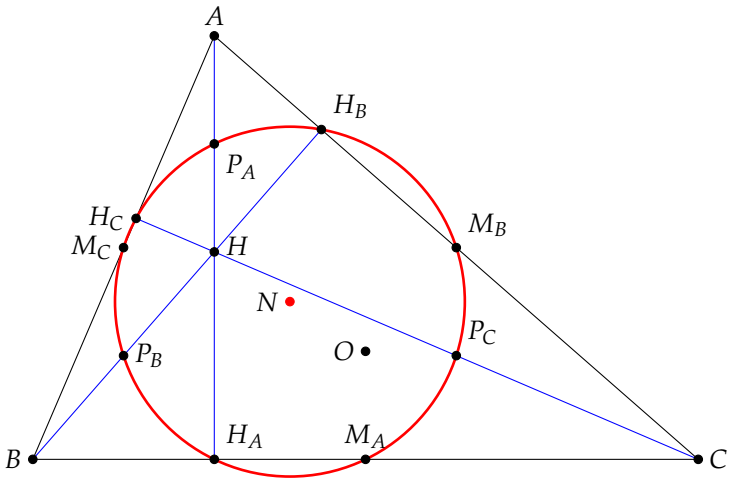
Demostración automática de teoremas geométricos



Martín Padrón Rodríguez
Nerea Civeira Corral

Trabajo realizado en el bachillerato **STEMBach**
Bienio 2020/2022
Bajo la dirección conjunta de:

Dr. Manuel Ladra González
Departamento de Matemáticas
USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Dr. Carlos Ferreiro García
IES Blanco Amor de Ourense

Ourense, enero de 2022

Resumen

En este trabajo consideraremos una aplicación reciente de conceptos y técnicas de la geometría algebraica en áreas de la informática. Desarrollaremos un enfoque sistemático que utiliza variedades algebraicas y teoría de ideales. Aplicaremos los algoritmos desarrollados en las primeras secciones basados en el método de las bases de Gröbner de Buchberger al estudio de la demostración automática de teoremas geométricos, un área que ha sido de interés para los investigadores en inteligencia artificial. Cuando las hipótesis de un teorema geométrico pueden expresarse como ecuaciones polinómicas que relacionan las coordenadas cartesianas de puntos en el plano euclidiano, las proposiciones geométricas deducibles de las hipótesis incluirán todos los enunciados que puedan expresarse como polinomios en el ideal generado por las hipótesis. Explicamos cómo se pueden demostrar teoremas geométricos interesantes de forma automática y eficaz utilizando las bases de Gröbner. Para ilustrar las ideas subyacentes se ofrecen varios ejemplos, entre ellos los teoremas de Pappus, Steiner-Fermat y Apolonio.

Como aplicación de las bases de Gröbner, proponemos un modelo booleano para la regulación de la expresión génica del operón lactosa. Hemos utilizado el sistema de álgebra SageMath para realizar nuestros cálculos.

Abstract

In this work we will consider a recent application of concepts and techniques from algebraic geometry in areas of computer science. We will develop a systematic approach that uses algebraic varieties and ideal theory. We will apply the algorithms developed in the first sections based on Buchberger's Gröbner bases method to the study of automatic geometric theorem proving, an area that has been of interest to researchers in artificial intelligence. When the hypotheses of a geometric theorem can be expressed as polynomial equations relating the cartesian coordinates of points in the Euclidean plane, the geometrical propositions deducible from the hypotheses will include all the statements that can be expressed as polynomials in the ideal generated by the hypotheses. We explain how interesting geometric theorems may be proved automatically and effectively by using Gröbner bases. Several examples including theorems named after Pappus, Steiner-Fermat and Apollonius are provided to illustrate the underlying ideas.

As an application of Gröbner bases, we propose a Boolean model for the regulation of gene expression of the lactose operon. We have used the SageMath algebra system to do our computations.

Índice

1 Introducción

La geometría fue una de las primeras áreas de las matemáticas en ser desarrollada, desde la época de los griegos. Aunque el interés por esta fue desapareciendo a lo largo de los años, resurgió recientemente en el contexto del álgebra conmutativa algorítmica. Se desarrollaron una serie de métodos para dar respuesta de forma automática a si una serie de hipótesis geométricas implican un conjunto de conclusiones o tesis.

Esta teoría comienza con los trabajos del matemático polaco Alfred Tarski [9] en la década de los 30, pero no fue propiamente desarrollada hasta 1978, año en que el matemático chino Wen-Tsün Wu [10] introdujo su método, que fue continuado por Shang-Ching Chou [1]. En la misma década, también se desarrolló otra aproximación algorítmica distinta que utiliza las bases de Gröbner, desarrollada en profundidad por Deepak Kapur [4].

En este trabajo, comenzamos estudiando el marco teórico necesario para entender cómo se realizan estas pruebas automáticas. Este incluye las bases de Gröbner y el algoritmo de Buchberger para la construcción efectiva de estas. Visto esto, nos centramos en estudiar varios problemas de la geometría, tanto desde un punto de vista clásico como de uno algebraico. Entre estos, se incluyen el teorema de Pappus, el teorema de la circunferencia de los 9 puntos y el teorema del punto de Fermat. Por último, estudiamos una aplicación de las técnicas algebraicas anteriormente mencionadas a la biología molecular.

2 Antecedentes teóricos

2.1 Polinomios en varias variables

Un **monomio** en $x_1, x_2, \ldots, x_n$ es un producto de la forma $x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ donde los exponentes $\alpha_1, \alpha_2, \ldots, \alpha_n$ son enteros no negativos. El **grado total** del monomio es la suma $\alpha_1 + \alpha_2 + \cdots + \alpha_n$ de los exponentes. Un **polinomio** en $x_1, x_2, \ldots, x_n$ sobre el cuerpo K es una combinación lineal finita (con coeficientes en K) de monomios. El grado total de un polinomio f es el máximo de los grados de los términos de f que tienen coeficientes no nulos. Por ejemplo, $3x_1 + 4x_1x_2^3x_3^5 - 17x_3x_4$ es un polinomio sobre $\mathbb{R}$, cuyo grado total es 9.

El conjunto de todos los polinomios en $x_1, x_2, \ldots, x_n$ sobre el cuerpo K se denota por $K[x_1, x_2, \ldots, x_n]$, y tiene estructura de anillo conmutativo bajo la suma y la multiplicación.

Simplificaremos la notación para monomios como sigue: sea $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ una n-tupla de enteros no negativos. A continuación, establecemos

$$x^\alpha = x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

Si $\alpha = (0, \ldots, 0)$, observe que $x^\alpha = 1$. También denotamos $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$ el grado total del monomio x^α. Con esta notación, un polinomio f podemos escribirlo de la forma

$$f = \sum_\alpha a_\alpha x^\alpha, \ a_\alpha \in K$$

2.2 Ideales y variedades afines

Definición 1. Un **ideal** I del anillo de polinomios $K[x_1, x_2, \ldots, x_n]$ es un subconjunto del anillo que satisface:

(i) $0 \in I$.

(ii) Si $f, g \in I$, entonces $f + g \in I$.

(iii) Si $f \in I$ y $h \in K[x_1, x_2, \ldots, x_n]$, entonces $hf \in I$.

Ejemplo 1. Suponga que $K = \mathbb{R}$, y sea $I = \{f \in K[x_1, x_2] \mid f(2, -2) = 0\}$, es decir, el ideal I formado por los polinomios en el anillo que tienen a $x_1 = 2$ y $x_2 = -2$ como ceros; $x_1^3 - x_1 x_2^2$ y $4x_1 - 4x_1 x_2 + 3x_2^3$ son ejemplos de elementos de I. Es fácil comprobar que I es un ideal.

Definición 2. Sean $f_1, \ldots, f_s$ polinomios en $K[x_1, x_2, \ldots, x_n]$. El **ideal (finitamente) generado** por $\{f_1, \ldots, f_s\}$ se denota por $\langle f_1, \ldots, f_s \rangle$ y se define como

$$\langle f_1, \ldots, f_s \rangle = \left\{ \sum_{i=1}^{s} h_i f_i \mid h_i \in K[x_1, x_2, \ldots, x_n] \right\}$$

Se comprueba fácilmente que, en efecto, $\langle f_1, \ldots, f_s \rangle$ es un ideal de $K[x_1, x_2, \ldots, x_n]$.

El ideal $\langle f_1, \ldots, f_s \rangle$ tiene una útil interpretación en términos de ecuaciones polinómicas. Podemos construir el sistema de ecuaciones:

$$\left.\begin{array}{c} f_1 = 0 \\ \vdots \\ f_n = 0 \end{array}\right\}$$

A partir de esas ecuaciones, podemos derivar otras mediante operaciones algebraicas. Por ejemplo, $3f_1 + 2x_3^5 f_2 = 0$, lo cual es una consecuencia del sistema original. Nótese que $3f_1 + 2x_3^5 f_2$ es un miembro del ideal $\langle f_1, \ldots, f_s \rangle$. Así, podemos pensar en $\langle f_1, \ldots, f_s \rangle$ como todas las "consecuencias polinomiales" de las ecuaciones $f_1 = f_2 = \cdots = f_s = 0$.

Si $I = \langle f_1, \ldots, f_s \rangle$, entonces decimos que $f_1, \ldots, f_s$ es una *base* para I. El término "base" no tiene aquí las mismas implicaciones que en álgebra lineal. Los polinomios de una base no tienen por qué ser linealmente independientes, y aunque una base constituye un conjunto generador para el ideal, no tiene porque ser un conjunto generador mínimo.

Ejemplo 2. Considere $K[x, y]$ y el ideal $I = \langle f_1, f_2 \rangle$, con $f_1 = x^2 + 2xy^2$ y $f_2 = xy + 2y^3 - 1$. Observe que $x = yf_1 - xf_2$, así que $x \in I$. Entonces $\{x, f_1, f_2\}$ es también una base de I .

A continuación, definiremos los objetos geométricos básicos estudiados en este proyecto.

Definición 3. Para un entero positivo n definimos el **espacio afín**

$$K^n = \{(a_1, \ldots, a_n) \in K^n \mid a_i \in K,\ i = 1, \ldots, n\}$$

Definición 4. Dado un conjunto $\{f_1,\ldots,f_s\} \subseteq K[x_1,x_2,\ldots,x_n]$, la **variedad afín** definida por $f_1,\ldots,f_s$ es el conjunto

$$\mathbf{V}(f_1,\ldots,f_s) := \{(a_1,\ldots,a_n) \in K^n \mid f_i(a_1,\ldots,a_n) = 0,\ i = 1,\ldots,s\}$$

De esta forma, una variedad afín $\mathbf{V}(f_1,\ldots,f_s)$ es el conjunto de los ceros comunes de los polinomios $f_1,\ldots,f_s$. Por ejemplo, en $\mathbb{R}[x,y]$, $\mathbf{V}(x^2-y)$ es una parábola; $\mathbf{V}(x^2-y,x-y)$ es la intersección de la parábola $x^2-y=0$ y la recta $x-y=0$, es decir, los puntos $(0,0)$, $(1,1)$, etc. Un ejemplo de variedad en $\mathbb{R}^3$ es la esfera $\mathbf{V}(x^2+y^2+z^2-1)$.

2.3 Órdenes monomiales

En una variable es natural que un término como x^7 sea mayor que x^4. En más de una variable no hay una forma obvia de ordenar los monomios. En dos variables, por ejemplo, ¿cómo deberíamos comparar términos como x^3y y x^4? Esto se formaliza en la noción de orden monomial. El precio que se paga por comparar monomios en más de una variable es que hay infinitas formas naturales de hacerlo, mientras que en el caso de una variable hay una única forma.

Definición 5. Un orden monomial $>$ en $K[x_1,x_2,\ldots,x_n]$ es una relación de orden que cumple:

1. $>$ es un orden total en $K[x_1,x_2,\ldots,x_n]$
2. $x^\alpha > x^\beta \implies x^\alpha \cdot x^\gamma > x^\beta \cdot x^\gamma$
3. $>$ es un buen orden (todo subconjunto de $K[x_1,x_2,\ldots,x_n]$ tiene elemento mínimo)

Definición 6 (Orden lexicográfico)**.** Sea $\alpha = (\alpha_1,\alpha_2,\ldots,\alpha_n)$ y $\beta = (\beta_1,\beta_2,\ldots,\beta_n)$ en $\mathbb{N}^n$. Decimos $\alpha >_{lex} \beta$ si la primera componente no nula del vector $\alpha - \beta \in \mathbb{Z}^n$ es positiva. Escribimos $x^\alpha >_{lex} x^\beta$, si $\alpha >_{lex} \beta$

Definición 7. Sea $f = \sum_\alpha a_\alpha x^\alpha$ un polinomio no nulo de $K[x_1,x_2,\ldots,x_n]$ y sea $>$ un orden monomial.

1. El multigrado de f es

   $$\text{multideg}(f) = \text{máx}(\alpha \in \mathbb{N}^n \mid a_\alpha \neq 0)$$

   (el máximo es tomado respecto a >)

2. El coeficiente principal de f es

   $$\text{LC}(f) = a_{\text{multideg}(f)} \in \mathbb{K}$$

3. El monomio principal de f es

   $$\text{LM}(f) = x^{multideg(f)}$$

   (con coeficiente 1)

4. El término principal de f es

   $$\text{LT}(f) = \text{LC}(f) \cdot \text{LM}(f)$$

2.4 Algoritmo de la división en polinomios de varias variables

Fijado un orden monomial en $K[x_1, x_2, \ldots, x_n]$, podemos dividir todo polinomio g entre una s-upla ordenada de polinomios $F = (f_1, f_2, \ldots, f_s)$.

1. Encontramos el menor i tal que $\mathrm{LT}(f_i) \mid \mathrm{LT}(g)$
   - Si existe i, añadimos a a_i el polinomio $\frac{\mathrm{LT}(g)}{\mathrm{LT}(f_i)}$, y restamos a g el polinomio $\frac{\mathrm{LT}(g)}{\mathrm{LT}(f_i)} \cdot f_i$.
   - Si no existe i, pasamos el $\mathrm{LT}(g)$ al resto r.
2. Repetimos hasta que $g = 0$. Obtenemos los cocientes a_i y el resto r, y se verifica que $g = a_1 \cdot f_1 + a_2 \cdot f_2 + \cdots + a_s \cdot f_s + r$, donde $a_i, r \in K[x_1, x_2, \ldots, x_n]$ y, o bien $r = 0$ o bien r es CL de monomios, ninguno de los cuales es divisible por $\mathrm{LT}(f_i)$. Más aún, $\mathrm{multideg}(f) \geq \mathrm{multideg}(a_i f_i)$, si $a_i f_i \neq 0$.

Ejemplo 3.

$$\begin{array}{rrr|l|l}
xy^2 & & +1 & xy+1 & y+1 \\
\cline{4-5}
-xy^2 & -y & & y & -1 \\
\cline{1-2}
 & -y & +1 & & \\
 & y & +1 & & \\
\cline{2-3}
 & & 2 & &
\end{array}$$

$$xy^2 + 1 = y(xy+1) + (-1)(y+1) + 2$$

Proposición 1. *Si el resto de dividir g entre $F = (f_1, f_2, \ldots, f_s)$ es 0, entonces $g \in \langle f_1, f_2, \ldots, f_s \rangle$.*

Demostración. Por el algoritmo de la división existen polinomios $a_1, a_2, \ldots, a_s \in K[x_1, x_2, \ldots, x_n]$ con $g = \sum_{i=1}^{s} a_i \cdot f_i$, esto es, $g \in \langle f_1, f_2, \ldots, f_s \rangle$. ■

Sin embargo, el recíproco de la Prop. 1 no es cierto. Esto se debe a que aunque si se mantiene el orden de la s-upla F, los cocientes y el resto son únicos, si se cambia el orden, los cocientes y el resto pueden variar. Es decir, dar resto $r = 0$ no caracteriza a los elementos de $I = \langle f_1, f_2, \ldots, f_s \rangle$; esto motiva que busquemos un sistema de generadores de $I : g_1, g_2, \ldots, g_t$ tal que el termino principal $\mathrm{LT}(f)$ de cualquier $f \in I$ sea múltiplo de algún $\mathrm{LT}(g_i)$.

