Corollary 1 follows that any reactive organization, except the ones produced from the inflow of the network, contains at least one cycle. Or in other words only the smallest might not contain a



**Figure 3. Particular autocatalytic network.** (A) Original network presented in [13] and (B) a generalization, used for our analysis.  $M, T_{in}, T_{out}, W, C_1, C_2, \dots, C_n$  represent species. Arrows indicate reactions between the species, arrows ending with a circle denote catalysis (to reduce its complexity, in Panel (B) the catalysts but not the arrows are shown). The parameters  $k_{in}$  and  $k_{out}$  represent reaction rates. Observe that  $M$  and  $W$  are the external supplies and waste species,  $T_{in}$  and  $T_{out}$  represent transport molecules, whereas  $C_1, C_2, \dots, C_n$  conform a cycle, i.e.  $C_1 C_2 \dots C_n C_1$ . doi:10.1371/journal.pone.0045772.g003

cycle, like in the previous example. The shifting or moving from one reactive organization and therefore from one potentially stable constellation of molecules to another was already described in [24]. According to Lemma 0 this movement must involve the addition or removal of at least one cycle. In particular this can be achieved in two ways. Combining the species of two organizations with disjoint cycles is the first possibility. The second is adding one or more species to an organization to create new cycle. As an example we refer to Section “A Particular Autocatalytic Network” The reverse operations lead to the “destruction” of organizations. Thus, any qualitative transition from a stable state to another with a different composition of reacting molecular species must encompass the removal or addition of a cycle.

Note that Corollary 1 is also true for *overproducible* organizations, i.e. organizations where each species can be produced at a strictly positive rate, because overproducible organizations are reactive sets.

Moreover, from Corollary 1 we can also infer that every organization is uniquely identified by its set of cycles, and if the set of cycles of an organization  $O_1$  contains the set of cycles of another organization  $O_2$ , then  $O_2 \subseteq O_1$ . This implies that the hierarchy of the set of organizations [16] can be mapped to a hierarchy of the set of cycles.

Considering RAF set theory, not every RAO set is necessarily a reactive set. The other hypothesis is satisfied by all RAOs though, since for every catalyst  $C \in O$  the subset  $\{C\}$  is closed. Therefore Lemma 0 and the cycle theorem proven in [13] are complementary results.

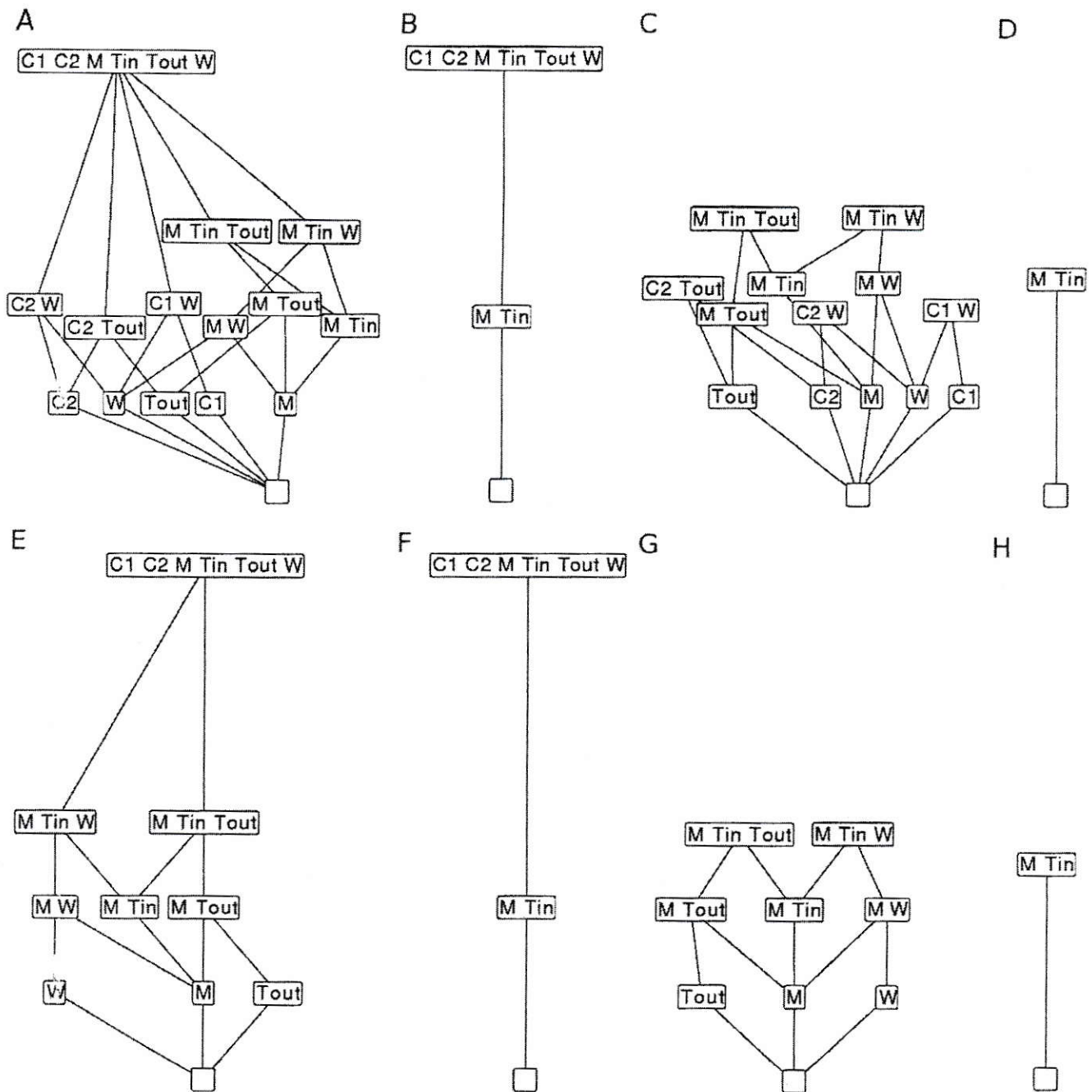
### Theoretical and Biological Applications

**A Particular Autocatalytic Network.** In this section we apply Lemma 1 and show how our algebraic analysis can be complemented by a dynamical analysis. We introduce a prototypical model (Figure 3) that captures properties of reaction