Ejemplo 4.

$$\begin{array}{rrrrr|l|l|l}
x^2y & +xy^2 & & +y^2 & & xy-1 & y^2-1 & r \\
\cline{6-8}
-x^2y & & +x & & & x+y & 1 & \\
\cline{1-3}
 & xy^2 & +x & +y^2 & & & & \\
 & -xy^2 & & & +y & & & \\
\cline{2-4}
 & & x & +y^2 & +y & & & \\
\cline{3-5}
 & & & y^2 & +y & & & x \\
 & & & -y^2 & & +1 & & \\
\cline{4-6}
 & & & & y & +1 & & \\
\cline{5-6}
 & & & & & 1 & & x+y \\
\cline{6-6}
 & & & & & 0 & & x+y+1
\end{array}$$

$$x^2y + xy^2 + y^2 = (x+y)(xy-1) + 1(y^2-1) + (x+y+1)$$

$$\begin{array}{rrrl|l|l|l}
x^2y & +xy^2 & & +y^2 & y^2-1 & xy-1 & r \\
\cline{5-7}
-x^2y & & +x & & x+1 & x & \\
\cline{1-3}
 & xy^2 & +x & +y^2 & & & \\
 & -xy^2 & +x & & & & \\
\cline{2-3}
 & & 2x & +y^2 & & & \\
\cline{3-4}
 & & & y^2 & & & 2x \\
 & & & -y^2 & +1 & & \\
\cline{4-5}
 & & & & 1 & & \\
\cline{5-5}
 & & & & 0 & & 2x+1
\end{array}$$

$$x^2y + xy^2 + y^2 = (x+y)(y^2-1) + x(xy-1) + (2x+1)$$

Ejemplo 5.

$$\begin{array}{rrl|l|l|l}
xy^2 & -x & & xy-1 & y^2-1 & r \\
\cline{4-5}
-xy^2 & & +y & y & & \\
\cline{1-3}
 & -x & +y & & & \\
\cline{2-3}
 & & 0 & & & -x+y
\end{array}$$

$$xy^2 - x = y(xy-1) + 0(y^2-1) + (-x+y) \implies xy^2 - x \in \langle y^2-1, xy-1 \rangle$$

$$\begin{array}{rrl|l|l}
xy^2 & -x & & y^2-1 & xy-1 \\
\cline{4-5}
-xy^2 & +x & & x & \\
\cline{1-3}
 & 0 & & &
\end{array}$$

$$xy^2 - x = x(y^2-1) + 0(xy-1) + 0 \implies xy^2 - x \in \langle y^2-1, xy-1 \rangle$$

Esto quiere decir que dar resto 0 no caracteriza a los elementos de I, lo que motiva que busquemos un sistema de generadores de I (fijado un orden) $f_1, f_2, \ldots, f_s$ tal que el término principal $\mathrm{LT}(f)$ de cualquier polinomio $f \in I$ sea múltiplo de algún $\mathrm{LT}(f_i)$.

Por tanto, dado un $f \in I$, cualquier dividendo intermedio está también en I. Por eso, con este sistema de generadores, dar resto 0 si caracteriza a los elementos de I.

Definición 8. Un ideal $I \subset K[x_1, x_2, \ldots, x_n]$ es un **ideal monomial** si existe un subconjunto $A \subset \mathbb{N}^n$ tal que

$$I = \langle x^\alpha \mid \alpha \in A \rangle = \left\{ \sum_{finita} h_\alpha x^\alpha \mid \alpha \in A, h_\alpha \in K[x_1, x_2, \ldots, x_n] \right\}$$

Lema 1. *Sea $I = \langle x^\alpha \mid \alpha \in A \rangle$ un ideal monomial. Entonces, un monomio x^β pertenece a I si y solo si $x^\alpha \mid x^\beta$ para algún $\alpha \in A$.*

Demostración. Se sigue de las definiciones (ver [2, Pág. 71]). ■

Dada una n-tupla $v \in \mathbb{N}^n$, definimos $v + \mathbb{N}^n = \{v + w \mid w \in \mathbb{N}^n\}$. Necesitaremos el siguiente resultado crucial, demostrado por L. E. Dickson en un artículo de Teoría de Números [3].

Lema 2. ***(Dickson)*** *Sea $S \subseteq \mathbb{N}^n$. Entonces existe un conjunto finito de n-tuplas $v_1, \ldots, v_r \in S$ tales que*

$$S \subseteq (v_1 + \mathbb{N}^n) \cup \ldots \cup (v_r + \mathbb{N}^n)$$

Demostración. Puede verse en [2]. ■

2.5 Bases de Gröbner

Una base de Gröbner es un conjunto de polinomios en varias variables que tiene propiedades algorítmicas deseables. Todo conjunto de polinomios puede transformarse en una base de Gröbner. Este proceso generaliza tres técnicas conocidas: la eliminación gaussiana para resolver sistemas de ecuaciones lineales, el algoritmo euclidiano para calcular el máximo común divisor de dos polinomios en una variable, y el algoritmo Simplex para la programación lineal. Las bases de Gröbner son poderosas herramientas para calcular con polinomios en varias variables y tienen una gran variedad de usos, incluyendo aplicaciones en robótica, estadística, teoría de control, teoría de códigos, biología molecular (como veremos más adelante), y muchos otros campos.

Definición 9. Sea I un ideal no nulo en $K[x_1, x_2, \ldots, x_n]$. Se define $LT(I)$ como el conjunto de los términos principales de I, esto es, $LT(I) = \{LT(f) \mid f \in I\}$.

Definición 10. Denotamos por $\langle LT(I) \rangle$ el ideal generado por los elementos de $LT(I)$.

Observe que si $I = \langle f_1, f_2, \ldots, f_s \rangle$, entonces $\langle LT(f_1), LT(f_2), \ldots, LT(f_s) \rangle \subset \langle LT(I) \rangle$, pero los dos ideales no son, en general, iguales, como muestra el siguiente ejemplo:

Ejemplo 6. Sea $I = \langle f_1, f_2 \rangle$ donde $f_1 = x^2 + 2xy^2$ y $f_2 = xy + 2y^3 - 1$. Usando el orden lexicográfico en $K[x, y]$, se tiene $LT(f_1) = x^2$, $LT(f_2) = xy$. Puesto que $y(x^2 + 2xy^2) - x(xy + 2y^3 - 1) = x$, $x \in I$. Pero $LT(x) = x \notin \langle LT(f_1), LT(f_2) \rangle$, pues no es divisible ni por $LT(f_1)$, ni por $LT(f_2)$.

Proposición 2. *Sea $I \subset K[x_1, x_2, \ldots, x_n]$ un ideal.*

(i) *$\langle LT(I) \rangle$ es un ideal monomial.*

(ii) *Existen $g_1, g_2, \ldots, g_t$ tales que $\langle LT(I) \rangle = \langle LT(g_1), LT(g_2), \ldots, LT(g_t) \rangle$.*

Demostración. Consecuencia del lema de Dickson 2. ■

Teorema 1. ***(Teorema de la Base de Hilbert)****. Todo ideal $I \subset K[x_1, x_2, \ldots, x_n]$ tiene un conjunto finito de generadores.*

Demostración. Puede consultarse en [2, Pág. 77]. ■

Con este resultado, no solo tenemos que todo ideal tiene una base finita, sino que la base usada en la prueba cumple una propiedad muy especial: y es que $\langle LT(I)\rangle = \langle LT(g_1), LT(g_2), \ldots, LT(g_t)\rangle$. Esto motiva la siguiente definición.

Definición 11. Dado un orden monomial fijo, un subconjunto $G = \{g_1, g_2, \ldots, g_t\}$ de un ideal I se dice que es una **base de Gröbner** de I si

$$\langle LT(I)\rangle = \langle LT(g_1), LT(g_2), \ldots, LT(g_t)\rangle$$

Esto significa que un subconjunto $G = \{g_1, g_2, \ldots, g_t\}$ de un ideal I es una base de Gröbner si y solo si el término principal de cada elemento de I es divisible por uno de los $LT(g_i)$, $1 \leq i \leq t$.

Corolario 1. *Fijado un orden monomial, todo ideal $I \subset K[x_1, x_2, \ldots, x_n]$ distinto del $\{0\}$ tiene una base de Gröbner. Además, cualquier base de Gröbner para un ideal I es base de I.*

No existe ningún requisito de minimalidad para ser una base de Gröbner. Si G es una base de Gröbner para I, entonces cualquier subconjunto finito de I que contenga a G es también una base de de Gröbner. Para remediar esta no minimalidad, decimos que G es una base de Gröbner *reducida* si:

1. Para cada $g \in G$, el coeficiente $LC(g)$ es 1
2. $\{LT(g) \mid g \in G\}$ es un conjunto generador minimal de $\langle LT(I)\rangle$
3. Para todo $g \in G$, ningún monomio no principal de g pertenece a $\langle LT(I)\rangle$

Con esta definición, tenemos el siguiente teorema: fijado un orden monomial, todo ideal I en $K[x_1, x_2, \ldots, x_n]$ tiene una única base de Gröbner reducida [2, pag. 93].

Ejemplo 7. En el ejemplo 6, $\langle f_1, f_2\rangle$ no es una base de Gröbner.

Ejemplo 8. Ahora, considere el ideal $J = \langle g_1, g_2\rangle$ donde $g_1 = x + z$ y $g_2 = y - z$. Afirmamos que $\{g_1, g_2\}$ es una base Gröbner para J usando el orden lexicográfico en $K[x, y, z]$ con $K = \mathbb{R}$. Debemos probar que el término principal de todo elemento no nulo de J es divisible por $LT(g_1) = x$ o $LT(g_2) = y$. Para ello, consideremos cualquier $f = Ag_1 + Bg_2 \in J$. Supongamos que f es no nulo y $LT(f)$ no es divisible ni por x ni por y. Entonces, por definición de orden lexicográfico, f debe ser un polinomio en z solamente. Sin embargo, si consideramos el subespacio lineal L de $\mathbb{R}^3$ determinado por las soluciones del sistema de ecuaciones

$$\left.\begin{aligned} x + z &= 0 \\ y - z &= 0 \end{aligned}\right\}$$

es fácil comprobar que $L = \{(-t, t, t) : t \in \mathbb{R}\}$. Esto es, $(x, y, z) = (-t, t, t)$ para cierto número real t. El único polinomio en $\mathbb{Z}$ que se anula en todos esos puntos es el polinomio nulo, lo cual es una contradicción. Se deduce que $\{g_1, g_2\}$ es una base (más aún, reducida) de Gröbner para J.

Hay una forma más sistemática de detectar cuándo una base es de Gröbner, pero desafortunadamente requiere más definiciones y más maquinaria. Veremos, a continuación, algunas propiedades de las bases de Gröbner.

Proposición 3. *Sea $G = \{g_1, g_2, \ldots, g_t\}$ una base de Gröbner de un ideal $I \subset K[x_1, x_2, \ldots, x_n]$ y sea $f \in K[x_1, x_2, \ldots, x_n]$. Existe un único $r \in K[x_1, x_2, \ldots, x_n]$ que cumple las siguientes condiciones:*

(i) *Ningún término de r es divisible por ningún $LT(g_1), LT(g_2), \ldots, LT(g_t)$.*

(ii) *Existe un $g \in I$ tal que $f = g + r$.*

En particular, r es el resto de la división de f por G, sin importar cómo están ordenados los polinomios al aplicar el algoritmo de la división.

A pesar de que el resto siempre es igual, los "cocientes" pueden variar si ordenamos los elementos de la base de maneras distintas.

Ejemplo 9. Sea $G = \{x + z, y - z\}$ una base de Gröbner con el orden lexicográfico. Si dividimos el polinomio xy entre $\{x + z, y - z\}$, obtenemos que $xy = y(x + z) - z(y - z) - z^2$, y si dividimos el mismo polinomio entre $\{y - z, x + z\}$, resulta que $xy = x(y - z) + z(x + z) - z^2$. Es decir, el resto permanece igual, ya que G es una base de Gröbner, pero los "cocientes" varían.

Corolario 2. *Sea $G = \{g_1, g_2, \ldots, g_t\}$ una base de Gröbner de un ideal $I \subset K[x_1, x_2, \ldots, x_n]$ y sea $f \in K[x_1, x_2, \ldots, x_n]$. Entonces $f \in I$ si y solo si el resto de dividir f entre G es 0.*

Este corolario, nos da una solución al problema de pertenecer a un ideal, siempre que tengamos una base de Gröbner de dicho ideal. Para resolver el problema completamente, tendremos que dar un algoritmo para encontrar bases de Gröbner. La base reducida de Gröbner G puede ser calculada de forma constructiva a partir de cualquier conjunto generador de I por un algoritmo que fue introducido en la tesis de Bruno Buchberger en 1965. Buchberger nombró su método en honor a su tutor, Wolfgang Gröbner. Buchberger estableció un criterio que da las condiciones necesarias y suficientes para que un conjunto de polinomios G sea una base de Gröbner. A partir de ese criterio, derivó el algoritmo.

Una propiedad a tener en cuenta es que, dada una base de Gröbner, si cambiamos el orden del anillo de polinomios, puede dejar de ser base de Gröbner.

2.6 Resolución de sistemas de ecuaciones polinomiales

En este apartado nos centraremos en el uso de las bases de Gröbner para resolver sistemas de ecuaciones polinómicas, con el objetivo de aplicarlo a los modelos de redes booleanas de los mecanismos de regulación génica. El siguiente ejemplo revela una relación entre las bases de Gröbner y la resolución de sistemas de ecuaciones.

Ejemplo 10. Consideremos el sistema de ecuaciones lineales

$$\left.\begin{array}{c} 3x + 2y + 4z = 10 \\ 4x + 3y + 5z = 4 \end{array}\right\}$$

Por el método de eliminación de Gauss, obtenemos el siguiente sistema equivalente:

$$\left.\begin{array}{c} x + 2z - 22 = 0 \\ y - z + 28 = 0 \end{array}\right\}$$

El primer sistema determina el ideal:

$$I = \langle f_1, f_2 \rangle \text{ donde } f_1 = 3x + 2y + 4z - 10 \text{ y } f_2 = 4x + 3y + 5z - 4$$

Una base de Gröbner (bajo el orden lexicográfico) para este ideal es:

$$g_1 = x + 2z - 22,\ g_2 = y - z + 28$$

La ventaja evidente de la base de base de Gröbner es que facilita la descripción del conjunto de soluciones del sistema de ecuaciones.

Ejemplo 11. Ahora se quiere resolver el siguiente sistema de ecuaciones no lineales:

$$\left.\begin{aligned} x^2 + y^2 + z^2 - 8 &= 0 \\ x^2 + y^2 + z^2 - 4y &= 0 \\ x + 3y + 2z &= 0 \end{aligned}\right\}$$

El método de Gauss es inadecuado en este caso. Sin embargo, una base de Gröbner puede facilitar la resolución del sistema. Haciendo $J = \langle f_1, f_2, f_3 \rangle$ con

$$\begin{aligned} f_1 &= x^2 + y^2 + z^2 - 8 \\ f_2 &= x^2 + y^2 + z^2 - 4y \\ f_3 &= x + 3y + 2z \end{aligned}$$

y usando el código SageMath:

```
[7]: R.<x,y,z>=QQ[ ]

[8]: R

[8]: Multivariate Polynomial Ring in x, y, z over Rational Field

[9]: J=(x^2+y^2+z^2-8,x^2+y^2+z^2-4*y,x+3*y+2*z)*R

[10]: J.groebner_basis( )

[10]: [z^2 + 24/5*z + 32/5, x + 2*z + 6, y - 2]
```

Una base de Gröbner para J es $\{g_1, g_2, g_3\}$, donde:

$$\begin{aligned} g_1 &= x + 2z + 6 \\ g_2 &= y - 2 \\ g_3 &= z^2 + \frac{24}{5}z + \frac{32}{5} \end{aligned}$$

Comenzando con $g_2 = 0$ y $g_3 = 0$, y luego sustituyendo en $g_1 = 0$, se obtiene el conjunto solución.

2.7 Ideales radicales y la correspondencia Ideal-Variedad

Sabemos que todo ideal I en $K[x_1,x_2,\ldots,x_n]$ es finitamente generado (Teorema 1), más aún, es fácil ver que si dos bases del mismo ideal tales que $\langle f_1,\ldots,f_s\rangle = \langle g_1,\ldots,g_s\rangle$, entonces $V(f_1,\ldots,f_s) = V(g_1,\ldots,g_s)$.

Definición 12. Dado un ideal $I \subseteq K[x_1,x_2,\ldots,x_n]$ la **variedad afín** de I es el conjunto

$$\mathbb{V}(I) := \{(a_1,\ldots,a_n) \in \mathbb{K}^n \mid f(a_1,\ldots,a_n) = 0, \text{ para todo } f \in I\}$$

El Teorema de la base de Hilbert nos permite relacionar las dos definiciones de variedad dadas, como se indica en la siguiente:

Proposición 4. *Si $I = \langle f_1,\ldots,f_s\rangle$, entonces $\mathbb{V}(I) = V(f_1,\ldots,f_s)$.*

Por otro lado, definimos:

Definición 13. Dada una variedad afín $V \subseteq \mathbb{K}^n$, se define el conjunto

$$\mathbb{I}(V) := \{f \in K[x_1,x_2,\ldots,x_n] \mid f(a_1,\ldots,f_n) = 0 \text{ para todo } (a_1,\ldots,a_n) \in V\}$$

Es inmediato probar que, $\mathbb{I}(V)$ es un ideal, llamado **ideal de** V.
Las dos proposiciones siguientes son consecuencias de las definiciones.

Proposición 5.

(a) Si $I,J \subseteq K[x_1,x_2,\ldots,x_n]$ son dos ideales, $I \subseteq J \implies \mathbb{V}(I) \supseteq \mathbb{V}(J)$

(b) Si $V,W \subseteq K[x_1,x_2,\ldots,x_n]$ son dos variedades afines, $V \subseteq W \implies \mathbb{I}(V) \supseteq \mathbb{I}(W)$

Proposición 6. *Dada una variedad afín V, se cumple $\mathbb{V}(\mathbb{I}(V)) = V$.*

En este punto es natural preguntarse si, recíprocamente, dado un ideal I, es válido $\mathbb{I}(\mathbb{V}(I)) = I$. Es decir, ¿$\mathbb{I}(\mathbb{V}(f_1,\ldots,f_s)) = \langle f_1,\ldots,f_s\rangle$?
Desafortunadamente, no siempre es cierto, sin embargo podemos afirmar:

Lema 3. *Sean $f_1,\ldots,f_s \in K[x_1,x_2,\ldots,x_n]$. Entonces $\langle f_1,\ldots,f_s\rangle \subseteq \mathbb{I}(\mathbb{V}(f_1,\ldots,f_s))$.*

Demostración. Si $f \in \langle f_1,\ldots,f_s\rangle$, entonces $f = \sum_{i=1}^n h_i f_i$ con $h_1,\ldots,h_s \in K[x_1,x_2,\ldots,x_n]$. Puesto que $f_1,\ldots,f_s$ se anulan en $\mathbb{V}(f_1,\ldots,f_s)$, también lo hará $\sum_{i=1}^n h_i f_i$. Así f se anula en $\mathbb{V}(f_1,\ldots,f_s)$, lo que prueba que $f \in \mathbb{I}(\mathbb{V}(f_1,\ldots,f_s))$. ■

Ejemplo 12. Se puede comprobar que $\mathbb{I}(\mathbb{V}(x^2,y^2)) = \langle x,y\rangle$, sin embargo, $x \notin \langle x^2,y^2\rangle$, por lo que

$$\langle x^2,y^2\rangle \subsetneq \langle x,y\rangle$$

Para cuerpos arbitrarios, la relación entre $\langle f_1,\ldots,f_s\rangle$ y $\mathbb{I}(\mathbb{V}(f_1,\ldots,f_s))$ es muy sutil, sin embargo, sobre cuerpos algebraicamente cerrados, como $\mathbb{C}$, hay una relación sencilla entre esos ideales, dada por el Nullstellensatz de Hilbert que veremos más adelante.

Para seguir explorando la relación entre ideales y variedades, es natural preguntarse sobre la naturaleza de los tipos de ideales que aparecen como ideales de una variedad. Es decir, ¿podemos identificar los ideales $\mathbb{I}(V)$ que consisten en todos los polinomios que se

anulan en alguna variedad V? Observemos que si $f^s \in \mathbb{I}(V)$, entonces $(f(a_1,\ldots,a_n))^s = 0$, $\forall (a_1,\ldots,a_n) \in V$, lo cual solo puede ser si $f(a_1,\ldots,a_n) = 0$, $\forall (a_1,\ldots,a_n) \in V$ y, por tanto, $f \in \mathbb{I}(V)$. Así, un ideal que consiste en todos los polinomios que se anulan en una variedad V tiene la propiedad de que si alguna potencia de un polinomio pertenece al ideal, entonces el propio polinomio debe pertenecer al ideal. Esto nos lleva a la siguiente definición.

Definición 14. Sea $I \subseteq K[x_1, x_2, \ldots, x_n]$ un ideal. El **radical** de I, denotado por $\sqrt{I}$, es el conjunto

$$\sqrt{I} := \{f \in K[x_1, x_2, \ldots, x_n] \mid f^s \in I \text{ para cierto entero } s \geq 1\}$$

Por los comentarios previos, es claro que $\sqrt{I}$ es también un ideal y se cumple $\sqrt{I} \subseteq \mathbb{I}(\mathbb{V}(I))$.

Definición 15. Diremos que I es un **ideal radical** si $\sqrt{I} = I$.

Proposición 7. *Dada una variedad afín V y un ideal I, se cumple:*

1. $\mathbb{I}(V)$ *es un ideal radical.*
2. $\sqrt{I}$ *es un ideal radical.*
3. $\mathbb{V}(\sqrt{I}) = \mathbb{V}(I)$.

Demostración. Se puede ver en [2, Lema 5. Pág. 182] ■

A continuación presentamos dos versiones de un mismo resultado que será esencial para completar la correspondencia entre ideales y variedades, cuyas demostraciones pueden verse en [2].

Teorema 2 (Nullstellensatz débil de Hilbert)**.** *Sea $\mathbb{K}$ un cuerpo algebraicamente cerrado. Dado $I \subseteq K[x_1, x_2, \ldots, x_n]$ un ideal tal que $\mathbb{V}(I) = \emptyset$, entonces $I = K[x_1, x_2, \ldots, x_n]$.*

Teorema 3 (Nullstellensatz fuerte de Hilbert)**.** *Sea $\mathbb{K}$ algebraicamente cerrado y sea $I \subseteq K[x_1, x_2, \ldots, x_n]$ un ideal. Entonces*

$$\mathbb{I}(\mathbb{V}(I)) = \sqrt{I}$$

A la vista de las definiciones y teoremas anteriores, podemos asignar a cada ideal I la variedad $\mathbb{V}(I)$ del ideal I,

$$I \longrightarrow \mathbb{V}(I)$$

y recíprocamente, a cada variedad V, el ideal $\mathbb{I}(V)$ de la variedad V,

$$V \longrightarrow \mathbb{I}(V)$$

A continuación resumimos la correspondencia entre variedades afines e ideales.

Teorema 4 (Correspondencia Ideal-Variedad)**.** *Sea $\mathbb{K}$ un cuerpo arbitrario.*

1. *Las aplicaciones*

$$\mathbb{I} : \{\textit{Variedades afines}\} \longrightarrow \{\textit{Ideales}\}$$

y

$$\mathbb{V} : \{\textit{Ideales}\} \longrightarrow \{\textit{Variedades afines}\}$$

invierten inclusiones. Más aún, se da la igualdad $\mathbb{V}(\mathbb{I}(V)) = V$*, con lo que* $\mathbb{I}$ *es siempre inyectiva.*

2. *Si* $\mathbb{K}$ *es algebraicamente cerrado, y consideramos los ideales radicales, las aplicaciones*

$$\mathbb{I} : \{\textit{Variedades afines}\} \longrightarrow \{\textit{Ideales radicales }\}$$

y

$$\mathbb{V} : \{\textit{Ideales radicales }\} \longrightarrow \{\textit{Variedades afines}\}$$

invierten inclusiones y son biyecciones inversas una de otra.

Observemos, entonces, que mientras la aplicación $\mathbb{I}$ es siempre inyectiva, es decir, si $V \neq W$, entonces $\mathbb{I}(V) \neq \mathbb{I}(W)$, la aplicación $\mathbb{V}$ no lo es. Por ejemplo, si $I = \langle x, y \rangle$ y $J = \langle x^2, y^3 \rangle$ en $\mathbb{K}[x,y]$, es claro que $\mathbb{V}(I) = \{(0,0)\} = \mathbb{V}(J)$. Este inconveniente puede ser evitado si consideramos radicales de ideales en lugar de ideales. Adicionalmente, si el cuerpo no es algebraicamente cerrado, puede pasar que para ideales como $J = \langle x^2 + x \rangle$ en $\mathbb{R}[x]$, la variedad $\mathbb{V}(J) = \emptyset$. Esto se evita trabajando sobre cuerpos algebraicamente cerrados.

3 Hipótesis de trabajo y objetivos de la investigación

La demostración automática de teoremas se ocupa de enunciados geométricos de la forma $H \Longrightarrow T$, donde H denota el conjunto de hipótesis y T el de tesis o conclusiones. El primer paso es traducir el planteamiento geométrico inicial en uno algebraico. Para ello, se introduce en el plano euclidiano un sistema de coordenadas que facilite el problema a resolver, y a continuación se expresa tanto las hipótesis como las conclusiones del teorema en forma de ecuaciones polinómicas sobre $\mathbb{R}$. Dichas ecuaciones relacionarán las coordenadas de los puntos que intervengan en el planteamiento del problema. Tendremos un cierto número de coordenadas arbitrarias o variables independientes denotadas por $u_1, \ldots u_m$, y otro conjunto de coordenadas fijas o variables dependientes $x_1, \ldots, x_n$. Las hipótesis del teorema estarán representadas por una colección de ecuaciones polinómicas en las u_i, x_j. Como observaremos en los ejemplos de este trabajo, es típico de un teorema correctamente planteado que el número de hipótesis sea igual al número de variables dependientes. Esto tiene por objeto que, fijados unos valores para las variables independientes $u_1, \ldots u_m$, solo exista un número finito de combinaciones de las x_j que satisfagan las ecuaciones. Por lo que escribiremos las hipótesis como

$$\left.\begin{array}{r} h_1(u_1, \ldots, u_m, x_1, \ldots, x_n) = 0 \\ \vdots \\ h_n(u_1, \ldots, u_m, x_1, \ldots, x_n) = 0 \end{array}\right\} \tag{1}$$

Las conclusiones del teorema se expresarán también como polinomios en las u_i, x_j. Basta con considerar el caso de una conclusión, ya que si hay más, podemos simplemente tratarlas de una en una. Por lo tanto, escribiremos la conclusión como

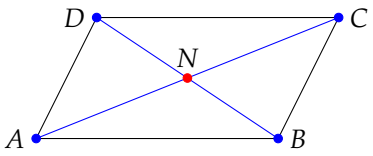
$$g(u_1,\dots,u_m,x_1,\dots,x_n) = 0$$

donde $h_1,\dots,h_n,g \in \mathbb{R}[u_1,\dots u_m,x_1,\dots,x_n]$.

Notemos que una hipótesis no puede involucrar únicamente variables independientes, puesto que una de ellas pasaría automáticamente a depender de las demás. La formulación algebraica del teorema lleva implícita una comprensión del mismo por parte del individuo que la lleva a cabo, y que le permitirá distinguir entre estos dos subconjuntos. Es este un paso difícil de automatizar. Finalmente, una manera razonable de expresar el teorema en términos algebraicos será pedir que siempre que se verifiquen las ecuaciones hipótesis, se cumpla también la ecuación conclusión. Es importante destacar que ni la elección de los puntos arbitrarios, ni la de las hipótesis y conclusiones, ni la traducción de estas, será, por lo general, única, y diferentes interpretaciones pueden llevar a distintos niveles de complejidad del problema o, incluso, a distintas "clases de verdad" del teorema (ver [2], Capítulo 6).

Visualizaremos estas ideas mediante un simple, pero representativo, ejemplo tomado de [2, Pág. 319].

Ejemplo 13. Sean A, B, C, D los vértices de un paralelogramo en el plano. Es sabido que sus dos diagonales se intersecan en el punto medio, N, de ambas.



Planteemos este teorema en términos algebraicos. El sistema de coordenadas más práctico será el cartesiano (y será el elegido en este trabajo). En primer lugar, asignaremos coordenadas a los puntos. Sin pérdida de generalidad, podemos hacer coincidir el punto A con el origen de coordenadas, así como situar el B sobre el eje de abscisas, puesto que las propiedades de un paralelogramo son invariantes por rotaciones y traslaciones. Entonces, tomamos $A = (0,0)$ y $B = (u_1,0)$, donde $u_1 \in \mathbb{R}$ representa, en valor absoluto, la longitud del segmento $\overline{AB}$, y escogemos $C = (u_2,u_3)$ arbitrariamente. El vértice restante, D, se encuentra ahora completamente determinado por A, B y C, y los cuatro vértices determinarán a su vez las coordenadas de N. Por tanto, escribimos $D = (x_1,x_2)$ y $N = (x_3,x_4)$. Notemos que, con la elección de A y B, hemos disminuido las variables potencialmente empleadas de diez a siete. Siempre intentaremos minimizar el número de variables haciendo coincidir todos los puntos posibles con los ejes. A continuación, relacionamos los puntos. Una hipótesis del teorema es que el cuadrilátero $ABCD$ es un paralelogramo, lo cual se traduce en que sus lados opuestos son paralelos, o, equivalentemente, tienen la misma pendiente:

$$\overline{AB} \parallel \overline{CD} \iff 0 = \frac{u_3 - x_2}{u_2 - x_1}$$
$$\overline{AD} \parallel \overline{BC} \iff \frac{x_2}{x_1} = \frac{u_3}{u_2 - u_1}$$

si $x_1 \neq u_2$, $x_1 \neq 0$ y $u_1 \neq u_2$. Eliminado denominadores, las ecuaciones resultan en:

$$\begin{aligned} h_1 &= u_3 - x_2 = 0 \\ h_2 &= x_2(u_2 - u_1) - u_3 x_1 = 0 \end{aligned} \tag{2}$$

Notemos que en caso de darse $x_1 = 0$, la condición $\overline{AD} \parallel \overline{BC}$ se transforma en $u_1 = u_2$, y recíprocamente. Además, si $x_1 = u_2$, $\overline{AB} \parallel \overline{CD}$ implica que $u_1 = 0$, o equivalentemente, $A = B$. Entonces, la misma hipótesis obliga a que $C = D$, con lo que se cumple $u_3 = x_2$. Así, $ABCD$ será un paralelogramo si y solo si las ecuaciones (2) se verifican.

La otra hipótesis de nuestro teorema es que las diagonales $\overline{AC}$ y $\overline{BD}$ se intersecan en N, lo cual es equivalente a afirmar que las ternas A, N, C y B, N, D son colineales. Hagamos el desarrollo para la segunda. Los tres puntos pertenecen a la misma recta vertical si y solo si $u_1 = x_1 = x_3$; a una misma recta no vertical cuando

$$\frac{x_2}{x_1 - u_1} = \frac{x_4}{x_3 - u_1}$$

igualando las pendientes de $\overline{BD} = \overline{BN}$. Los dos casos aparecen recogidos en

$$h_3 = x_2(x_3 - u_1) - x_4(x_1 - u_1) = 0 \tag{3}$$

Similarmente, la terna A, N, C es colineal si y solo si

$$h_4 = u_2 x_4 - u_3 x_3 = 0 \tag{4}$$

Finalmente, traduciremos la conclusión del teorema utilizando que N, alineado con A y C, es el punto medio entre estos puntos si y solo si $AN = NC$, donde AN es la longitud del segmento $\overline{AN}$. Empleando la fórmula de la distancia euclidiana y elevando al cuadrado, se tiene:

$$AN = NC \iff x_3^2 + x_4^2 = (u_2 - x_3)^2 + (u_3 - x_4)^2$$

Del mismo modo, para la diagonal $\overline{BD}$:

$$BN = ND \iff x_4^2 + (u_1 - x_3)^2 = (x_3 - x_1)^2 + (x_2 - x_4)^2$$

Desarrolando los cuadrados, obtenemos las conclusiones:

$$\begin{aligned} g_1 &= u_2^2 + u_3^2 - 2u_2 x_3 - 2u_3 x_4 = 0 \\ g_2 &= u_1^2 - x_1^2 - x_2^2 - 2u_1 x_3 + 2x_1 x_3 + 2x_2 x_4 = 0 \end{aligned} \tag{5}$$

Así, la traducción del teorema a ecuaciones polinómicas se enuncia de la siguiente manera: siempre que se verifiquen las ecuaciones (2), (3) y (4), deben cumplirse también las ecuaciones (5).

Nótese que en la construcción de la figura no hemos especificado que $u_1 \neq 0$ ni $u_3 \neq 0$. Estamos, por tanto, permitiendo realizaciones que no se corresponden con la idea que tenemos de un paralelogramo. Estas configuraciones intuitivamente degeneradas pueden ser estudiadas con una maquinaria que no incluiremos en este trabajo, pero, en cualquier caso, las incluiremos siempre en un primer planteamiento de un teorema.