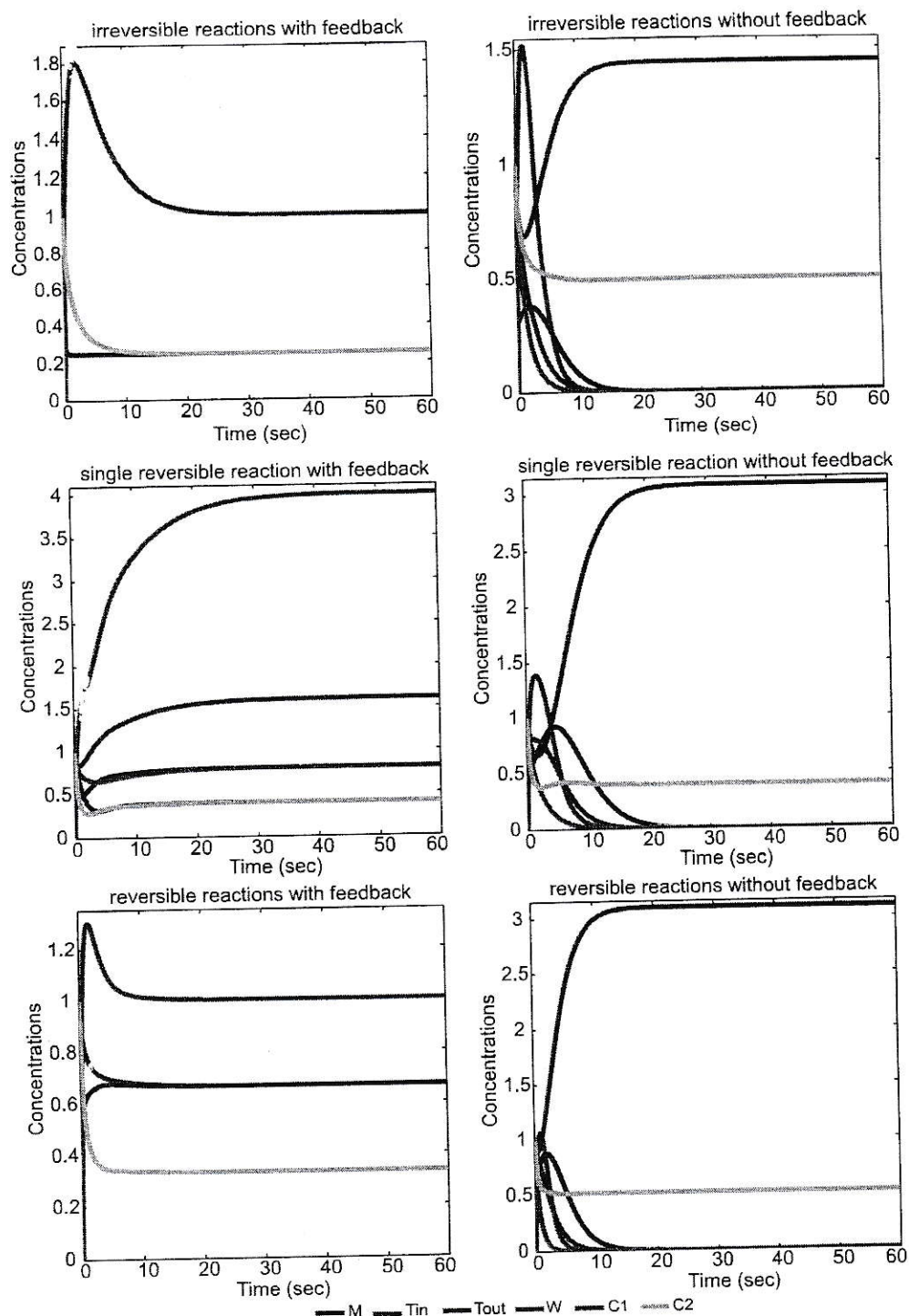
networks found in origin of life and protocell research [25]. This model is a generalization of the RAF model by Contreras et al [13]. It incorporates a basic metabolism built from enzymatic reactions and includes some environment variables for the transport, production and decomposition of the species. Hence the environment, even if in a relatively simple way, is considered. Furthermore we consider five variants of our generalization (case 1). Since in biological applications there are often also reversible reaction given, i.e. due to thermodynamics, the network in case 3 incorporates reversible reactions for all the reaction given in the network. However not all reactions in biological models need to be reversible [19]. Hence we also consider a model in which one of the reactions is not reversible (case 5). In order to examine the effect of the positive feedback given by the cycle we took, without loss of generality, one reaction out of the set of possible reactions from the networks in cases 1, 3, 5. This constitutes the networks 2, 4, 6 respectively.

Figure 4 illustrates how our main result, Lemma 1, can be applied to analyze the particular autocatalytic network model Figure 3. This network has three reactive organizations (Figure 4B). The smallest organization does not contain any molecule and thus also no cycles. According to the main result Lemma 1 all other reactive organizations must contain cycles. First, the organization  $\{M, T_{in}\}$  emerges by adding the trivial cycle  $\{T_{in}\}$  (Figure 4B). Second, the largest organization contains a cycle with the catalysts  $\{C_1, C_2\}$  (Figure 4B). The catalytic effect is necessary for the organizations to exist. If we remove the cycle's catalytic effect, the largest organization vanishes, as can be seen in the reactive organizations of the network with the cycle (case 1, Figure 4B) compared the ones of the network without the cycle (case 2, Figure 4D). Due to the special structure of the model considered here we are able to provide a further analysis of dynamical aspects using ODEs.

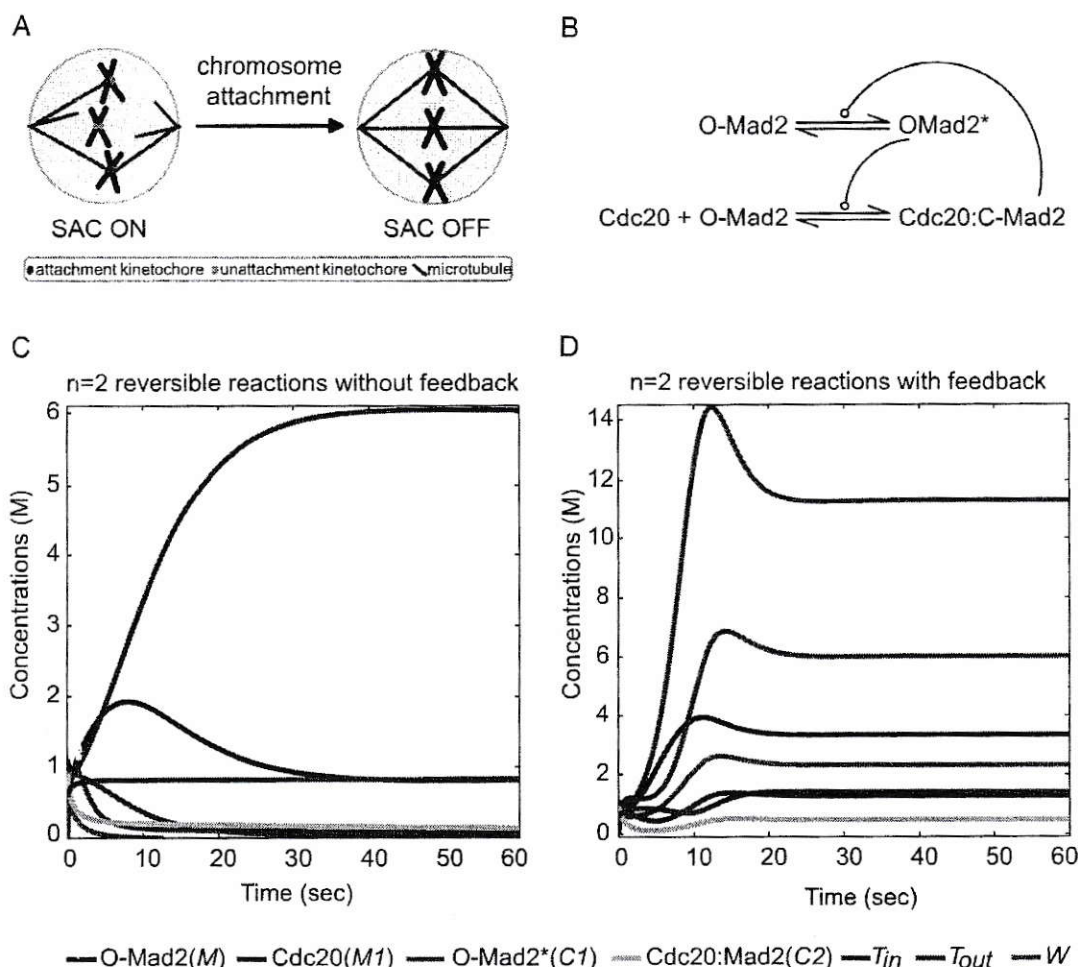




**Figure 4. Hasse diagrams of organizations for the  $n=2$  case of the particular autocatalytic network with and without reversible reactions.** A and C show the Hasse diagrams of the organizations with feedback (closed cycle) and without feedback (open cycle) respectively for the particular autocatalytic network without reversible reactions. When adding reversible reactions, we have panels E and G which show the Hasse diagrams of the organizations with feedback (closed cycle) and without feedback (open cycle) respectively. In B, D, F and H we show only reactive organizations of the panels A, C, E and G respectively (see Section "Methods"). For example  $\{M\}$  is an organization in both cases, feedback and no feedback, but  $M$  alone does not react to anything. Hence it is not a reactive organization. doi:10.1371/journal.pone.0045772.g004



**Figure 5. Simulations of all particular autocatalytic network variants ( $n=2$ ) with and without feedback.** We chose initial conditions as follows:  $M=0.3$ ,  $T_{in}=1$ ,  $T_{out}=1$ ,  $w=0$ ,  $C1=1$ ,  $C2=1$ . All forward reaction coefficients are set to 1 while reversible rates are set to 0.1. We tested the effect of different relative reversible rates and we found that they have no influence on the systems (see Appendix "Fixed Point and Stability Analysis for the Particular Autocatalytic Network") for details.  
doi:10.1371/journal.pone.0045772.g005



**Figure 5. The Template model of the mitotic spindle assembly checkpoint (SAC) mechanism.** (A) Pictorial representation of the SAC mechanism. (B) The biochemical reaction of the Template model which is the interface of SAC mechanism. We added the species  $T_{in}$ ,  $T_{out}$  and  $W$  to this core model so that we end up with the particular autocatalytic network Figure 3A, where the species  $M$ ,  $M1$ ,  $C1$  and  $C2$  refer to O-Mad2, Cdc20, O-Mad2\* and Cdc20:Mad2 respectively. (C) Numerical simulations of ordinary differential equations where all species concentrations are presented. All species reach steady state after about 30 seconds. (D) Same as in (C) after removing the positive feedback from Cdc20:C-Mad2 to activate O-Mad2 (refers to Mad2\*). doi:10.1371/journal.pone.0045772.g006

In order to define the particular autocatalytic network under consideration, we use the molecular species

$$\mathcal{M} = \{M, T_{in}, T_{out}, W, C_1, C_2, C_3, \dots, C_{n-1}, C_n\}$$

and the set of reactions

$$\mathcal{R} = \left\{ \begin{aligned} &T_{in} \xrightarrow{k_{in}} M + T_{in}, W + T_{out} \xrightarrow{k_{out}} T_{out}, M + C_1 \rightarrow T_{in} + C_1, \\ &M + C_n \rightarrow T_{out} + C_n, T_{in} + C_n \rightarrow W + C_n, T_{out} + C_1 \rightarrow W + C_1, \\ &M + C_n \rightarrow C_1 + C_n, M + C_1 \rightarrow C_2 + C_1, M + C_2 \rightarrow C_3 + C_2, \\ &\dots, M + C_{n-1} \rightarrow C_n + C_{n-1}, C_1 + C_n \rightarrow W + C_n, \\ &C_2 + C_1 \rightarrow W + C_1, C_3 + C_2 \rightarrow W + C_2, \dots, \\ &C_n + C_{n-1} \rightarrow W + C_{n-1} \end{aligned} \right\}.$$

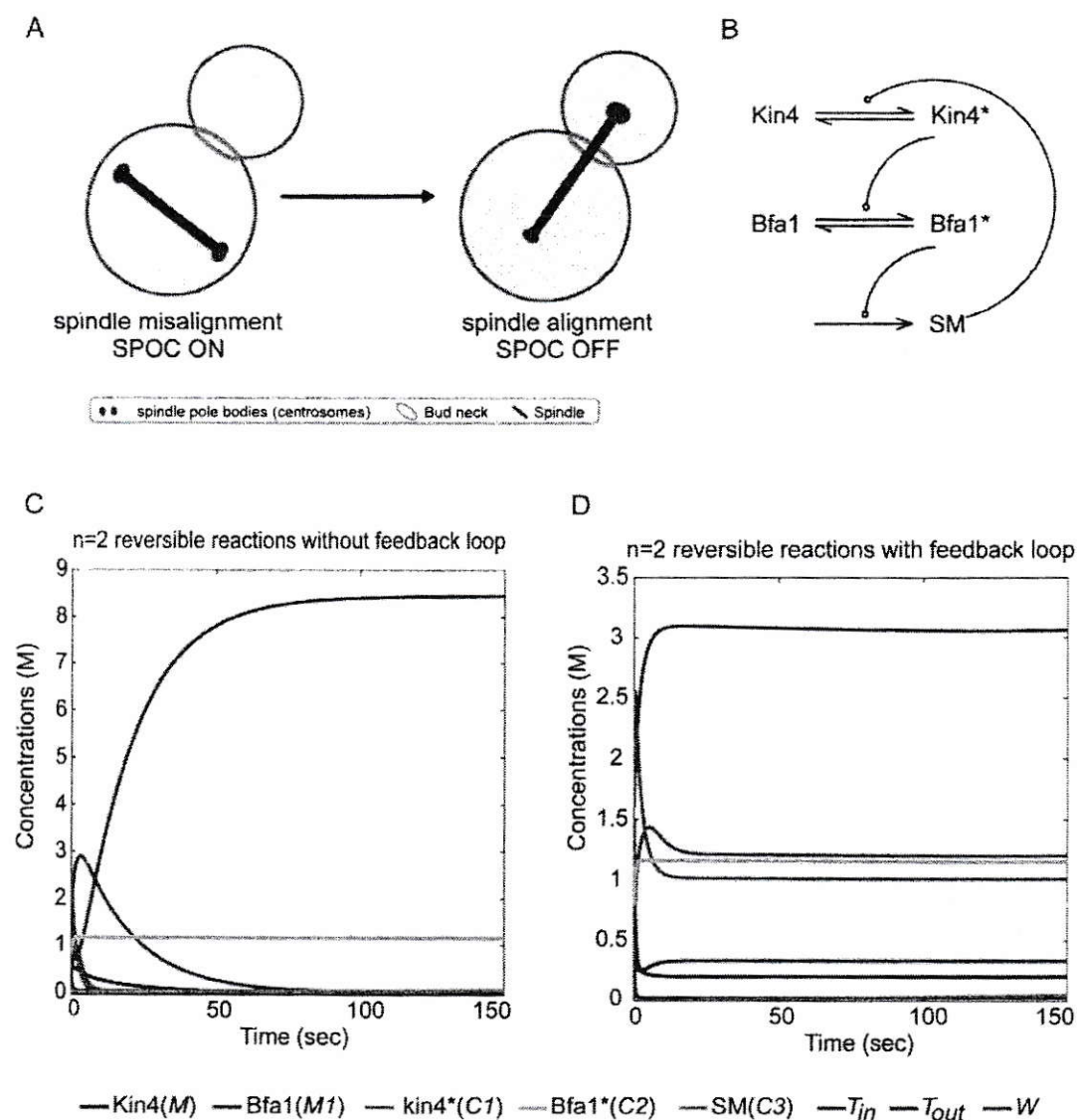
The set of reverse reactions is

$$\mathcal{R}_{rev} = \{C_1 \xrightarrow{a} M, C_2 \xrightarrow{a} M, C_3 \xrightarrow{a} M, \dots, C_n \xrightarrow{a} M\}.$$

We decided to use reaction rates equal to 1 in most of the reactions in order to have a simple and usable model. Also the main influence of reaction rates should be given by the catalysts as suggested in RAF set theory. The parameter  $a$  gives control over the strength of the reverse reactions and  $k_{in}$  and  $k_{out}$  govern the inflow and degradation, respectively.