Veamos, ahora, otra construcción alternativa, lo que muestra que hay unicidad. En primer lugar, no tendríamos por qué haber tomado como arbitrarios tres vértices del paralelogramo, dejando determinado el cuarto. Escogiendo A y B como fue explicado arriba, podríamos haber seleccionado arbitrariamente la ordenada de C y la abscisa de D, quedando determinados las coordenadas restantes. Estas coordenadas corresponden a vértices distintos. En cuanto a la traducción de las hipótesis, se puede obtener una más sencilla teniendo en cuenta la ley del paralelogramo para la suma de vectores. Si volvemos a la construcción original de la figura, comprobamos que el vector $C = (u_2, u_3)$ se obtiene sumando los vectores $B = (u_1, 0)$ y $D = (x_1, x_2)$. Así, obtendríamos las hipótesis alternativas:

$$\begin{aligned} h_1' &= u_1 + x_1 - u_2 = 0 \\ h_2' &= x_2 - u_3 = 0 \end{aligned} \tag{6}$$

Estas son más sencillas que las de (2), y permitirán una simplificación del problema que será muy apreciada cuando se trabaje con técnicas complejas. Sin embargo, las primeras mantienen una ventaja sobre las ecuaciones (6), y es que se basan en la afirmación más general de $\overline{AB} \parallel \overline{CD}$, $\overline{AD} \parallel \overline{BC}$, aplicable a muchos otros teoremas. Más adelante (Punto de Steiner-Fermat), hará patente la posibilidad de escoger entre distintos conjuntos de hipótesis y conclusiones, así como la problemática que esta elección conlleva.

La siguiente proposición enumera algunos de los enunciados geométricos más comunes que pueden ser traducidos a ecuaciones polinómicas, cuya demostración elemental omitiremos por razón de espacio.

Proposición 8. *Sean $A = (a_1, a_2)$, $B = (b_1, b_2)$, $C = (c_1, c_2)$, $D = (d_1, d_2)$, $E = (e_1, e_2)$ y $F = (f_1, f_2)$ puntos en el plano. Los siguientes enunciados geométricos pueden ser expresados mediante una o varias ecuaciones polinómicas:*

(i) *$\overline{AB}$ es paralelo a $\overline{CD}$.*

(ii) *$\overline{AB}$ es perpendicular a $\overline{CD}$.*

(iii) *A, B, C son colineales.*

(iv) *Las distancias AB y CD son iguales.*

(v) *C pertenece a la circunferencia con centro A y radio AB.*

(vi) *D pertenece a la circunferencia que pasa por A, B, C.*

(vii) *$\overline{AD}$ es tangente a la circunferencia que pasa por A, B, C.*

(viii) *C es el punto medio de $\overline{AB}$.*

(ix) *El ángulo agudo $\angle ABC$ es igual al ángulo agudo $\angle DEF$.*

(x) *$\overline{BD}$ biseca al ángulo $\angle ABC$.*

Un teorema geométrico es *admisible* cuando tanto sus hipótesis como sus conclusiones se pueden expresar en forma de ecuaciones polinómicas de la forma que lo hemos expuesto. En este trabajo, cada vez que se haga referencia a un teorema geométrico, sobrentenderemos que es admisible y que admite una formulación algebraica del tipo (1).

Una observación final es que la geometría ordinaria se interpreta sobre los números reales, pero muchos métodos de demostración automática son completos solo sobre los números complejos. Por lo tanto, tales métodos deben fallar necesariamente en teoremas válidos que no son verdaderos sobre $\mathbb{C}$. Es un hecho, que comprobamos en nuestros experimentos, que la mayoría de los teoremas de la geometría euclidiana son válidos sobre los números complejos, por lo que esta situación no se da con mucha frecuencia.

4 Materiales y métodos

La demostración automática de teoremas es la derivación de teoremas matemáticos mediante un programa informático. En este trabajo utilizaremos el sistema algebraico computacional SageMath basado en el lenguaje de programación Python. En especial, usamos la versión en línea CoCalc que permite el trabajo colaborativo en grupo.

Este apartado tratará acerca de uno de los principales métodos para probar teoremas geométricos: el de las bases de Gröbner, desarrollado por Kapur [4]. Comprobaremos que el planteamiento inicial presentado en (1) resulta no ser suficiente en muchas ocasiones, pues los enunciados de los teoremas pueden llevar implícitas hipótesis accesorias que no son traducidas. Será necesario evitar configuraciones intuitivamente degeneradas (como un triángulo colapsado a un segmento), para las que las hipótesis carecen de sentido o resultan absurdas, o restringirse a ciertas familias de configuraciones entre las no degeneradas. Estas diferencias entre los teoremas y entre sus formulaciones llevarán a definir una serie de clases de verdad para catalogar los teoremas según los tipos de configuraciones en que sean ciertos. En este trabajo estudiaremos aquellos teoremas que se verifiquen sobre todas las configuraciones de puntos posibles. Para una explicación detallada se puede consultar [6].

4.1 Teoremas universalmente ciertos

La idea natural detrás de la formulación de (1) es verificar que g se anule cada vez que lo hagan $h_1,\dots,h_n$. Mantendremos esta notación en todo este apartado. Si denotamos $V := \mathbf{V}(h_1,\dots,h_n) \subseteq \mathbb{R}^{m+n}$ (variedad hipótesis), siendo $V \neq \emptyset$, se tiene la siguiente definición.

Definición 16. Un teorema se denomina **universalmente cierto** si $g \in \mathbb{I}(V)$.

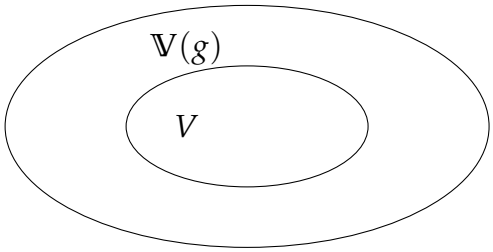
Notemos que la condición de $V \neq \emptyset$ es crucial para evitar casos triviales. En caso de que $V = \emptyset$, se tendría $\mathbb{I}(V) = \mathbb{R}[u_1,\dots,u_m,x_1,\dots,x_n] := \mathbb{R}[\mathbf{u},\mathbf{x}]$, y todo polinomio g verificaría la Definición 16. Debemos ser cuidadosos planteando hipótesis no contradictorias.

Proposición 9. *Un teorema es universalmente cierto si y solo si* $V \subseteq \mathbb{V}(g)$.

Demostración. Sea $g \in \mathbb{I}(V)$. Por definición de $\mathbb{I}(V)$,

$$\begin{aligned}&(a_1,\dots,a_m,b_1,\dots,b_n) \in V \Longrightarrow g\,(a_1,\dots,a_m,b_1,\dots,b_n) = 0\\ &\Longrightarrow (a_1,\dots,a_m,b_1,\dots,b_n) \in \mathbb{V}(g).\end{aligned}$$

Recíprocamente, $V \subseteq \mathbb{V}(g) \Longrightarrow \mathbb{I}(V) \supseteq \mathbb{I}(\mathbb{V}(g))$. Así, $g \in \mathbb{I}(\mathbb{V}(g)) \Longrightarrow g \in \mathbb{I}(V)$.



■

Desafortunadamente no se conoce un método efectivo para determinar $\mathbb{I}(V)$. No obstante, se dispone de algunas condiciones suficientes.

Proposición 10. *Si $g \in \sqrt{\langle h_1,\ldots,h_n\rangle} \subseteq \mathbb{R}[u_1,\ldots,u_m,x_1,\ldots,x_n]$, entonces el teorema $H \Longrightarrow T$ es universalmente cierto.*

Demostración. Vimos que siempre se cumple $\sqrt{I} \subseteq \mathbb{I}(\mathbb{V}(I))$, y en este caso $\sqrt{\langle h_1,\ldots,h_n\rangle} \subseteq \mathbb{I}(V)$. ■

Si $g \in \sqrt{\langle h_1,\ldots,h_n\rangle} \subseteq \mathbb{R}[\mathbf{u},\mathbf{x}]$, entonces $g^s \in \langle h_1,\ldots,h_n\rangle \subseteq \mathbb{R}[\mathbf{u},\mathbf{x}]$ para cierto $s \geq 1$. En consecuencia, existen polinomios $A_j \in \mathbb{R}[\mathbf{u},\mathbf{x}]$ tales que $g^s = \sum_{j=1}^n A_j h_j$. Esto implica que si todo h_j se anula en algunos puntos de $\mathbb{C}^{m+n}$, entonces g^s y, por tanto, también g se anulan en esos puntos, y $g \in \mathbb{I}(V_{\mathbb{C}})$. Recíprocamente, usando el Nullstellensatz fuerte, si $g \in \mathbb{I}(V_{\mathbb{C}}) = \sqrt{\langle h_1,\ldots,h_n\rangle} \subseteq \mathbb{C}[\mathbf{u},\mathbf{x}]$, entonces $g^s \in \langle h_1,\ldots,h_n\rangle$ para cierto $s \geq 1$, con lo que $g^s = \sum_{j=1}^n A_j h_j$ para $A_j \in \mathbb{C}[\mathbf{u},\mathbf{x}]$. Expresando $A_j = A'_j + iA''_j$ como suma de su parte real e imaginaria, resulta $g^s = \sum_{j=1}^n A'_j h_j + i\sum_{j=1}^n A''_j h_j$. Como $g^s, h_j \in \mathbb{R}[\mathbf{u},\mathbf{x}]$, se debe tener $g^s = \sum_{j=1}^n A'_j h_j$ y $g \in \sqrt{\langle h_1,\ldots,h_n\rangle} \subseteq \mathbb{R}[\mathbf{u},\mathbf{x}]$. Finalmente, hemos probado que

$$g \in \sqrt{\langle h_1,\ldots,h_n\rangle} \subseteq \mathbb{R}[\mathbf{u},\mathbf{x}] \Longleftrightarrow g \in \mathbb{I}(V_{\mathbb{C}}) \subseteq \mathbb{C}[\mathbf{u},\mathbf{x}].$$

Así, $g \in \sqrt{\langle h_1,\ldots,h_n\rangle}$ significa que el teorema es "universalmente cierto sobre $\mathbb{C}$".

Dados un ideal I y una variedad afín V, no es fácil calcular $\mathbb{V}(I)$ ni, especialmente, $\mathbb{I}(V)$. Así, el Nullstellensatz fuerte cobra especial importancia, pues traslada este último problema a calcular $\sqrt{I}$ en los cuerpos algebraicamente cerrados ($\mathbb{C}$). Aunque existen algoritmos para calcularlo, aquí presentaremos un criterio de pertenencia que será útil en la metodología que utilizaremos en este trabajo.

Proposición 11. *(Algoritmo de pertenencia a un radical). Sea K un cuerpo arbitrario, y sea $I = \langle f_1,\ldots,f_t\rangle \subseteq K[x_1,x_2,\ldots,x_n]$ un ideal. Entonces, $f \in \sqrt{I}$ si y solo si $\{1\}$ es la base de Gröbner reducida de $\langle f_1,\ldots,f_t, f\cdot y - 1\rangle \subseteq K[x_1,\ldots,x_n,y]$ respecto de un orden monomial cualquiera, siendo y una variable auxiliar.*

Demostración. Si $f \in \sqrt{I}$, entonces $f_1,\ldots,f_t, f\cdot y - 1$, considerados como polinomios de $K[x_1,\ldots,x_n,y]$, no tienen ceros comunes en ninguna extensión de K (tampoco algebraicamente cerrada). Es decir, $\mathbb{V}(f_1,\ldots,f_t, f\cdot y - 1) = \emptyset$, y por el Nullstellensatz débil, $\langle f_1,\ldots,f_t, f\cdot y - 1\rangle = \{1\}$, esto es, la base de Gröbner reducida de $\{f_1,\ldots,f_t, f\cdot y - 1\}$ es $\{1\}$. En ese caso, existen $q_i \in K[x_1,\ldots,x_n,y]$ tales que

$$q_1 f_1 + \cdots q_n f_n + q_{n+1}(f\cdot y - 1) = 1$$

Haciendo $y = 1/f$ y eliminando denominadores, se obtiene

$$a_1 f_1 + \cdots a_n f_n = f^s$$

para cierto entero $s \geq 1$ y polinomios $a_i \in K[x_1, x_2, \ldots, x_n]$. En consecuencia, $f \in \sqrt{I}$. ■

De esta forma, $H \Longrightarrow T$ es universalmente cierto si y solo si la base de Gröbner reducida de $\{f_1, \ldots, f_t, f \cdot y - 1\}$ es $\{1\}$.

4.1.1 Algoritmo de Buchberger

Una base $H = \{h_1, \ldots, h_n\}$ no suele ser una base de Gröbner. El algoritmo de Buchberger puede utilizarse para convertir una base en una base de Gröbner. Sea $mcm(a, b)$ el mínimo común múltiplo de los monomios a y b. La idea básica del algoritmo de Buchberger es la siguiente:

- Ordenar $h_1, \ldots, h_n$, tal que $LT(h_1) < \ldots < LT(h_n)$.
- Para cada par h_i, h_j, añadir el siguiente polinomio a la base:

$$\frac{mcm(LT(h_i), LT(h_j))}{LT(h_i)} h_i - \frac{mcm(LT(h_i), LT(h_j))}{LT(h_j)} h_j$$

- Para cada $i, j \in \{1, \ldots, n\}$ (en orden arbitrario), sustituir h_i por $hi \ (mod\ h_j)$.
- Eliminar los polinomios que son 0.

Estos pasos se repiten hasta que la base ya no cambia. Entonces se llega a una base de Gröbner.

El principal método de demostración automática de teoremas geométricos que utilizaremos está basado en los teoremas anteriores y funciona como sigue:

Método. Asumimos que $x_1 < \cdots < x_n$ según un orden lexicográfico entre monomios de $\mathbb{R}[x_1, \ldots, x_n]$.

- Paso 1. Calcular la base de Gröbner de $h_1, \ldots, h_n$ en $\mathbb{R}[x_1, \ldots, x_n]$; luego se reduce g para examinar si $g \in H = \langle h_1, \ldots, h_n \rangle$. Si es así, $H \Longrightarrow T$ es universalmente cierto. Si no es así, entonces pasamos al:
- Paso 2. Calcular la base de Gröbner de $h_1, \ldots, h_n, g \cdot y - 1$ en $\mathbb{R}[x_1, \ldots, x_n, y]$. Así, $H \Longrightarrow T$ es universalmente cierto si y solo si esta base es $\{1\}$.

El Paso 1 puede ser considerado como una primera aproximación, sin embargo, en la gran mayoría de teoremas que hemos encontrado en la práctica, $I = \sqrt{I}$. Un ejemplo en el que $I \subsetneq \sqrt{I}$ es el Teorema del pentágono planario que lo hemos estudiado pero no lo incluimos por limitaciones en la extensión del trabajo.

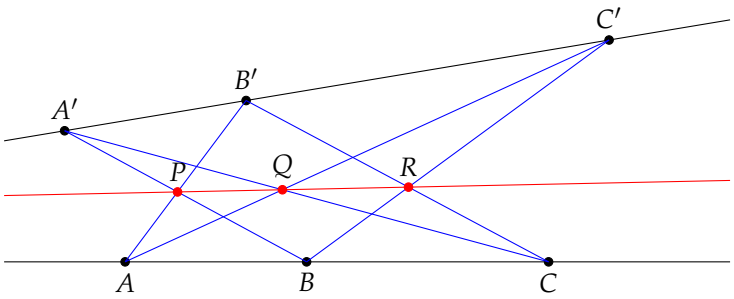
5 Resultados

5.1 Teorema de Pappus

Teorema 5 (Pappus). *Sean A, B, C y A', B', C' dos ternas de puntos alineados, y sean*

$$P = \overline{AB'} \cap \overline{A'B}$$
$$Q = \overline{AC'} \cap \overline{A'C}$$
$$R = \overline{BC'} \cap \overline{B'C}$$

como se muestra en la siguiente figura:



entonces los puntos P, Q, R están alineados.

Demostración. Sin perder generalidad, consideremos $A(0,0)$, $B(u_1,0)$ y $C(u_2,0)$ con $u_1 \neq u_2$ y distintos de 0.

Sean $A'(u_3,u_4)$, $B'(u_5,u_6)$ y $C'(u_7,x_1)$ otros tres puntos distintos entre si.

Llamemos $P(x_2,x_3)$, $Q(x_4,x_5)$ y $R(x_6,x_7)$ a los puntos que queremos demostrar que están alineados.

Convenimos en llamar u_i a las coordenadas arbitrarias y x_j a las coordenadas dependientes (determinadas por u_i).