We will focus on the following reaction networks:

1.  $(\mathcal{M}, \mathcal{R})$  – generalized cycle
2.  $(\mathcal{M}, \mathcal{R} \setminus \{M + C_n \rightarrow C_1 + C_n\})$  – generalized open cycle (the feedback is missing)
3.  $(\mathcal{M}, \mathcal{R} \cup \mathcal{R}_{rev})$  – generalized cycle with reversible reactions
4.  $(\mathcal{M}, \mathcal{R} \cup \mathcal{R}_{rev} \setminus \{M + C_n \rightarrow C_1 + C_n\})$  – generalized open cycle with reversible reactions (the feedback is missing)
5.  $(\mathcal{M}, \mathcal{R} \cup \mathcal{R}_{rev} \setminus \{C_1 \rightarrow M\})$  – generalized cycle with all reactions reversible except of one



**Figure 7. The spindle position checkpoint in budding yeast (SPOC).** (A) Pictorial representation of the SPOC mechanism. (B) A rather simplified model for SPOC mechanism. In response to mis-orientations, Bfa1 is activated through Kin4 kinase. Active Bfa1 preventing mitotic exit and enhance SM signaling. The later regulate Kin4 activations. We added the species  $T_{in}$ ,  $T_{out}$  and  $W$  to this mechanism so that we end up with the generalized particular autocatalytic network Figure 3B, where the species  $M$ ,  $M1$ ,  $C1$ ,  $C2$  and  $C3$  refer to Kin4, Bfa1, Kin4\*, Bfa1\* and SM respectively. (C) Biochemical reactions in (B) have been translated to mathematical language of ODEs and simulated without positive feedback. Almost all species are vanishing after 40 seconds. (D) same as in (C) with positive feedback presented. doi:10.1371/journal.pone.0045772.g007

6.  $(M, \mathcal{R} \cup \mathcal{R}_{rev} \setminus \{M + C_n \rightarrow C_1 + C_n, C_1 \rightarrow M\})$  – generalized open cycle with all reactions reversible (the feedback is missing) except of one

It is possible to compute the fixed points and to analyze their stability analytically in the first case. The analysis with the help of chemical organizations yields that there cannot be a positive fixed point in the cases 2, 4, 6. The other cases 3, 5 can only be treated numerically (analytical solution was not found using Mathematica [26]) but yield similar results and are more realistic for biological applications.

In summary we see that the three reaction networks 1, 3, 5 with a cycle have a stable positive fixed point whereas the networks 2, 4, 6 without the feedback via the cycle do not exhibit a positive fixed point at all. So the feedback of the cycle ensures the existence and stability of a fixed point. Numerical simulations of all particular

autocatalytic network variants are shown in Figure 5. A detailed account is provided in the Appendix.

**Mitotic Spindle Assembly Checkpoint Network Model.** The mitotic Spindle Assembly Checkpoint (SAC) ensures accurate chromosome segregation by restraining cell-cycle progression from entering anaphase until all chromosomes have made proper bipolar attachments to the mitotic spindle (Figure 6A). It is thought that unattached or misaligned kinetochores catalyze the formation of a “wait-anaphase” signal which then diffuses to counter the activation of the ubiquitin ligase APC by its coactivator Cdc20. Activation of APC by Cdc20 triggers chromosome segregation by ubiquitination of securin and cyclin B. Dysfunction of the SAC leads to aneuploidy [27,28] and its reliable function is important for tumor suppression [29,30].

**Table 1.** Number of species of the reactive organizations.

	O <sub>0</sub>	O <sub>1</sub>	O <sub>2</sub>	O <sub>3</sub>
specie	31	487	532	547

We number the reactive organizations by size in ascending order from 0 to 3. doi:10.1371/journal.pone.0045772.t001

The Cdc20-binding protein Mad2 was suggested as playing crucial and major part in the “wait-anaphase” signal, as it is stabilized in a conformation with increased affinity to Cdc20 specifically at unattached kinetochores. This module was called the “Template” model [31]. According to this model, Mad2 in its open conformation (O-Mad2) is recruited to unattached kinetochores by Mad1-bound Mad2 in its close conformation (C-Mad2) to form the ternary complex Mad1:C-Mad2:O-Mad2\*. In this complex O-Mad2\* is the “activated” Mad2, i.e. it is stabilized in a conformation which can interact with Cdc20 to form Cdc20:C-Mad2. It has been proposed [32] that there exists an autocatalytic amplification (or a cycle, according to our definition) of Cdc20:C-Mad2 formation by which in addition to the activation via kinetochore-bound Mad1:C-Mad2, O-Mad2 can likewise be activated by Cdc20:C-Mad2 (Figure 6B).

The dynamics of the full Template model both *in-silico* and *in-vitro* have been recently discoursed in detail by several mathematical models [33–35]. In this study, we first reduce the Template model to solely two reversible reactions with an autocatalytic amplification which is suggested in [31]. This was observed *in vitro* [35] and has been widely discussed in [36,37]. Subsequently, we incorporate it into the particular autocatalytic network model presented in Section “A Particular Autocatalytic Network” by adding species and reactions forming production and degradation pathways for all species (compare Figure 6B with Figure 3A). This is justified since we know that all proteins are synthesized during the cell-cycle and at least some are degraded [38]. In contrast, all models in the literature [21,33–35,39,40] did not consider the production and degradation of Mad2 and Cdc20. We used arbitrary values of the kinetic parameters and avoid using published or optimized kinetic data from Budding yeast [35] or humans [31,34] because our approach indicates that the kinetic parameters for the small cycles play no role (see last section of Appendix) and our model is a reduction of the published models where the current reactions have unknown values. We stopped the simulations after one minute which was enough to reach the steady state for the system (Figure 6C and Figure 6D). This time can be thought of as the period required for attaching a single kinetochore to the microtubule (e.g. in humans during mitosis, we have 92 kinetochores). The simulation analysis is in the same concert as our theory where the model has no positive fixed point in the absence of a positive feedback provided by a cycle (Figure 6D) while in the

**Table 2.** Number of species and reactions in the genome-scale network and the cycles.

	M	Cyc <sub>1</sub> <sup>01</sup>	Cyc <sub>0</sub> <sup>01</sup>	Cyc <sub>1</sub> <sup>12</sup>	Cyc <sub>0</sub> <sup>12</sup>	Cyc <sub>1</sub> <sup>23</sup>	Cyc <sub>0</sub> <sup>23</sup>
species	762	463	2	599	2	523	2
reactions	1939	837	2	899	2	913	2

M denotes the full network, Cyc<sub>1</sub><sup>mn</sup> the bigger of the two newly appearing cycles between O<sub>n</sub> and O<sub>m</sub>, Cyc<sub>0</sub><sup>mn</sup> the smaller one.

doi:10.1371/journal.pone.0045772.t002

**Table 3.** Critical and non-critical reactions in the large cycle compared to the ones outside the cycle.

	Cyc <sub>1</sub> <sup>01</sup>	M\Cyc <sub>1</sub> <sup>01</sup>	Cyc <sub>1</sub> <sup>12</sup>	M\Cyc <sub>1</sub> <sup>12</sup>	Cyc <sub>1</sub> <sup>23</sup>	M\Cyc <sub>1</sub> <sup>23</sup>
total reactions	837	1102	899	1040	913	1026
critical reactions	353	13	386	11	400	11
non-critical reactions	484	1089	513	1029	513	1015

doi:10.1371/journal.pone.0045772.t003

presence of the cycle the system has a stable positive fixed point (Figure 6C). Only the positive fixed points are relevant for our given biological application, since we want all the species to be maintained. Hence our simple model of SAC supports the idea that the autocatalytic amplification (or cycle) is necessary for the topology of the Template model.