Las hipótesis se pueden poner como polinomios

$$h_j \in k[u_1,u_2,\ldots,u_7,x_1,x_2,\ldots,x_7]$$

Por hipótesis, A', B' y C' están alineados, entonces

$$\frac{u_5-u_3}{u_7-u_3} = \frac{u_6-u_4}{x_1-u_4} \Longrightarrow (x_1-u_4)(u_5-u_3) = (u_6-u_4)(u_7-u_3)$$

y nos queda la primera hipótesis:

$$h_1 = (x_1-u_4)(u_5-u_3) - (u_6-u_4)(u_7-u_3) = 0$$

Como $P \in \overline{AB'}$, tenemos

$$\frac{x_2}{u_5} = \frac{x_3}{u_6} \Longrightarrow x_2u_6 - x_3u_5 = 0$$

y resulta la segunda hipótesis: $h_2 = x_2u_6 - x_3u_5 = 0$

También $P \in \overline{A'B}$ y entonces

$$\frac{x_2 - u_3}{u_3 - u_1} = \frac{x_3 - u_4}{u_4}$$

es la tercera hipótesis: $h_3 = (x_3 - u_4)(u_3 - u_1) - u_4(x_2 - u_3) = 0$

Análogamente

$$\begin{aligned}
Q \in \overline{AC'} &\Longrightarrow h_4 = x_1x_4 - x_5u_7 = 0 \\
Q \in \overline{A'C} &\Longrightarrow h_5 = (x_5 - u_4)(u_3 - u_2) - u_4(x_4 - u_3) = 0 \\
R \in \overline{BC'} &\Longrightarrow h_6 = x_1(x_6 - u_1) - x_7(u_7 - u_1) = 0 \\
R \in \overline{B'C} &\Longrightarrow h_7 = (x_7 - u_6)(u_5 - u_2) - u_6(x_6 - u_5) = 0
\end{aligned}$$

Queremos demostrar que P, Q y R están alineados, o lo que es lo mismo, la tesis puede escribirse como: $g = (x_5 - x_3)(x_6 - x_2) - (x_4 - x_2)(x_7 - x_3) = 0$

Ocurre siempre, que el número de ecuaciones de las hipótesis (h_j) es igual al número de variables dependientes (x_j).

Tenemos que demostrar que g se anula siempre que se anulen las hipótesis h_j con $j = 1,2,\ldots,7$ o lo que es lo mismo, g está en el ideal generado por las hipótesis.

El programa SageMath nos ayuda a calcular las Bases de Gröbner y también a ver si $g \in \mathbb{I} = \langle h_1, h_2, \ldots, h_7 \rangle$

Podemos ver una implementación en SageMath de este teorema.

■

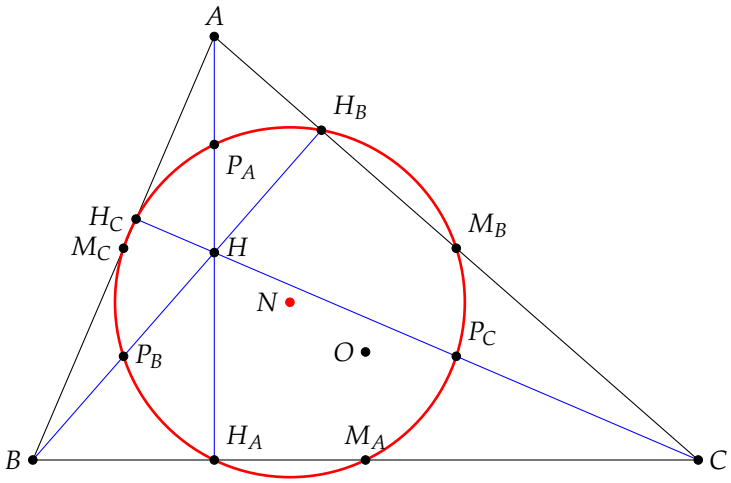
```
[1]: P.<u1,u2,u3,u4,u5,u6,u7> = PolynomialRing(QQ)

[2]: F=Frac(P) # o F=P.fraction_field()

[3]: F

[3]: Fraction Field of Multivariate Polynomial Ring in u1, u2, u3,␣
      ↪u4, u5, u6, u7
     over Rational Field

[4]: R.<x1,x2,x3,x4,x5,x6,x7>= PolynomialRing(F)

[5]: h1=(x1-u4)*(u5-u3)-(u6-u4)*(u7-u3)

[6]: h2=x2*u6-x3*u5

[7]: h3=(x3-u4)*(u3-u1)-u4*(x2-u3)

[8]: h4=x1*x4-x5*u7

[9]: h5=(x5-u4)*(u3-u2)-u4*(x4-u3)

[10]: h6=x1*(x6-u1)-x7*(u7-u1)
```

```
[11]: h7=(x7-u6)*(u5-u2)-u6*(x6-u5)

[12]: H_tilde=ideal(h1,h2,h3,h4,h5,h6,h7)

[13]: H_tilde.groebner_basis()

[13]: [x1 + (u4*u5 - u3*u6 - u4*u7 + u6*u7)/(u3 - u5), x2 +␣
      ↪(-u1*u4*u5)/(u4*u5 + u1*u6
      - u3*u6), x3 + (-u1*u4*u6)/(u4*u5 + u1*u6 - u3*u6), x4 +␣
      ↪(-u2*u3*u4*u7 +
      u2*u4*u5*u7)/(-u2*u4*u5 + u3*u4*u5 + u2*u3*u6 - u3^2*u6 +␣
      ↪u2*u4*u7 - u4*u5*u7 -
      u2*u6*u7 + u3*u6*u7), x5 + (u2*u4^2*u5 - u2*u3*u4*u6 -␣
      ↪u2*u4^2*u7 +
      u2*u4*u6*u7)/(-u2*u4*u5 + u3*u4*u5 + u2*u3*u6 - u3^2*u6 +␣
      ↪u2*u4*u7 - u4*u5*u7 -
      u2*u6*u7 + u3*u6*u7), x6 + (-u1*u2*u4*u5 + u1*u4*u5^2 +␣
      ↪u1*u2*u5*u6 -
      u1*u3*u5*u6 + u1*u2*u4*u7 - u1*u4*u5*u7 - u1*u2*u6*u7 +␣
      ↪u2*u3*u6*u7 +
      u1*u5*u6*u7 - u2*u5*u6*u7)/(u2*u4*u5 - u4*u5^2 + u1*u3*u6 -␣
      ↪u2*u3*u6 - u1*u5*u6
      + u3*u5*u6 - u2*u4*u7 + u4*u5*u7 + u2*u6*u7 - u3*u6*u7), x7 +␣
      ↪(u1*u4*u5*u6 -
      u2*u4*u5*u6 - u1*u3*u6^2 + u2*u3*u6^2 - u1*u4*u6*u7 +␣
      ↪u2*u4*u6*u7 + u1*u6^2*u7 -
      u2*u6^2*u7)/(u2*u4*u5 - u4*u5^2 + u1*u3*u6 - u2*u3*u6 -␣
      ↪u1*u5*u6 + u3*u5*u6 -
      u2*u4*u7 + u4*u5*u7 + u2*u6*u7 - u3*u6*u7)]

[14]: g=(x5-x3)*(x6-x2)-(x4-x2)*(x7-x3)

[15]: g in H_tilde

[15]: True
```

5.2 Circunferencia de 9 puntos

Teorema 6 (Circunferencia de 9 puntos). *En todo triángulo, los puntos medios de los lados, los pies de las alturas y los puntos medios que unen cada vertice con el ortocentro están sobre una circunferencia con centro el punto medio del ortocentro y el circuncentro.*

Sean A, B, C los vertices de un triángulo, M_A, M_B, M_C los puntos medios de los lados, H_A, H_B, H_C los pies de las alturas, P_A, P_B, P_C los puntos medios de los segmentos que unen cada vértice con el ortocentro H, y sean O, el circuncentro, y N, el punto medio de O y H.



Tomamos los puntos $A = (u_2, u_3)$; $B = (0,0)$ y $C = (u_1, 0)$. El resto de puntos quedan determinados a partir de estos:

$$\begin{array}{lll} H_A = (u_2, 0) & H_B = (x_1, x_2) & H_C = (x_3, x_4) \\ M_A = \left(\frac{u_1}{2}, 0\right) & M_B = \left(\frac{u_1+u_2}{2}, \frac{u_3}{2}\right) & M_C = \left(\frac{u_2}{2}, \frac{u_3}{2}\right) \\ H = (u_2, x_5) & O = \left(\frac{u_1}{2}, x_6\right) & N = \left(\frac{2u_2+u_1}{4}, \frac{x_5+x_6}{2}\right) \\ P_A = \left(u_2, \frac{u_3+x_5}{2}\right) & P_B = \left(\frac{u_2}{2}, \frac{x_5}{2}\right) & P_C = \left(\frac{u_1+u_2}{2}, \frac{x_5}{2}\right) \end{array}$$

Las variables dependientes se determinan en función de las siguientes condiciones para cierto $\lambda \in \mathbb{R}$:

- H_B pertenece a la recta AC y $H_B B \perp AC$:

$$\overrightarrow{CH_B} = \lambda \overrightarrow{CA} \implies (x_1 - u_1, x_2) = \lambda(u_2 - u_1, u_3) \implies \frac{x_1 - u_1}{u_2 - u_1} = \frac{x_2}{u_3}$$

$$\overrightarrow{BH_B} \cdot \overrightarrow{CA} = 0 \implies (x_1, x_2) \cdot (u_2 - u_1, u_3) = 0$$

  Tenemos las dos primeras hipótesis:

$$\begin{aligned} h_1 &= u_3(x_1 - u_1) - x_2(u_2 - u_1) = 0 \\ h_2 &= x_1(u_2 - u_1) + x_2 u_3 = 0 \end{aligned}$$

- H_C pertenece a la recta AB y $CH_C \perp AB$:

$$\overrightarrow{BH_C} = \lambda \overrightarrow{BA} \implies (x_3, x_4) = \lambda(u_2, u_3) \implies \frac{x_3}{u_2} = \frac{x_4}{u_3}$$

$$\overrightarrow{CH_C} \cdot \overrightarrow{BA} = 0 \implies (x_3 - u_1, x_4) \cdot (u_2, u_3) = 0$$

  Tercera y cuarta hipótesis:

$$\begin{aligned} h_3 &= u_3 x_3 - x_4 u_2 = 0 \\ h_4 &= u_2(x_3 - u_1) + x_4 u_3 = 0 \end{aligned}$$

- $BH \perp AC$

$$\overrightarrow{BH} \cdot \overrightarrow{CA} = 0 \implies (u_2, x_5) \cdot (u_2 - u_1, u_3) = 0$$
$$h_5 = u_2(u_2 - u_1) + x_5 u_3 = 0$$

- $OM_C \perp AB$

$$\overrightarrow{M_C O} \cdot \overrightarrow{BA} = 0 \implies \left(\frac{u_2 - u_1}{2}, \frac{u_3}{2} - x_6\right) \cdot (u_2, u_3) = 0$$

$$h_6 = u_2\frac{u_2 - u_1}{2} + u_3\left(\frac{u_3}{2} - x_6\right) = 0$$

Con esto, ya tenemos las 6 hipótesis. Las tesis son que la distancia desde N hasta los puntos $H_A, H_B, H_C, M_A, M_B, M_C, P_A, P_B, P_C$ es la misma. Para obtener polinomios, calculamos los cuadrados de las distancias:

$$\overline{NH_A}^2 = \left(\frac{2u_2 + u_1}{4} - u_2\right)^2 + \left(\frac{x_5 + x_6}{2}\right)^2 = \left(\frac{u_1 - 2u_2}{4}\right)^2 + \left(\frac{x_5 + x_6}{2}\right)^2$$

$$\overline{NM_A}^2 = \left(\frac{2u_2 + u_1}{4} - \frac{u_1}{2}\right)^2 + \left(\frac{x_5 + x_6}{2}\right)^2 = \left(\frac{2u_2 - u_1}{4}\right)^2 + \left(\frac{x_5 + x_6}{2}\right)^2$$

$$\overline{NH_B}^2 = \left(\frac{2u_2 + u_1}{4} - x_1\right)^2 + \left(\frac{x_5 + x_6}{2} - x_2\right)^2$$

$$\overline{NM_B}^2 = \left(\frac{2u_2 + u_1}{4} - \frac{u_1 + u_2}{2}\right)^2 + \left(\frac{x_5 + x_6}{2} - \frac{u_3}{2}\right)^2 = \left(\frac{u_1}{4}\right)^2 + \left(\frac{x_5 + x_6}{2} - \frac{u_3}{2}\right)^2$$

$$\overline{NH_C}^2 = \left(\frac{2u_2 + u_1}{4} - x_3\right)^2 + \left(\frac{x_5 + x_6}{2} - x_4\right)^2$$

$$\overline{NM_C}^2 = \left(\frac{2u_2 + u_1}{4} - \frac{u_2}{2}\right)^2 + \left(\frac{x_5 + x_6}{2} - \frac{u_3}{2}\right)^2 = \left(\frac{u_1}{4}\right)^2 + \left(\frac{x_5 + x_6}{2} - \frac{u_3}{2}\right)^2$$

$$\overline{NP_A}^2 = \left(\frac{2u_2 + u_1}{4} - u_2\right)^2 + \left(\frac{x_5 + x_6}{2} - \frac{u_3 + x_5}{2}\right)^2 = \left(\frac{u_1 - 2u_2}{4}\right)^2 + \left(\frac{x_6 - u_3}{2}\right)^2$$

$$\overline{NP_B}^2 = \left(\frac{2u_2 + u_1}{4} - \frac{u_2}{2}\right)^2 + \left(\frac{x_5 + x_6}{2} - \frac{x_5}{2}\right)^2 = \left(\frac{u_1}{4}\right)^2 + \left(\frac{x_6}{2}\right)^2$$

$$\overline{NP_C}^2 = \left(\frac{2u_2 + u_1}{4} - \frac{u_1 + u_2}{2}\right)^2 + \left(\frac{x_5 + x_6}{2} - \frac{x_5}{2}\right)^2 = \left(\frac{u_1}{4}\right)^2 + \left(\frac{x_6}{2}\right)^2$$

Se puede observar a simple vista que $\overline{NH_A}^2 = \overline{NM_A}^2$, $\overline{NM_B}^2 = \overline{NM_C}^2$ y $\overline{NP_B}^2 = \overline{NP_C}^2$, por lo que podemos reducir las tesis a 5: $\overline{NH_A}^2 = \overline{NH_B}^2$, $\overline{NH_A}^2 = \overline{NM_B}^2$, $\overline{NH_A}^2 = \overline{NH_C}^2$, $\overline{NH_A}^2 = \overline{NP_A}^2$ y $\overline{NH_A}^2 = \overline{NP_B}^2$.

$$g_1 = \left(\frac{u_1 - 2u_2}{4}\right)^2 + \left(\frac{x_5 + x_6}{2}\right)^2 - \left(\frac{2u_2 + u_1}{4} - x_1\right)^2 - \left(\frac{x_5 + x_6}{2} - x_2\right)^2 = 0$$

$$g_2 = \left(\frac{u_1 - 2u_2}{4}\right)^2 + \left(\frac{x_5 + x_6}{2}\right)^2 - \left(\frac{u_1}{4}\right)^2 - \left(\frac{x_5 + x_6}{2} - \frac{u_3}{2}\right)^2 = 0$$

$$g_3 = \left(\frac{u_1 - 2u_2}{4}\right)^2 + \left(\frac{x_5 + x_6}{2}\right)^2 - \left(\frac{2u_2 + u_1}{4} - x_3\right)^2 - \left(\frac{x_5 + x_6}{2} - x_4\right)^2 = 0$$

$$g_4 = \left(\frac{u_1 - 2u_2}{4}\right)^2 + \left(\frac{x_5 + x_6}{2}\right)^2 - \left(\frac{u_1 - 2u_2}{4}\right)^2 - \left(\frac{x_6 - u_3}{2}\right)^2 = 0$$

$$g_5 = \left(\frac{u_1 - 2u_2}{4}\right)^2 + \left(\frac{x_5 + x_6}{2}\right)^2 - \left(\frac{u_1}{4}\right)^2 - \left(\frac{x_6}{2}\right) = 0$$

[1]:
```
P.<u1,u2,u3>= PolynomialRing(QQ)
```

[2]:
```
F=Frac(P)
```

[3]:
```
R.<x1,x2,x3,x4,x5,x6>= PolynomialRing(F)
```

[4]:
```
h1=(x1-u1)*u3-x2*(u2-u1)
h2=x1*(u2-u1)+x2*u3
h3=x3*u3-x4*u2
h4=(x3-u1)*u2+x4*u3
h5=u2*(u2-u1)+x5*u3
h6=u2*((u2-u1)/2)+u3*(u3/2-x6)
```