**Spindle Position Checkpoint Network Model.** Spindle orientation with respect to the polarity axis is crucial during asymmetric cell division [41–45]. The budding yeast *S. cerevisiae* is a unicellular organism which undergoes asymmetric cell division and has been widely used to study polarized cell growth and asymmetric cell division [46]. If the spindle fails to align properly, a remarkable surveillance mechanism called the spindle position checkpoint (SPOC) (Figure 7A), delays cells from exiting mitosis until correct spindle orientation is achieved [46–49]. SPOC keeps the activity of the Bfa1-Bub2 GAP complex under tight control. Upon spindle misalignment, the kinase Kin4 phosphorylates Bfa1, preventing its inhibitory phosphorylation by another kinase called Cdc5. Kin4 is therefore crucial for maintaining the GAP complex active [46,50,51]. Here, we build a simple mathematical model for SPOC active state. Analogously to the SAC model, we incorporated the cycle (Figure 7B) into the structure of the particular autocatalytic network (Figure 3B). Our aim is to see the effect of the positive feedback given by the cycle when we include a production and a degradation pathway for all species. The simulations show that the closed cycle (the signal SM that targets the kinase Kin4) leads to a stable positive fixed point whereas the open cycle does not (compare Figure 7D and Figure 7C).

**Genome-Scale Network.** In an earlier study the organizations of four versions of a genome-scale metabolic reaction network model ( $\mathcal{M}, \mathcal{R}$ ) of *Escherichia coli* [22] were already computed [23]. The fourth of the scenarios consists of 762 species and 1939 reactions, where reversible reactions are counted as two single reactions. It exhibits four reactive organizations O<sub>0</sub>, O<sub>1</sub>, O<sub>2</sub> and O<sub>3</sub> of which the biggest one O<sub>3</sub> contains 547 species.

For this study we developed a Java tool (available at <http://www.biosys.uni-jena.de/Services.html>) to find cycles. Our theory predicts the existence of at least one additional or changed cycle if we look at the difference between two reactive organizations. The results are summarized in Tables 1, 2 and 3. In the following we only describe our analysis for the newly appearing cycles in the largest (O<sub>3</sub>) compared to the second largest reactive organization (O<sub>2</sub>) in detail. There is one large cycle Cyc<sub>1</sub><sup>23</sup> and a small one Cyc<sub>0</sub><sup>23</sup> consisting of only two species. In the following we only consider the large cycle since the small one has no apparent influence on the system as a whole. The large cycle consists of 523 different species and there are 913 different reactions constituting it.

In order to measure how important the cycle is for the system to be able to maintain a high amount of different species, namely for a potential fixed point to exist, we performed the following

experiment. For each reaction in the cycle we check whether the largest organization still exists in the network with the reaction switched off. The same procedure is applied to the reactions not part of the cycle.

In summary, we can say that 400 of the 913 reactions being part of the cycle were identified as crucial for the presence of the largest organization. If one of these reactions is switched off the largest organization cannot be found. In case of the reactions outside the cycle, only 11 of 1026 of reactions are crucial for the existence of the organization. This indicates that the structure of the cycle is important for the existence of the stable composition of species formed by the largest organization.

The runtime of our measurement is neglectable compared to the problem of finding the reactive organizations in such large networks [52]. After a preparatory step to construct a graph from the reaction network, we use Tarjan's algorithm for detecting the strongly connected components of a graph [53] to find our cycles. Both steps have a runtime of  $O(|\mathcal{M}| + |\mathcal{R}|)$ . In order to check whether a set of molecules  $\mathbf{O}$  remains an organization when a reaction is switched off, we need to solve a linear programming problem of size  $O(|\mathbf{O}| + |\mathcal{R}_{\mathbf{O}}| - 1)$ . This is done using Ip\_solve [54].

## Discussion

The occurrence of specific motifs and their functionality in networks have been studied largely in biological, ecological and social systems [11]. Moreover, there are many works that present results, linking reaction networks with cyclic structures and the capacity to generate stationary states [1,2,4]. These studies either focus on isolated specific motifs or on the effect of feedback loops on a system's behavior. Here however, we focus on the role of cycles in systems that undergo a change in the composition of molecular species. In Lemma 1 we showed that one possibility to achieve a qualitative transition, namely from a closed (not necessarily self-maintaining) set to an organization, is to add a cycle.

On the one hand, this transition can be interpreted as a way of creating an autopoietic system from a non-autopoietic one. A chemical reaction system in which the present molecules are not able to react such that new species occur (as stated in our definition of closure) may not be sufficient to have an autopoietic property. In origins of life research one is interested in the transition from non-living to living chemical systems. Our results suggest that the addition of a cycle is a necessary condition. Furthermore the evolution from one autopoietic system to another one then includes the additions, changes and/or deletions of cycles.

On the other hand, a change of composition of molecular species in a reaction network also gives rise to new behavior and in particular to new fixed points or a change in the stability of already existing ones. In terms of COT, altering the present molecular species means a movement between organizations. The results on cycles suggest two ways of how a system moves from one organization to another and therefore potentially between stable states. Namely, a cycle can be added or eliminated (see Results section). The results presented here (Lemma 1 and its Corollary 1) provide a framework to study how the inner and outer perturbations lead to qualitative changes in the composition of a chemical system. We emphasize, that this is not only a change of state going from stable state to another one, but a change in the molecular species present at the stable state. As shown in the example on the *E. coli* sugar metabolism model this analysis can even be employed in cases where ODEs fail to predict properties. In particular, it indicates how to identify reactions and species responsible for the stability of the dynamical behavior of the

network. Furthermore, organizations can be decomposed in groups of self-maintaining clusters, that are connected through catalysts and overproduced molecules [55] and hence give more insight into the systems' dynamical structure.

An examination of qualitative properties like fixed points and stability of cycles has not been a subject in the mentioned studies. Also COT can only disprove the existence of a fixed point but not prove it. Therefore we need to use classical methods (like the Jacobian). This task can be performed for concrete classes of networks only. We took advantage of an already existing prototypical autocatalytic network containing a cycle of length 2 studied in [13]. One possible generalization without altering the original structure too much is the cycle of length  $n$  used here. We also included certain reverse reactions which often occur. On the one hand, we chose the simplest generalization (many more are thinkable), on the other hand this variant still matches closely to biological applications. The positive feedback provided by the cycle is crucial for the existence of a positive fixed point and its stability in all cycle variants we have discussed. We emphasize that in the particular autocatalytic network discussed, a positive and not a negative feedback leads to the stability property of the fixed point. This is due to the fact that the molecular species are all degrading. Therefore there needs to be a positive feedback to replace the vanished molecules. In our study a cycle provides the structure for the network to reproduce all the molecules and not only one type by a single positive feedback.

There are many biological applications that have similar cycles to those presented in this study, e.g. MAPK (mitogen-activated protein kinase) cascades and cell-cycle control models. We have used two examples of mitotic control mechanisms, namely the Spindle Assembly (SAC) and Position (SPOC) Checkpoints. Both checkpoints have high importance in cancer research [46,56]. We took advantage of the well studied models on SAC [33–35] while for SPOC, we built a very simple model that is able to grasp the basic SPOC mechanism. Our analysis of these two models is consistent with our theory and shows that the feedback guarantees a positive stable fixed point. Additionally, the published models did neglect all productions and degradations reactions for all proteins and complexes. These reactions could be important for the effect on the model's prediction. For instance in a SAC model, neglecting these reactions led to an underestimate of the effect of the feedback loop at least with respect to the stability of the system [34].

As future work we aim at the inclusion of non-stoichiometric information, such as the kinetic rate relations or relative molecular concentrations. These might allow us to establish concrete results about the existence and stability of fixed points, and of other stationary regimes of higher dimension, such as periodic orbits and limit cycles [17]. Indeed, a novel molecular species (or a set of them) introduced to a reaction network can reinforce or breakdown the stability of a chemical system.

## Acknowledgments

We acknowledge Florian Centler and Fouzia Ahmad for critical reading the manuscript; Gerd Grünert and Dennis Görlich for technical assistance.

## Author Contributions

Conceived and designed the experiments: PK BR GE TV BI PD. Performed the experiments: PK GE BI. Analyzed the data: PK BR GE BI PD. Contributed reagents/materials/analysis tools: PK GE BI. Wrote the paper: PK BR GE TV BI PD.

## Appendix

### Chemical Organization Theory

Let  $\mathcal{M}$  be a set and  $\mathcal{R}$  be a subset of  $\mathcal{P}_{mult}(\mathcal{M}) \times \mathcal{P}_{mult}(\mathcal{M})$  where  $\mathcal{P}_{mult}(\mathcal{M})$  denotes the set of multisets over  $\mathcal{M}$ . The pair  $(\mathcal{M}, \mathcal{R})$  is called *reaction network* and we call  $\mathcal{M}$  the set of *molecules* (or *species*) and  $\mathcal{R}$  the set of *reactions*.

Like in chemistry, for  $(l, r) \in \mathcal{R}$  we also write  $l \rightarrow r$  or

$$\sum_{M \in \mathcal{M}} l_M M \rightarrow \sum_{M \in \mathcal{M}} r_M M$$

where we denote by  $l_M, r_M \in \mathbb{N}_0$  the multiplicity of  $M$  in  $l, r$  respectively. A reaction  $(l, r) \in \mathcal{R}$  occurs when the molecules in  $l$  are consumed to produce the molecules in  $r$  (what this actually means has to be made precise for every dynamics separately, e.g. the application of a reaction in stochastic dynamics is different from continuous dynamics).

Furthermore we define the *support* and the *product* of  $(l, r)$  by

$$\text{supp}(l, r) := \{M \in \mathcal{M} \mid l_M > 0\},$$

$$\text{prod}(l, r) := \{M \in \mathcal{M} \mid r_M > 0\}.$$

Let  $\mathbf{A}$  be a subset of  $\mathcal{M}$ . We define  $\mathcal{R}_{\mathbf{A}}$ , the set of reactions applicable to  $\mathbf{A}$ , by setting

$$\mathcal{R}_{\mathbf{A}} := \{(l, r) \in \mathcal{R} \mid \text{supp}(l, r) \subseteq \mathbf{A}\}.$$

Abusing notation we use a reaction  $(l, r) \in \mathcal{R}$  as an index as well and define the *stoichiometric matrix*  $N_{\mathbf{A}} \in \mathbb{R}^{|\mathbf{A}| \times |\mathcal{R}_{\mathbf{A}}|}$  for  $\mathbf{A}$  by

$$(N_{\mathbf{A}})_{A, (l, r)} = r_A - l_A, \quad A \in \mathbf{A}, (l, r) \in \mathcal{R}_{\mathbf{A}}.$$

$\mathbf{C}$  being closed means that by applying reactions from  $\mathcal{R}_{\mathbf{C}}$  we do not get molecules outside  $\mathbf{C}$ . Formally speaking, a subset  $\mathbf{C}$  of  $\mathcal{M}$  is *closed* if for all reactions  $(l, r) \in \mathcal{R}_{\mathbf{C}}$  we have  $\text{prod}(l, r) \subseteq \mathbf{C}$ , i.e. if  $(\mathbf{C}, \mathcal{R}_{\mathbf{C}})$  is a reaction network [14].

$\mathbf{S}$  being semi-self-maintaining means that if a reaction application consumes a species, there is a reaction producing this species. More formally, a subset  $\mathbf{S}$  of  $\mathcal{M}$  is *semi-self-maintaining* if for every  $A \in \mathbf{S}$  and  $(l, r) \in \mathcal{R}_{\mathbf{S}}$  with  $l_A - r_A > 0$  there is a  $(l', r') \in \mathcal{R}_{\mathbf{S}}$  with  $r'_A - l'_A > 0$ .

A subset of  $\mathcal{M}$  is a *semi-organization* if it is closed and semi-self-maintaining.

Semi-organizations entail a topological (non-stoichiometric) notion of stability in reaction networks. On the one hand, the closure property ensures that a semi-organization will not produce novel species. On the other hand the semi-self-maintenance ensures that the species that are consumed within the network are also produced. However, although semi-organization is a necessary condition for the occurrence of a fixed point, semi-organizations cannot be maintained over time in general, i.e. consider the reaction network  $(\{A, B\}, \{A \rightarrow B, 2B \rightarrow A\})$  with the semi-organization  $\{A, B\}$ .

In order to capture more of the notion of fixed states in a reaction network, we need to consider the stoichiometry of the reactions, and verify if the network can maintain itself, i.e. produce all the species at a higher or equal rate to what they are consumed.

A subset  $\mathbf{S}$  of  $\mathcal{M}$  is *self-maintaining* if there is a vector  $v \in \mathbb{R}^{|\mathcal{R}_{\mathbf{S}}|}$  with strictly positive entries such that  $N_{\mathbf{S}} v \in \mathbb{R}^{|\mathbf{S}|}$  has only non-negative entries [16].  $\mathbf{S}$  being self-maintaining means that all the reactions from  $\mathcal{R}_{\mathbf{S}}$  can occur at a certain strictly positive rate without decreasing the concentration of any species of  $\mathbf{S}$ .

A subset of  $\mathcal{M}$  is a *chemical organization* if it is closed and self-maintaining [14,16].

The term organization refers to a more stringent condition than semi-organization. Indeed, every organization is a semi-organization [16]. Moreover, organizations bridge the reaction network with its dynamical behavior. Given a reaction network and a kinetic law (e.g. mass action kinetics), it has been proven that the fixed points of the obtained ordinary differential equation (ODE) system, relate to the set of organizations as follows. For every fixed point the set of molecules with concentration higher than zero in that fixed point is an organization of the reaction network [16,17].

### Fixed Point and Stability Analysis for the Particular Autocatalytic Network

The stoichiometric matrix  $N_{\mathcal{M}}$  of the system  $(\mathcal{M}, \mathcal{R})$  with the generalized cycle (case 1) defined in Section “A Particular Autocatalytic Network” is

$$N_{\mathcal{M}} = \begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 0 & -1 & -1 & -1 & \dots & -1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & \dots & 0 & 1 & 1 & 1 & \dots & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 & -1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & -1 \end{pmatrix}$$

It is possible to guess a suitable vector to show the self-maintenance property of the whole system. For  $v = (n+2, 1, 1, 1, 1, \dots, 1)$  we get  $N_{\mathcal{M}} v = (0, 0, 0, n+1, 0, 0, 0, \dots, 0)$ . Hence the reaction network is an organization.

In a similar fashion we see that the full network is not an organization in the cases 2, 4, 6 since the row 7 is missing and species  $C_1$  can no longer be maintained. In case of the cycle with reversible reactions (cases 3, 5) we also find the network to be an organization with the help of a similar argument. More precisely, the stoichiometric matrix needs to be enlarged by the reverse reactions and the flux vector slightly adjusted (see Figure 4 for typical Hasse diagrams of these cases).