[5]:
```
H_tilde=ideal(h1,h2,h3,h4,h5,h6)
```

[6]:
```
H_tilde.groebner_basis()
```

[6]:
```
[x1 + (-u1*u3^2)/(u1^2 - 2*u1*u2 + u2^2 + u3^2), x2 + (-u1^2*u3
 ↪+
u1*u2*u3)/(u1^2 - 2*u1*u2 + u2^2 + u3^2), x3 + (-u1*u2^2)/(u2^2
 ↪+ u3^2), x4 +
(-u1*u2*u3)/(u2^2 + u3^2), x5 + (-u1*u2 + u2^2)/u3, x6 + (u1*u2
 ↪- u2^2 -
u3^2)/(2*u3)]
```

[7]:
```
g1=((u1-2*u2)/4)^2+((x5+x6)/2)^2-((u1+2*u2)/4-x1)^2-((x5+x6)/
 ↪2-x2)^2
g2=((u1-2*u2)/4)^2+((x5+x6)/2)^2-((u1)/4)^2-((x5+x6)/2-u3/2)^2
g3=((u1-2*u2)/4)^2+((x5+x6)/2)^2-((u1+2*u2)/4-x3)^2-((x5+x6)/
 ↪2-x4)^2
g4=((u1-2*u2)/4)^2+((x5+x6)/2)^2-((u1-2*u2)/4)^2-((x6-u3)/2)^2
g5=((u1-2*u2)/4)^2+((x5+x6)/2)^2-((u1)/4)^2-(x6/2)^2
```

```
[8]: g1 in H_tilde
```

```
[8]: True
```

```
[9]: g2 in H_tilde
```

```
[9]: True
```

```
[10]: g3 in H_tilde
```

```
[10]: True
```

```
[11]: g4 in H_tilde
```

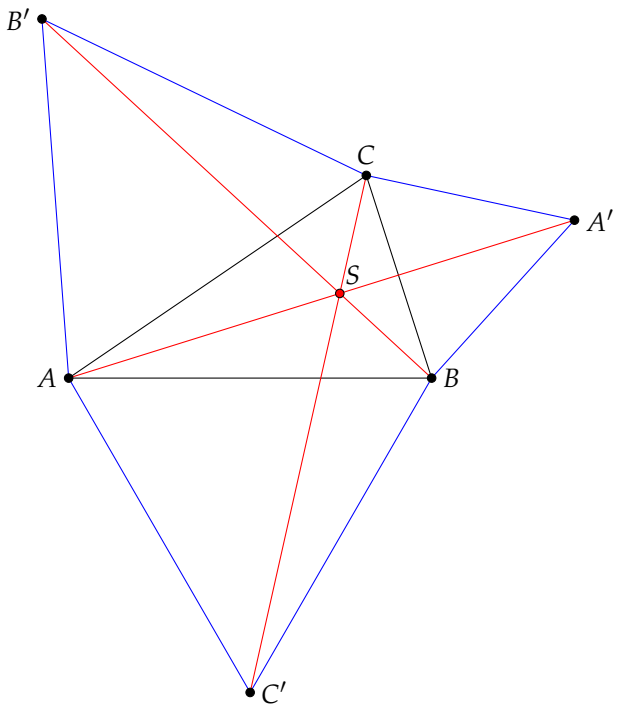
```
[11]: True
```

```
[12]: g5 in H_tilde
```

```
[12]: True
```

5.3 Teorema de Fermat

Teorema 7 (Teorema del punto de Fermat). *Sea $\triangle ABC$ un triángulo en el plano. Construimos los puntos A', B', C' de tal forma que los triángulos $\triangle A'BC$, $\triangle AB'C$ y $\triangle ABC'$ sean equiláteros.*



Los tres segmentos $\overline{AA'}$, $\overline{BB'}$ y $\overline{CC'}$ se cortan en un único punto S que llamamos Punto de Fermat *del triángulo.*

Demostración. Asignamos coordenadas a los puntos: $A(0,0)$, $B(u_1,0)$, $C(u_2,u_3)$, $A'(x_1,x_2)$, $B'(x_3,x_4)$, $C'(x_5,x_6)$ y $S(x_7,x_8)$.

Dado que el triángulo $\triangle ABC'$ es equilátero, la longitud de sus lados es la misma, entonces:

$$\overline{AB} = \overline{AC'} \Longrightarrow h_1 = x_5^2 + x_6^2 - u_1^2 = 0$$
$$\overline{AB} = \overline{BC'} \Longrightarrow h_2 = x_5^2 + x_6^2 - 2u_1x_5 = 0$$

De forma análoga, para los triángulos equiláteros $\triangle A'BC$, $\triangle AB'C$ nos quedaría:

$$\overline{AC} = \overline{AB'} \Longrightarrow h_3 = x_3^2 + x_4^2 - u_2^2 - u_3^2 = 0$$
$$\overline{AC} = \overline{CB'} \Longrightarrow h_4 = x_3^2 + x_4^2 - 2u_2x_3 - 2u_3x_4 = 0$$
$$\overline{BC} = \overline{BA'} \Longrightarrow h_5 = x_1^2 + x_2^2 - 2u_1x_1 - u_2^2 - u_3^2 + 2u_1u_2 = 0$$
$$\overline{BC} = \overline{CA'} \Longrightarrow h_6 = x_1^2 + x_2^2 - 2u_2 - x_1 - 2u_3x_2 - u_1^2 + 2u_1u_2 = 0$$

Como los puntos A, S y A' son colineales, nos queda $h_7 = x_1x_8 - x_2x_7 = 0$.
Por la colinealidad de B, S y B', tendremos $h_8 = (x_7 - u_1)x_4 + (u_1 - x_3)x_8 = 0$.
La conclusión sería la colinealidad de C, S y C', o lo que es lo mismo:

$$g = (x_7 - x_5)u_3 + (u_2 - x_7)x_6 + (x_5 - u_2)x_8 = 0$$

Cuando implementamos en SageMath estas hipótesis, la conclusión que nos reporta es que g no está en el ideal. ¿Cuál es el motivo? La construcción que aparece en la figura no es la única. Hay dos posibilidades para A', en el plano, formando un triángulo equilátero con el segmento $\overline{BC}$. Lo mismo se podría decir para los puntos B' y C' y, parece ser, que no para todas ellas es cierto el teorema.

Cambiemos, entonces, el enunciado del teorema para exigir que los triángulos equiláteros estén apuntando hacia fuera del triángulo inicial, tal y como aparecen en la imagen anterior. También sería un teorema válido si los tres triángulos apuntasen hacia dentro, como pudimos comprobar.

Con las mismas coordenadas de los puntos, definamos los siguientes vectores

$$\overrightarrow{v_1} = \overrightarrow{AB} = (u_1,0)^T \qquad \overrightarrow{w_1} = \overrightarrow{AC'} = (x_5,x_6)^T$$
$$\overrightarrow{v_2} = \overrightarrow{BC} = (u_2 - u_1, u_3)^T \qquad \overrightarrow{w_2} = \overrightarrow{BA'} = (x_1 - u_1, x_2)^T$$
$$\overrightarrow{v_3} = \overrightarrow{AB'} = (x_3,x_4)^T \qquad \overrightarrow{w_3} = \overrightarrow{AC} = (u_2,u_3)^T$$

Tenemos que los vectores $\overrightarrow{w_i}$ se obtienen de los $\overrightarrow{v_i}$ con una rotación de $\theta = \pi/3$ radianes, en sentido horario.

La matriz de rotación

$$R(-\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

nos permite obtener las coordenadas de $\overrightarrow{w_i}$ a partir de $\overrightarrow{v_i}$, así:

$$\begin{pmatrix} x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ 0 \end{pmatrix} \Longrightarrow \begin{matrix} x_5 = \frac{1}{2}u_1 \\ x_6 = -\frac{\sqrt{3}}{2}u_1 \end{matrix}$$

$$\begin{pmatrix} x_1 - u_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} u_2 - u_1 \\ u_3 \end{pmatrix} \Longrightarrow \begin{matrix} x_1 - u_1 = \frac{1}{2}(u_2 - u_1) + \frac{\sqrt{3}}{2}u_3 \\ x_2 = -\frac{\sqrt{3}}{2}(u_2 - u_1 + \frac{1}{2}u_3 \end{matrix}$$

$$\begin{pmatrix} u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} \Longrightarrow \begin{matrix} u_2 = \frac{1}{2}x_3 + \frac{\sqrt{3}}{2}x_4 \\ u_3 = -\frac{\sqrt{3}}{2}x_3 + \frac{1}{2}x_4 \end{matrix}$$

Quitando denominadores y agrupando, tenemos las primeras seis hipótesis, en el orden que fueron apareciendo:

$$\begin{aligned} h_1 &= 2x_5 - u_1 \\ h_2 &= 2x_6 + \sqrt{3}u_1 \\ h_3 &= 2x_1 - u_1 - u_2 - \sqrt{3}u_3 \\ h_4 &= 2x_2 + \sqrt{3}(u_2 - u_1) - u_3 \\ h_5 &= 2u_2 - x_3 - \sqrt{3}x_4 \\ h_6 &= 2u_3 + \sqrt{3}x_3 - x_4 \end{aligned}$$

Las hipótesis h_7 y h_8 son las mismas que teníamos y la conclusión g también.

Al implementar todo esto en SageMath tendremos que, ahora sí, $g \in \langle h_1, h_2, \ldots, h_8 \rangle$ con lo cual quedaría demostrado el teorema con el nuevo planteamiento que le hemos dado. ■

En las nuevas hipótesis aparece la expresión $\sqrt{3}$ que no es racional y tuvimos que extender el conjunto $\mathbb{Q}$ con las soluciones del polinomio $x^2 - 3 = 0$.

Una vez que tenemos las hipótesis con los tres triángulos equiláteros apuntando hacia fuera podemos demostrar más cosas, por ejemplo, que la distancia entre A y A' es la misma que entre B y B' y también que entre C y C'. Para el primer caso construimos la tesis

$$g_1 = x_1^2 + x_2^2 - (x_3 - u_1)^2 - x_4^2$$

y para el otro

$$g_2 = x_1^2 + x_2^2 - (x_5 - u_2)^2 - (x_6 - u_3)^2$$

En ambos casos SageMat nos confirma que $g_i \in \langle h_1, h_2, \ldots, h_8 \rangle$.

Una demostración visual, muy bonita, de este último resultado podemos hacerla apoyando el compás en el punto B y viendo que con un giro de $\pi/3$ radianes en sentido antihorario llevamos A' en C y A en C'. De forma análoga se puede hacer para el otro segmento.

Otra solución para el problema es utilizando una tercera dimensión y el producto vectorial. Mantenemos el triángulo en el plano XY, y el eje Z nos indica en qué sentido está un ángulo entre 2 vectores.

Vamos a situar cada punto en la mediatriz de los lados, y utilizando el producto vectorial y el hecho de que los ángulos del triángulo son de 60°:

$$AC' = BC' \implies x_5^2 + x_6^2 = (x_5 - u_1)^2 + x_6^2$$

$$\frac{\overrightarrow{AB} \times \overrightarrow{AC'}}{\|\overrightarrow{AB}\|^2 \cdot \text{sen}(60^\circ)} = (0,0,-1)$$

$$BA' = CA' \implies (x_1 - u_1)^2 + x_2^2 = (x_1 - u_2)^2 + (x_2 - u_3)^2$$

$$\frac{\overrightarrow{BC} \times \overrightarrow{BA'}}{\|\overrightarrow{BC}\|^2 \cdot \text{sen}(60^\circ)} = (0,0,-1)$$

$$CB' = AB' \implies (x_3 - u_2)^2 + (x_4 - u_3)^2 = x_3^2 + x_4^2$$

$$\frac{\overrightarrow{CA} \times \overrightarrow{CB'}}{\|\overrightarrow{CA}\|^2 \cdot \text{sen}(60^\circ)} = (0,0,-1)$$

Obtenemos las seis primeras hipótesis:

$$h_1 = x_5^2 - (x_5 - u_1)^2 = 0$$

$$h_2 = x_6 + \frac{\sqrt{3}}{2} u_1 = 0$$

$$h_3 = (x_1 - u_1)^2 + x_2^2 - (x_1 - u_2)^2 - (x_2 - u_3) = 0$$

$$h_4 = (u_1 - u_1)x_2 - (x_1 - u_1)u_3 + \frac{\sqrt{3}}{2}((u_2 - u_1)^2 + u_3^2) = 0$$

$$h_5 = (x_3 - u_2)^2 + (x_4 - u_3)^2 - x_3^2 - x_4^2 = 0$$

$$h_6 = u_2(u_3 - x4) - u_3(u_2 - x_3) + \frac{\sqrt{3}}{2}(u_2^2 + u_3^2) = 0$$

Las dos últimas hipótesis son las mismas que teníamos al principio y también la tesis. Se pudo comprobar en SageMath que, también con esta nueva formulación de las hipótesis, manteniendo los triángulos apuntando hacia fuera, se cumple el teorema.

[1]:
```
E.<sqrt3> = QQ.extension(x^2 - 3)
```

[2]:
```
P.<u1,u2,u3> = PolynomialRing(E)
```

[3]:
```
F=Frac(P) # o F=P.fraction_field()
```

[4]:
```
F
```

[4]:
```
Fraction Field of Multivariate Polynomial Ring in u1, u2, u3␣
↪over Number Field
in sqrt3 with defining polynomial x^2 - 3
```

[5]:
```
R.<x1,x2,x3,x4,x5,x6,x7,x8>= PolynomialRing(F)
```

[6]:
```
h1=2*x5-u1
h2=2*x6+sqrt3*u1
h3=2*x1-u1-u2-sqrt3*u3
h4=2*x2+sqrt3*(u2-u1)-u3
h5=2*u2-x3-sqrt3*x4
```

```
h6=2*u3+sqrt3*x3-x4
h7=x1*x8-x2*x7
h8=(x7-u1)*x4+(u1-x3)*x8
```

[7]:
```
H_tilde=ideal(h1,h2,h3,h4,h5,h6,h7,h8)
```

[8]:
```
H_tilde.groebner_basis()
```

```
verbose 0 (3837: multi_polynomial_ideal.py, groebner_basis) Warning: falling
back to very slow toy implementation.
```

[8]:
```
[x1 - 1/2*u1 - 1/2*u2 + (-1/2*sqrt3)*u3, x2 + (-1/2*sqrt3)*u1 + (1/2*sqrt3)*u2 -
1/2*u3, x3 - 1/2*u2 + (1/2*sqrt3)*u3, x4 + (-1/2*sqrt3)*u2 - 1/2*u3, x5 -
1/2*u1, x6 + (1/2*sqrt3)*u1, x7 + ((-1/8*sqrt3)*u1^2*u2 + (-1/8*sqrt3)*u1*u2^2 -
1/8*u1^2*u3 - 1/2*u1*u2*u3 + (-1/8*sqrt3)*u1*u3^2)/((1/4*sqrt3)*u1^2 +
(-1/4*sqrt3)*u1*u2 + (1/4*sqrt3)*u2^2 + 3/4*u1*u3 + (1/4*sqrt3)*u3^2), x8 +
(-3/4*u1^2*u2 + 3/4*u1*u2^2 + (-1/4*sqrt3)*u1^2*u3 -
1/4*u1*u3^2)/((1/2*sqrt3)*u1^2 + (-1/2*sqrt3)*u1*u2 + (1/2*sqrt3)*u2^2 +
3/2*u1*u3 + (1/2*sqrt3)*u3^2)]
```

[9]:
```
g=(x7-x5)*u3+(u2-x7)*x6+(x5-u2)*x8
```

[10]:
```
g in H_tilde
```

[10]:
```
True
```

5.4 Una aplicación de las bases de Gröbner. Modelo algebraico del operón lac

> "Mathematics is Biology's next microscope, only better; Biology is Mathematics' next Physics, only Better."
>
> **— Joel E. Cohen, 2017**

En el anexo A se justifica el modelo matemático que se utilizará a continuación, así como una detallada descripción del operón lactosa.

A continuación, propondremos y justificaremos las funciones booleanas para cada una de las variables. En todas ellas, utilizaremos *estará* para referirnos al estado de una variable en el siguiente paso de tiempo, $t+1$. Por ejemplo, "X estará presente si Y y Z están

presentes" es expresado como $X(t+1) = Y(t) \wedge Z(t)$. Decimos que el estado de X es una función de los estados de Y y Z, y lo escribimos como $f_X = Y \wedge Z$.