For the determination of fixed points and their stability we employ the ODEs given by mass action kinetics for the reaction networks. The reaction rate constants we did not explicitly define are assumed to be all equal to 1.

The ODEs can be written down for all six cases at once using the parameter  $s$  which is equal to 1 if the cycle is closed (cases 1, 3, 5) and equal to 0 if open (cases 2, 4, 6). We also use the constant  $a$  for the rate of the reverse reactions.

1.  $s = 1, a_i = 0$  – generalized cycle
2.  $s = 0, a_i = 0$  – generalized open cycle (the feedback is missing)

3.  $s = 1, a_i = a$  – generalized cycle with reversible reactions
4.  $s = 0, a_i = a$  – generalized open cycle with reversible reactions (the feedback is missing)
5.  $s = 1, a_1 = 0, a_i = a$  – generalized cycle with all reactions reversible except of one
6.  $s = 0, a_1 = 0, a_i = a$  – generalized open cycle with all reactions reversible (the feedback is missing) except of one

The last section already showed that there is no fixed point in the cases 2, 4, 6. Only case 1 can be treated analytically. For the cases 3, 5 numerical simulations can be done.

The ODEs for our reaction network are given by

$$\dot{M} = k_{in}T_{in} - M \left( 2C_1 + \sum_{i=2}^{n-1} C_i + (s+1) \cdot C_n \right) + \sum_{i=1}^n a_i C_i$$

$$\dot{T}_{in} = MC_1 - T_{in}C_n$$

$$\dot{T}_{out} = MC_n - T_{out}C_1$$

$$\dot{W} = -k_{out}WT_{out} + T_{in}C_n + T_{out}C_1 + C_1C_n + \sum_{i=1}^{n-1} C_iC_{i+1}$$

$$\dot{C}_1 = (s \cdot M - C_1)C_n - a_1C_1$$

$$\dot{C}_2 = (M - C_2)C_1 - a_2C_2$$

$$\dot{C}_3 = (M - C_3)C_2 - a_3C_3$$

⋮

$$\dot{C}_{n-1} = (M - C_{n-1})C_{n-2} - a_{n-1}C_{n-1}$$

$$\dot{C}_n = (M - C_n)C_{n-1} - a_nC_n.$$

Its Jacobian matrix is

$$J = \begin{pmatrix} J_{1,1} & k_{in} & 0 & 0 & -2M & -M & -M & \dots & -M & -2M \\ C_1 & -C_n & 0 & 0 & M & 0 & 0 & \dots & 0 & -T_{in} \\ C_n & 0 & -C_1 & 0 & -T_{out} & 0 & 0 & \dots & 0 & M \\ 0 & C_n & J_{4,3} & -k_{out} & J_{4,4} & C_3 & C_4 & \dots & C_n & J_{4,n} \\ C_n & 0 & 0 & 0 & -C_n & 0 & 0 & \dots & 0 & M - C_1 \\ C_1 & 0 & 0 & 0 & M - C_2 & -C_1 & 0 & \dots & 0 & 0 \\ C_2 & 0 & 0 & 0 & 0 & M - C_3 & -C_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{n-2} & 0 & 0 & 0 & 0 & 0 & 0 & \dots & -C_{n-2} & 0 \\ C_{n-1} & 0 & 0 & 0 & 0 & 0 & 0 & \dots & M - C_n & -C_{n-1} \end{pmatrix},$$

where  $J_{1,1} = -(2C_1 + \sum_{i=2}^{n-1} C_i + 2C_n)$ ,  $J_{4,3} = -k_{out}W + C_1T_{out}$ ,  $J_{4,n} = T_{out} + C_n + C_2$ , and  $J_{4,n} = T_{in} + C_1 + C_{n-1}$ .

**References**

1. Thomas R, Kaufman M (2001) Multistationarity, the basis of cell differentiation and memory. I. structural conditions of multistationarity and other nontrivial behavior. *Chaos* 11: 170–179.
2. Soulé C (2006) Mathematical approaches to differentiation and gene regulation. *Comptes Rendus Biologies* 329: 13–20.
3. Xiong W, Ferrell J (2003) A positive-feedback-based bistable ‘memory module’ that governs a cell fate decision. *Nature* 426: 460–465.

For case 1 we look for positive steady states and set the system to 0. We immediately get the following relations:

$$M = \frac{k_{in}}{(n+2)}$$

$$C_1 = \dots = C_n = M$$

$$T_{in} = M$$

$$T_{out} = M$$

$$W = \frac{k_{in}}{k_{out}}$$

To determine the stability we use the Jacobian matrix which is in the fixed point given by

$$\frac{1}{(n+2)} k_{in} \begin{pmatrix} -(n+2) & n+2 & 0 & 0 & -2 & -1 & -1 & \dots & -1 & -2 \\ 1 & -1 & 0 & 0 & 1 & 0 & 0 & \dots & 0 & -1 \\ 1 & 0 & -1 & 0 & -1 & 0 & 0 & \dots & 0 & 1 \\ 0 & 1 & -(n+1) & -k_{out} & 3 & 1 & 1 & \dots & 1 & 3 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & -1 \end{pmatrix}.$$

After tedious computations we can find that the eigenvalues of the Jacobian are as follows

$$k_{in} \left( -1, -\frac{1}{n+2}, \dots, -\frac{1}{n+2}, -\frac{k_{out}}{n+2} \right).$$

This means that the fixed point we found is stable. For the remaining cases 3, 5 we employ a parameter scan and numerically (with Mathematica [26]) compute fixed points and their stability. For the case  $n = 2$  we used the following pairs of parameter values

$$k_{in} = 10\{-1, -0.4, 0.2, 0.8, 1.4, 2\},$$

$$k_{out} = 10\{-1, -0.4, 0.2, 0.8, 1.4, 2\},$$

$$a = \{-1, -0.8, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6, 0.8, 1\}.$$

We could verify that in these cases the results were as expected. Apparently the effects of the parameters are neglectable. All numerical simulations of cycle variants as ODEs ( $n = 2$ ) show the same expected results (Figure 5).



7. Piedrafitra G, Montero F, Morán F, Cárdenas M, Cornish-Bowden A (2010) A simple selfmaintaining metabolic system: Robustness, autocatalysis, bistability. *PLoS Computational Biology* 6: e1000872.
8. Li Y (2011) A generic model for open signaling cascades with forward activation. *J Math Biol* 65: 709–742.
9. Meléndez-Hevia E, Waddell T, Cascante M (1996) The puzzle of the krebs citric acid cycle: Assembling the pieces of chemically feasible reactions, and opportunism in the design of metabolic pathways during evolution. *Journal of Molecular Evolution* 43: 293–303.
10. Calvin M, Massini P (1952) The path of carbon in photosynthesis. *Cellular and Molecular Life Sciences* 8: 445–457.
11. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, et al. (2002) Network motifs: Simple building blocks of complex networks. *Science* 298: 824–827.
12. Knabe JF, Nehaniv CL, Schilstra MJ (2008) Do motifs react evolved function? – No convergent evolution of genetic regulatory network subgraph topologies. *Biosystems* 94: 68–74.
13. Contreras D, Pereira U, Hernández V, Reynaert B, Letelier C (2011) A loop conjecture for metabolic closure. In: Lenaerts T, Giacobini M, Bersini H, Bourguin P, Dorigo M, et al, editors, *Advances in Artificial Life, ECAL 2011: Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems*. MIT Press, pp. 176–183.
14. Fontana W, Buss L (1994) “The arrival of the fittest”: Toward a theory of biological organization. *Bulletin of Mathematical Biology* 56: 1–64.
15. di Fenizio P, Dittrich P, Banzhaf W, Ziegler J (2012) Towards a theory of organizations. In: Lange H, editor, *German Workshop on Artificial Life (GVAL 2007)*, Bayreuth (In press).
16. Dittrich P, di Fenizio P (2007) Chemical organisation theory. *Bulletin of Mathematical Biology* 69: 1199–1231.
17. Peter S, Dittrich P (2011) On the relation between organizations and limit sets in chemical reaction systems. *Advances in Complex Systems* 14: 77–96.
18. Hordijk W, Steel M (2004) Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology* 227: 451–461.
19. Le Novère N, Bornstein B, Broicher A, Courtot M, Donizelli M, et al. (2006) *BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems*. *Nucleic Acids Research* 34: D689–D691.
20. Kaleta C, Richter S, Dittrich P (2009) Using chemical organization theory for model-checking. *Bioinformatics* 25: 1915–1922.
21. Ibrahim B, Diekmann S, Schmitt E, Dittrich P (2008) In-silico modeling of the mitotic spindle assembly checkpoint. *PLoS ONE* 3: e1555.
22. Reed J, Vo T, Schilling C, Palsson B, et al. (2003) An expanded genome-scale model of *Escherichia coli* k-12 (jir904 gsm/gpr). *Genome Biol* 4: R54.
23. Centler F, Kaleta C, Speroni di Fenizio P, Dittrich P (2008) Computing chemical organizations in biological networks. *Bioinformatics* 24: 1611–1618.
24. Di Fenizio P, Dittrich P (2002) Artificial chemistry's global dynamics. movement in the lattice of organisation. *The Journal of Three Dimensional Images* 16: 160–163.
25. Rasmussen S, Bedau M, Chen L, Deamer D, Krakauer D, et al (2008) *Protocells – Bridging Nonliving and Living Matter*. MIT Press.
26. Wolfram Research Inc (2008) *Mathematica* version 7.0. Wolfram Research Inc.
27. Cimini D, Degraffi F (2005) Aneuploidy: a matter of bad connections. *Trends in Cell Biology* 15: 442–451.
28. Suijkerbuijk SJ, Kops GJ (2008) Preventing aneuploidy: The contribution of mitotic checkpoint proteins. *Biochim Biophys Acta* 1786: 24–31.
29. Holland AJ, Cleveland DW (2009) Boveri revisited: Chromosomal instability, aneuploidy and tumorigenesis. *Nat Rev Mol Cell Biol* 10: 478–87.
30. Li M, Fang X, Wei Z, York JP, Zhang P (2009) Loss of spindle assembly checkpoint-mediated inhibition of cdc20 promotes tumorigenesis in mice. *J Cell Biol* 185: 983–994.
31. DeMattoni A, Pearson CG, Cimini D, Canman JC, Sala V, et al. (2005) The mad1/mad2 complex as a template for mad2 activation in the spindle assembly checkpoint. *Curr Biol* 15: 214–25.
32. Simonetta M, Manzoni R, Mosca R, Mapelli M, Massimiliano L, et al. (2009) The influence of catalysis on mad2 activation dynamics. *PLoS Biol* 7: e10.
33. Dončić A, Ben-Jacob E, Barkai N (2005) Evaluating putative mechanisms of the mitotic spindle checkpoint. *Proc Natl Acad Sci U S A* 102: 6332–7.
34. Ibrahim B, Dittrich P, Diekmann S, Schmitt E (2008) Mad2 binding is not sufficient for complete cdc20 sequestering in mitotic transition control (an in silico study). *Biophys Chem* 134: 93–100.
35. Simonetta M, Manzoni R, Mosca R, Mapelli M, Massimiliano L, et al. (2009) The influence of catalysis on mad2 activation dynamics. *PLoS Biol* 7: e10.
36. Musacchio A, Salmon ED (2007) The spindle-assembly checkpoint in space and time. *Nat Rev Mol Cell Biol* 8: 379–393.
37. Ciliberto A, Shah JV (2009) A quantitative systems view of the spindle assembly checkpoint. *EMBO J* 28: 2162–2173.
38. Prinz S, Hwang ES, Visintin R, Amon A (1998) The regulation of cdc20 proteolysis reveals a role for APC components cdc23 and cdc27 during S phase and early mitosis. *Curr Biol* 8: 750–760.
39. Ibrahim B, Schmitt E, Dittrich P, Diekmann S (2009) In silico study of kinetochore control, amplification, and inhibition effects in MCC assembly. *BioSystems* 95: 35–50.
40. Ibrahim B, Dittrich P, Diekmann S, Schmitt E (2007) Stochastic effects in a compartmental model for mitotic checkpoint regulation. *Journal of Integrative Bioinformatics* 4. doi:10.2390/bicoll-jib-2007-66.
41. Chia W, Somers WG, Wang H (2008) *Drosophila* neuroblast asymmetric divisions: Cell cycle regulators, asymmetric protein localization, and tumorigenesis. *J Cell Biol* 180: 267–72.
42. Januschke J, Gonzalez C (2008) *Drosophila* asymmetric division, polarity and cancer. *Oncogene* 27: 6994–7002.
43. Segal M, Bloom K (2001) Control of spindle polarity and orientation in *Saccharomyces cerevisiae*. *Trends Cell Biol* 11: 160–6.
44. Shapiro L, McAdams HH, Losick R (2002) Generating and exploiting polarity in bacteria. *Science* 298: 1942–6.
45. Siller KH, Doe CQ (2009) Spindle orientation during asymmetric cell division. *Nat Cell Biol* 11: 365–74.
46. Caydasi AK, Ibrahim B, Pereira G (2010) Monitoring spindle orientation: Spindle position checkpoint in charge. *Cell Div* 5: 28.
47. Hoyt MA (2000) Exit from mitosis: Spindle pole power. *Cell* 102: 267–70.
48. Pereira GA, Bestetti RB, Leite MP, Santos RB, Ramos SG, et al. (2002) Portopulmonary hypertension syndrome in schistosomiasis mansoni. *Trans R Soc Trop Med Hyg* 96: 427–8.
49. Caydasi AK, Lohel M, Gruenert G, Dittrich P, Pereira G, et al. (2012) Dynamical model of the Spindle Position Checkpoint. *Molecular Systems Biology* 8: 582.
50. Caydasi AK, Pereira G (2009) Spindle alignment regulates the dynamic association of checkpoint proteins with yeast spindle pole bodies. *Dev Cell* 16: 146–56.
51. Pereira G, Schiebel E (2005) Kin4 kinase delays mitotic exit in response to spindle alignment defects. *Mol Cell* 19: 209–21.
52. Centler F, Kaleta C, Speroni di Fenizio P, Dittrich P (2010) A parallel algorithm to compute chemical organizations in biological networks. *Bioinformatics* 26: 1788–1789.
53. Tarjan R (1972) Depth-first search and linear graph algorithms. *SIAM Journal on Computing* 1: 146–160.
54. Berkelaar M, Eikland K, Notebaert P (2004). Ip solve 5.5, open source (mixedinteger) linear programming system. URL: <http://ipsolve.sourceforge.net/5.5/>. Accessed 2012 Aug 28.
55. Veloz T, Reynaert B, Rojas D, Dittrich P (2011) A decomposition theorem in chemical organizations. In: Gheorghe M, Hinze T, Păun G, editors, *Proc. of European Conference in Artificial Life (in press)*. LNCS Springer.
56. Silva P, Barbosa J, Nascimento AV, Faria J, Reis R, et al. (2011) Monitoring the fidelity of mitotic chromosome segregation by the spindle assembly checkpoint. *Cell Prolif* 44: 391–400.