Variables:

- M: ARNm
- P: permeasa lac
- B: β-galactosidasa
- C: proteína activadora de catabolitos (CAP)
- R: represor lac
- A: alolactosa
- A_l: mínimo nivel de alolactosa
- L: lactosa intracelular
- L_l: mínimo nivel de lactosa intracelular

En primer lugar, suponemos que las biomoléculas representadas por las variables se degradan en un *paso de tiempo*, con la excepción de las que existen en niveles altos, que se degradan en dos pasos de tiempo. Por ejemplo, si la β-galactosidasa está presente pero el ARNm no, entonces en el siguiente paso de tiempo, la concentración de β-galactosidasa se habrá degradado hasta los niveles basales. Sin embargo, si hay altos niveles de alolactosa pero no de lactosa, entonces en el siguiente paso de tiempo, se degradará a niveles medios, y luego a niveles traza en el siguiente paso de tiempo. Además, las funciones anteriores no son las únicas que podríamos haber propuesto. Por ejemplo, consideremos la afirmación niveles altos de la proteína represora estará presente si no hay alolactosa, que modelamos por $f_R = \overline{A} \wedge \overline{A_l}$. Podríamos haber elegido en cambio $f_R = \overline{A_l}$, porque si no hay ni siquiera niveles medios de lactosa presentes, no puede haber niveles altos. Así que se asumen los siguientes supuestos:

- La transcripción requiere de una "unidad de tiempo".
- Degradación de todas las proteínas y el ARNm ocurre en un "paso de tiempo".
- Altos niveles de lactosa o alolactosa en cualquier tiempo t implica un nivel mínimo de alolactosa para el siguiente paso de tiempo $t+1$.

Denotaremos por L_e y G_e los parámetros de lactosa y glucosa extracelular presente, respectivamente. Las funciones booleanas correspondientes son:

1. El ARNm se transcribirá si la proteína represora no está presente (es decir, está unida a la alolactosa) y el complejo proteico CAP está presente. Por lo tanto, $f_M(t+1) = \overline{R} \wedge C$.
2. Los productos génicos permeasa lac y β-galactosidasa estarán presentes si ARNm está presente. Por lo tanto, $f_P(t+1) = M$ y $f_B(t+1) = M$.

3. El complejo proteico CAP estará presente si no hay glucosa. Por lo tanto, $f_C(t+1) = \overline{G_e}$.

4. Los niveles de la proteína represora serán altos si no hay alolactosa. Así, $f_R(t+1) = \overline{A} \wedge \overline{A_l}$.

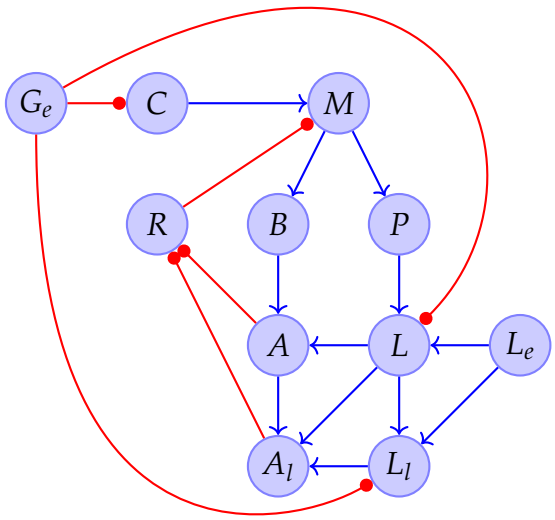
5. Habrá altos niveles de alolactosa si hay altos niveles de lactosa y β-galactosidasa disponibles. Así, $f_A(t+1) = L \wedge B$.

6. Habrá (al menos) niveles medios de alolactosa disponibles si hay (al menos) niveles medios de lactosa. En este caso, incluso los niveles basales de β-galactosidasa son suficientes para transformar la lactosa en alolactosa. Por lo tanto, $f_{A_l}(t+1) = A \vee L \vee L_l$.

7. Para tener altos niveles de lactosa, necesitamos altos niveles de lactosa extracelular, sin glucosa, y una permeasa de lactosa disponible. Así, $f_L(t+1) = \overline{G_e} \wedge P \wedge L_e$.

8. Para tener (al menos) niveles medios de lactosa, necesitamos que no haya glucosa extracelular y que se dé una de las dos situaciones siguientes (i) niveles elevados de lactosa extracelular, en cuyo caso una parte entrará en la célula a través del transporte de la permeasa basal, o (ii) niveles medios de lactosa extracelular y la proteína transportadora transportador. Así, $f_{L_l}(t+1) = \overline{G_e} \wedge (L \vee L_e)$.

En este punto, tenemos un modelo. Las variables y parámetros (nodos) y las funciones y relaciones (aristas en forma de flechas para una influencia positiva y terminadas en un círculo para una influencia negativa) se pueden representar mediante el grafo dirigido que se muestran en la siguiente figura.



Para analizar el modelo, será conveniente renombrar las variables utilizando números en lugar de letras, como sigue: $(M,P,B,C,R,A,A_l,L,L_l) = (x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9)$ y haciendo $f_{x_i}(t+1) = x_i$ para $1 \leq i \leq 9$, para encontrar los puntos fijos, se obtiene el siguiente sistema de ecuaciones en $\mathbb{Z}_2[x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9]$:

$$\left.\begin{aligned} x_1 + x_4x_5 + x_4 &= 0 \\ x_1 + x_2 &= 0 \\ x_1 + x_3 &= 0 \\ x_4 + (G_e+1) &= 0 \\ x_5 + x_6x_7 + x_6 + x_7 + 1 &= 0 \\ x_6 + x_3x_8 &= 0 \\ x_6 + x_7 + x_8 + x_9 + x_6x_8 + x_6x_9 + x_8x_9 + x_6x_8x_9 &= 0 \\ x_8 + x_2L_e(G_e+1) &= 0 \\ x_9 + (G_e+1)(x_8 + x_8L_e + L_e) &= 0 \end{aligned}\right\}$$

Se necesita resolver el sistema para las cuatro combinaciones:

$$(G_e,L_e) = (0,0),(0,1),(1,0),(1,1)$$

5.4.1 Lactosa presente, glucosa presente

Consideremos, en primer lugar, el caso $(G_e,L_e) = (1,1)$. A continuación utilizaremos los comandos de SageMath para realizar los cálculos.

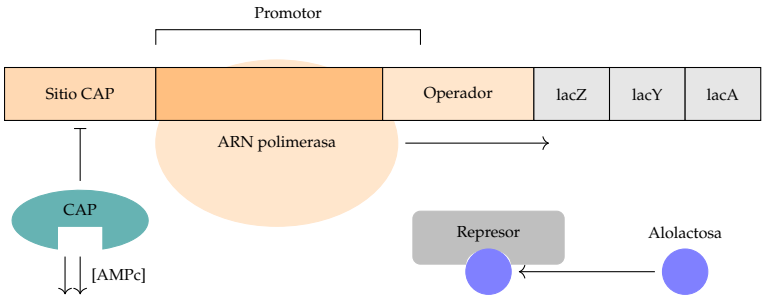


Figura 1: La alolactosa se une al represor y se usa antes la glucosa.

[1]:
```
P.<x1,x2,x3,x4,x5,x6,x7,x8,x9>= PolynomialRing(GF(2))
```

[2]:
```
Le=1
```

[3]:
```
Ge=1
```

[4]:
```
print ("Le=" , Le);
```

```
Le= 1
```

```
[5]: print ("Ge=" , Ge);
```

```
Ge= 1
```

```
[6]: f1=x1+x4+x4*x5
```

```
[7]: f2=x1+x2
```

```
[8]: f3=x1+x3
```

```
[9]: f4=x4+Ge+1
```

```
[10]: f5=x5+x6+x7+x6*x7+1
```

```
[11]: f6=x6+x3*x8
```

```
[12]: f7=x6+x7+x8+x9+x6*x8+x6*x9+x8*x9+x6*x8*x9
```

```
[13]: f8=x8+Le*x2+Ge*Le*x2
```

```
[14]: f9=x9+x8+Le+Le*x8+Ge*x8+Ge*Le+Ge*Le*x8
```

```
[15]: I=ideal(f1,f2,f3,f4,f5,f6,f7,f8,f9)
```

```
[16]: I
```

```
[16]: Ideal (x4*x5 + x1 + x4, x1 + x2, x1 + x3, x4, x6*x7 + x5 + x6 +␣
      ↪x7 + 1, x3*x8 +
      x6, x6*x8*x9 + x6*x8 + x6*x9 + x8*x9 + x6 + x7 + x8 + x9, x8,␣
      ↪x9) of
      Multivariate Polynomial Ring in x1, x2, x3, x4, x5, x6, x7, x8,␣
      ↪x9 over Finite
      Field of size 2
```

```
[17]: I.groebner_basis()
```

```
[17]: [x1, x2, x3, x4, x5 + 1, x6, x7, x8, x9]
```

De esta forma, resulta un sistema equivalente fácil de resolver con solución:

$$(M,P,B,C,R,A,A_l,L,L_l) = (0,0,0,0,1,0,0,0,0)$$

que tiene sentido biológico, y el operón es OFF.

5.4.2 Lactosa presente, glucosa ausente

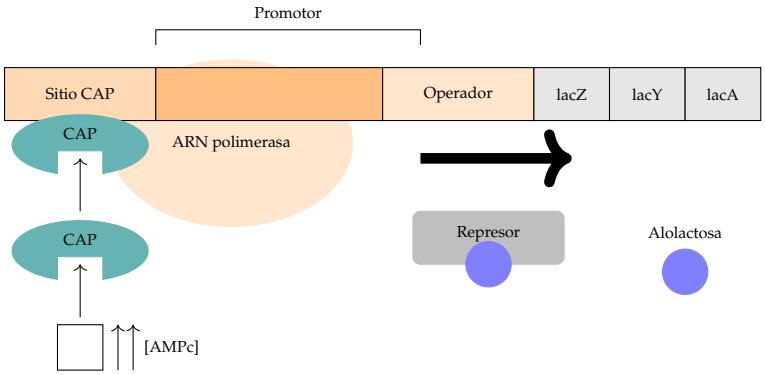


Figura 2: La alolactosa se une al represor y la ausencia de glucosa permite que el [AMPc] se una a la CAP, aumentando la tasa de transcripción.

En este caso $(G_e, L_e) = (0,1)$. De SageMath se obtiene:

[18]:
```
P.<x1,x2,x3,x4,x5,x6,x7,x8,x9>= PolynomialRing(GF(2))
```

[19]:
```
Le=1
```

[20]:
```
Ge=0
```

[21]:
```
print ("Le=" , Le);
```

```
Le= 1
```

[22]:
```
print ("Ge=" , Ge);
```

```
Ge= 0
```

[23]:
```
f1=x1+x4+x4*x5
```

[24]:
```
f2=x1+x2
```

[25]:
```
f3=x1+x3
```

[26]:
```
f4=x4+Ge+1
```

[27]:
```
f5=x5+x6+x7+x6*x7+1
```

[28]:
```
f6=x6+x3*x8
```

[29]:
```
f7=x6+x7+x8+x9+x6*x8+x6*x9+x8*x9+x6*x8*x9
```

[30]:
```
f8=x8+Le*x2+Ge*Le*x2
```

```
[31]: f9=x9+x8+Le+Le*x8+Ge*x8+Ge*Le+Ge*Le*x8

[32]: I=ideal(f1,f2,f3,f4,f5,f6,f7,f8,f9)

[33]: I

[33]: Ideal (x4*x5 + x1 + x4, x1 + x2, x1 + x3, x4 + 1, x6*x7 + x5 +␣
       ↪x6 + x7 + 1,
      x3*x8 + x6, x6*x8*x9 + x6*x8 + x6*x9 + x8*x9 + x6 + x7 + x8 +␣
       ↪x9, x2 + x8, x9 +
      1) of Multivariate Polynomial Ring in x1, x2, x3, x4, x5, x6,␣
       ↪x7, x8, x9 over
      Finite Field of size 2

[34]: I.groebner_basis()

[34]: [x1 + 1, x2 + 1, x3 + 1, x4 + 1, x5, x6 + 1, x7 + 1, x8 + 1, x9␣
       ↪+ 1]
```

De esta forma, resulta un sistema con solución:

$$(M,P,B,C,R,A,A_l,L,L_l) = (1,1,1,1,0,1,1,1,1)$$

que tiene sentido biológico, y el operón es ON.

5.4.3 Lactosa ausente, glucosa presente

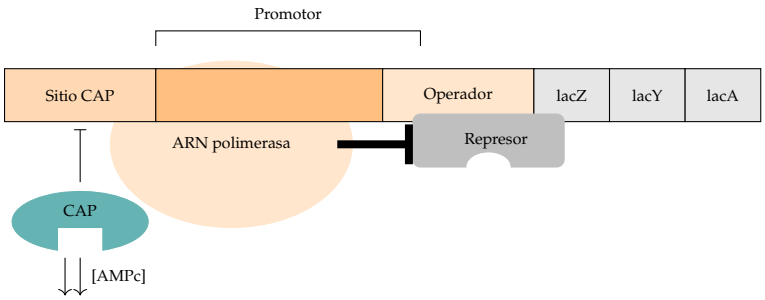


Figura 3: El represor se une al operador e impide la transcripción del ADN.

Input: $(G_e, L_e) = (1,0)$
Punto fijo: $(M,P,B,C,R,A,A_l,L,L_l) = (0,0,0,0,1,0,0,0,0)$.
Tiene sentido biológico, y el operón es OFF.

5.4.4 Lactosa ausente, glucosa ausente

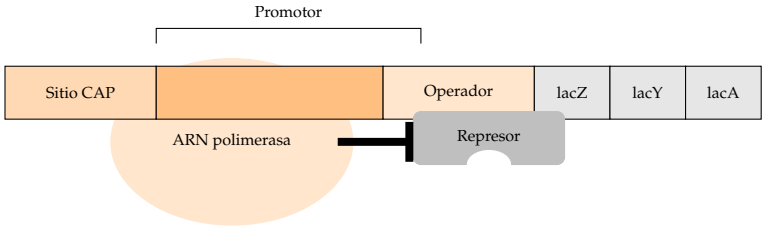


Figura 4: El represor se une al operador y el operón no funciona.

Input: $(G_e, L_e) = (0,0)$
Punto fijo: $(M,P,B,C,R,A,A_l,L,L_l) = (0,0,0,1,1,0,0,0,0)$.
Tiene sentido biológico, y el operón es OFF.

6 Conclusiones

Hemos aplicado las bases de Gröbner para demostrar varios teoremas de geometría elemental, como se vio en la sección Resultados. Los teoremas enunciados en los libros de texto a menudo, implícitamente, asumen condiciones de no degeneración sobre las figuras involucradas. Por ejemplo, cuando se habla de un triángulo ABC, a menudo se asume que ningún par de A,B,C coincide. Es posible encontrar automáticamente estos casos degenerados utilizando las bases de Gröbner. En la continuación de este trabajo, durante los próximos meses, trataremos estos casos mediante el llamado método de Wu, para lo cual hay que profundizar en la *seudodivisión polinómica* y la irreducibilidad de las variedades afines.

En este trabajo, se ofrece la aplicación de la teoría de demostración automática de teoremas geométricos importantes de la geometría elemental. En muchos ejemplos pudimos ver el núcleo de esta teoría que, especialmente en los últimos años, debido a la creciente potencia de los ordenadores y los nuevos algoritmos eficientes, cobra sentido. En todos los casos estudiados no hemos obviado las soluciones clásicas (vía geometría sintética) que no se incluyeron por razones de espacio, pero que descubren las ideas creativas del método puramente geométrico. En cualquier caso, no debemos albergar ninguna preferencia por ninguno de los dos métodos, ya que cada uno tiene sus propios puntos fuertes y débiles. Por lo tanto, no se puede decir que un método sea mejor que el otro. Por el contrario, ambos métodos deberían complementarse mutuamente.

También, existen casos en los que somos capaces de resolver un problema por un método clásico mientras que, por el contrario, un método informático falla. Este es un reto más para desarrollar la teoría de la demostración automática de teoremas para que siempre termine, es decir, que siempre podamos decir que una afirmación dada es verdadera o no (problema de decisión). Algunos problemas suponen un enorme coste de memoria computacional, lo que hace imposible la resolución de problemas más complejos.

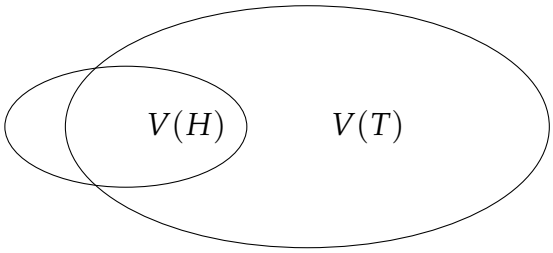
Un cierto desarrollo se logrará mediante la mejora de la eficiencia de ordenadores, pero el mayor y sustancial progreso en la resolución de problemas puede esperarse con

algoritmos mejores y más sofisticados y con formas más ingeniosas de solución de problemas.

Además de los dos métodos básicos de demostración mencionados —clásico y computacional—, nos queda pendiente aplicar los métodos vistos y algo más de maquinaria algebraica para estudiar la derivación automática y descubrimiento automático de teoremas.

Por derivación automática de teoremas, entendemos la búsqueda de fórmulas geométricas que se mantienen válidas entre las magnitudes geométricas prescritas y que se deducen de los supuestos dados. Esto se hace mediante el método de eliminación de variables (dependientes o independientes), obteniéndose el ideal de eliminación el cual contiene aquellas variables que no fueron eliminadas.

Por descubrimiento automático de teoremas, entendemos el proceso de tratar automáticamente con enunciados geométricos arbitrarios (es decir, enunciados que podrían ser, en general, falsos) y tratar de encontrar hipótesis complementarias tales que los enunciados sean verdaderos. La situación se puede ver en la siguiente figura para el esquema $H \Longrightarrow T$.



Podemos reaccionar ante esta situación de dos maneras: para que el enunciado $H \Longrightarrow T$ sea verdadero podemos "reducir" la variedad de hipótesis $V(H)$ o "ampliar" la variedad de conclusiones $V(T)$.

Agradecimientos

Queremos agradecer, en primer lugar, a nuestro director Dr. Carlos Ferreiro por aportarnos la base teórica necesaria para poder abordar este tema y por habernos ayudado y motivado tanto. Han sido muchas las tardes que pasamos en su despacho y muchos los aprendizajes que obtuvimos. Sin su inestimable ayuda, nada de esto hubiera sido posible.

Al Dr. Manuel Ladra por habernos propuesto este tema, por las directrices que nos fue marcando y los retos que nos iba proponiendo en cada momento.

A nuestro profesor de Biología, Dr. Francisco Boán, por las enseñanzas con el *Operón lac* y por estar siempre a nuestra disposición, incluso en vacaciones o festivos.

Bibliografía

[1] Shang-Ching Chou. *Mechanical Geometry Theorem Proving*. Mathematics and Its Applications. Springer, 1988.

[2] David Cox, John Little y Donal O'Shea. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*. Springer Science & Business Media, 2015.

[3] Leonard Eugene Dickson. «Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors». En: *American Journal of Mathematics* 35.4 (1913), págs. 413-422.

[4] Deepak Kapur. «Geometry theorem proving using Hilbert's Nullstellensatz». En: *Proceedings of the fifth ACM symposium on Symbolic and algebraic computation*. 1986, págs. 202-208.

[5] Stuart A Kauffman. «Metabolic stability and epigenesis in randomly constructed genetic nets». En: *Journal of theoretical biology* 22.3 (1969), págs. 437-467.

[6] María Pilar Páez. «Pruebas automáticas de teoremas geométricos». Trabajo de fin de grado dirigido por M. Ladra. Universidad de Santiago de Compostela, 2016.

[7] Pavel Pech. *Selected topics in geometry with classical vs. computer proving*. World Scientific Publishing Company, 2007.

[8] Raina Robeva y Terrell Hodge. *Mathematical concepts and methods in modern biology: using modern discrete models*. Academic Press, 2013.

[9] Alfred Tarski. «A Decision Method for Elementary Algebra and Geometry». En: (1948).

[10] Wen-Tsün Wu. *Mathematics Mechanization: Mechanical Geometry Theorem-Proving, Mechanical Geometry Problem-Solving and Polinomial Equations-Solving*. Kluwer, Dordrecht, 2001.

A Mecanismos de regulación de genes: modelos de redes booleanas del operón lactosa en *Escherichia coli*

Las matemáticas proporcionan un marco formal para organizar la abrumadora cantidad de datos experimentales dispares y para desarrollar modelos que reflejen las dependencias entre los componentes del sistema. Se han desarrollado diferentes tipos de modelos matemáticos para tratar de interpretar los mecanismos de regulación génica y su dinámica.

Los modelos deterministas generan exactamente los mismos resultados bajo un conjunto dado de condiciones iniciales, mientras que en los modelos estocásticos los resultados difieren debido a la aleatoriedad inherente. Los modelos dinámicos se centran en la evolución temporal de un sistema, mientras que los modelos estáticos no consideran el tiempo como parte del marco de modelización. Entre los modelos dinámicos, los modelos de tiempo continuo utilizan el tiempo como una variable continua, mientras que en los modelos discretos, el tiempo solo puede asumir valores enteros. Los modelos espacio-continuos se refieren a situaciones en las que las variables del modelo pueden asumir un continuo de valores, mientras que en los modelos espacio-discretos esas variables solo pueden asumir valores de un conjunto finito. Los modelos espacio-continuos de regulación génica suelen construirse en forma de ecuaciones diferenciales (en el caso del tiempo continuo) o de ecuaciones en diferencia (en el caso del tiempo discreto) y se centran en la cinética de las reacciones bioquímicas. Los modelos de tiempo discreto construidos a partir de funciones de variables de estado finito se denominan modelos algebraicos.

En un modelo continuo, todas las variables asumen valores dentro de rangos biológicamente factibles. Los modelizadores suelen necesitar un conocimiento exhaustivo de las interacciones entre las variables, que puede incluir información detallada de los mecanismos de control reconocidos, las tasas de producción y degradación, las concentraciones mínimas y máximas biológicamente relevantes, etc. En un modelo algebraico solo se permiten valores de un conjunto finito. El caso especial de una *red booleana* solo permite dos estados, por ejemplo, 0 y 1, que representan la ausencia o la presencia de productos génicos en un modelo de regulación génica. A diferencia de los modelos continuos, la información necesaria para construir un modelo booleano solo requiere una comprensión conceptual de los vínculos causales de dependencia. Así, en general, los modelos continuos son cuantitativos, mientras que los modelos booleanos son de naturaleza cualitativa.

Históricamente, los modelos continuos han sido el tipo preferido de modelos matemáticos utilizados en biología. Este tipo de modelos dinámicos ha demostrado ser esencial para los problemas de ecología, epidemiología, fisiología y endocrinología, entre muchos otros. Los modelos booleanos se introdujeron por primera vez en la biología en 1969 para estudiar las propiedades dinámicas de redes reguladoras de genes [5]. Son apropiados en los casos en que la dinámica de la red está determinada por la lógica de las interacciones y no por una cinética muy ajustada, que a menudo puede ser desconocida. El estado de un gen puede describirse mediante una variable booleana que exprese que está activo (*on*, 1) o inactivo (*off*, 0) y, por tanto, que sus productos están presentes o ausentes. Además, las interacciones entre elementos pueden representarse mediante funciones booleanas que calculan el estado de expresión de un gen a partir de la activación de otros genes. El resultado es una red booleana.

Dado que en biología pueden estar presentes trazas de diversas sustancias en todo

momento, la "ausencia" suele significar concentraciones inferiores a un determinado valor umbral que separa las concentraciones más altas de la línea de base, y "presencia" se interpreta como concentraciones superiores a este umbral. Puede parecer que, dado que las concentraciones químicas abarcan un rango continuo de valores, elegir un único umbral de corte puede no ser apropiado ya que dos valores de concentración pueden estar muy cercanos numéricamente y que uno de ellos esté por encima del umbral y el otro por debajo. Aunque esto puede ser una preocupación legítima en general, rara vez se aplicaría a un modelo de regulación génica. Cuando el gen se expresa, la concentración de la proteína que produce sería miles de veces mayor que las cantidades mínimas que se dan cuando éste no lo hace, presentando así una clara distinción entre presente y ausente (1 o 0). A lo largo de este trabajo, asumiremos sin más que un valor booleano de 0 indica una concentración cercana al nivel basal, mientras que un valor booleano de 1 significa una concentración notablemente superior.

En los modelos booleanos, la evolución dinámica del sistema se describe mediante funciones booleanas definidas en términos de las variables del modelo y de los operadores lógicos AND (denotado por el símbolo $\wedge$), OR (denotado por el símbolo $\vee$) y NOT (denotado por una barra sobre la variable).

En el contexto de la modelización de la dinámica de redes, suele ser útil considerar la siguiente interpretación intuitiva para las operaciones AND y OR: si los componentes x e y del sistema influyen (controlan) a un tercer componente z, entonces $z = x \wedge y$ significa que x e y deben estar presentes simultáneamente (tener valores 1) para afectar a z; por tro lado, $z = x \vee y$ significa que x e y influyen en z de forma independiente y que z se ve afectado cuando x o y (o ambos) están presentes. En ausencia de paréntesis, la jerarquía es la siguiente: el operador lógico NOT tiene la mayor precedencia, seguido por el AND, y luego por el OR.

Para más detalles, se puede consultar [8].

A.1 Modelización de la dinámica de una red booleana: Funciones de transición

Supongamos que una red booleana contiene n variables booleanas del modelo (nodos) denotadas por $x_1, x_2, \ldots, x_n$. Cada una de estas n variables puede tomar un valor 0 o 1, lo que resulta en un conjunto de n-tuplas $V = \{0,1\}^n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in \{0,1\}, i = 1,2,\ldots,n\}$, que contiene 2^n elementos, que representan todos los estados posibles para las variables del modelo. El tiempo es discreto y solo puede tomar valores $t = 0,1,2,\ldots$ Los valores de los nodos $x_1, x_2, \ldots, x_n$ cambian con el tiempo y escribimos $x_i = x_i(t)$ para el valor de la variable x_i en el tiempo t. Así, en cada paso de tiempo t, el sistema está representado por una n-tupla binaria del conjunto V, donde cada componente representa el valor de la respectiva variable booleana en el tiempo t. Las reglas de transición entre los estados en cada paso de tiempo están descritas por n funciones $f_{x_i}, i = 1,2,\ldots,n$, una función para cada variable del modelo. La expresión booleana que define la función f_{x_i}, escrita en términos de las operaciones booleanas AND, OR y NOT, describe de qué manera los valores de las variables $x_1, x_2, \ldots, x_n$ en el momento t afectan al valor de la variable x_i en el momento $t + 1$. Así, para cualquier valor de $t = 0,1,2,\ldots$, la "actualización" del sistema para la variable x_i desde el tiempo t hasta el tiempo $t + 1$ viene determinada

por

$$x_i(t+1) = f_{x_i}(x_1(t), x_2(t), \ldots, x_n(t)),\ i = 1,2,\ldots,n$$

Las funciones f_{x_i}, $i = 1,2,\ldots,n$, son las *funciones de transición* del modelo. Las funciones (actualizaciones) que utilizaremos aquí son sincrónicas, lo que significa que todas las variables x_i se calculan primero para el tiempo t y luego se utilizan para evaluar las funciones f_{x_i}. Si escribimos $x = (x_1, x_2, \ldots, x_n)$ y $f(x) = (f_{x_1}(x), \ldots, f_{x_n}(x))$, el espacio de estados del modelo queda definido por el grafo dirigido $\{V, T\}$, donde el conjunto $T = \{(x, f(x)) \mid x \in V\}$ representa el conjunto de aristas.

Ejemplo 14. Supongamos tres variables booleanas x_1, x_2 y x_3, cuyas funciones de transición vienen dadas por

$$\begin{aligned}
x_1(t+1) &= f_{x_1}(x_1(t), x_2(t), x_3(t)) = x_2(t) \\
x_2(t+1) &= f_{x_2}(x_1(t), x_2(t), x_3(t)) = x_1(t) \vee x_3(t) \\
x_3(t+1) &= f_{x_3}(x_1(t), x_2(t), x_3(t)) = x_1(t) \wedge x_2(t) \vee x_3(t)
\end{aligned}$$

Entendiendo que las variables del lado derecho se evalúan siempre en el tiempo t y que las variables del lado izquierdo representan los valores en el tiempo $t+1$, podemos simplificar la notación eliminando t y $t+1$ de las ecuaciones:

$$\begin{aligned}
x_1 &= f_{x_1}(x_1, x_2, x_3) = x_2 \\
x_2 &= f_{x_2}(x_1, x_2, x_3) = x_1 \vee x_3 \\
x_3 &= f_{x_3}(x_1, x_2, x_3) = x_1 \wedge x_2 \vee x_3
\end{aligned}$$

Supongamos que en el momento $t = 0$, los valores de las variables son $x_1 = 0$, $x_2 = 0$ y $x_3 = 1$ (*condiciones iniciales*). Utilizando estos valores para evaluar las funciones de transición anteriores, obtenemos los valores de las variables en el momento $t = 1$:

$$\begin{aligned}
x_1 &= f_{x_1}(0,0,1) = 0 \\
x_2 &= f_{x_2}(0,0,1) = 0 \vee 1 = 1 \\
x_3 &= f_{x_3}(0,0,1) = 0 \wedge 0 \vee 1 = 1
\end{aligned}$$

Ahora tomamos los nuevos valores $x_1 = 0$, $x_2 = 1$ y $x_3 = 1$. Estos valores se utilizan para evaluar las funciones de transición f_{xi} de nuevo, produciendo, para el tiempo $t = 2$, los valores $x_1 = 1$, $x_2 = 1$ y $x_3 = 1$. Si se introducen de nuevo estos valores en las funciones, se obtienen los mismos valores $x_1 = 1$, $x_2 = 1$ y $x_3 = 1$, que corresponden a los valores de las variables en el tiempo $t = 3$. Así, para cualquier valor futuro de t, los valores de las variables del modelo seguirán siendo $x_1 = 1$, $x_2 = 1$ y $x_3 = 1$. Decimos que hemos calculado la trayectoria del estado $(0,0,1) : (0,0,1) \longrightarrow (0,1,1) \longrightarrow (1,1,1)$, y que $(1,1,1)$ es un **punto fijo** para la red booleana. Consideraciones similares demuestran que $(0,0,0)$ también es un punto fijo en este ejemplo. El uso de diferentes valores iniciales para las variables booleanas dará lugar a diferentes trayectorias. Por ejemplo, el estado inicial $(0,1,0)$ genera el siguiente patrón de repetición: $(0,1,0) \longrightarrow (1,0,0) \longrightarrow (0,1,0) \longrightarrow (1,0,0) \longrightarrow (0,1,0) \longrightarrow (1,0,0)$. Decimos que los estados $(0,1,0)$ y $(1,0,0)$ forman un *ciclo* de longitud dos. Los puntos fijos pueden considerarse ciclos de longitud uno. Como el sistema está compuesto por tres variables, cada una de las cuales puede tomar los valores 0 y 1, hay $2^3 = 8$ estados posibles para el sistema. Calculando las trayectorias que parten de cada uno de estos

ocho estados iniciales, se visualizan todas las posibles trayectorias de la red booleana, que contienen todas las transiciones posibles entre los ocho estados y, por lo tanto, forma el diagrama de transición del espacio de estados de la red booleana. Evidentemente, para un número mucho mayor de variables, el cálculo de las trayectorias a mano sería imposible y el uso de software adecuado es esencial.

En las aplicaciones biológicas, un punto fijo equivale a un *estado estacionario o estable* del sistema. Desde un punto de vista aplicado, nos interesa determinar si el sistema biológico alcanza un punto fijo, entra en un ciclo límite más largo o presenta algún otro comportamiento dadas las condiciones iniciales del sistema. Sin embargo, como el espacio de estados crece exponencialmente con el número de variables del sistema, determinar el espacio de estados mediante cálculos de fuerza bruta no es computacionalmente factible para ninguna red de tamaño razonable. En la siguiente sección discutimos un método alternativo para determinar los puntos fijos de una red mediante álgebra computacional utilizando las técnicas estudiadas en este proyecto.

Las funciones booleanas pueden representarse como funciones polinómicas con coeficientes en $\mathbb{Z}_2 = \{0,1\}$. En efecto, si x e y son variables booleanas, tenemos las siguientes conversiones:

$$\begin{aligned} x \wedge y &:= xy \\ x \vee y &:= x + y + xy \\ \overline{x} &:= x + 1 \end{aligned}$$

donde $\wedge$ representa la operación AND, $\vee$ la operación OR y $\overline{(\ )}$ la operación NOT. Obsérvese que la suma no corresponde a la función lógica OR, sino a la función OR exclusiva, denotada por XOR.

A.2 Operón lactosa

El **operón lac** es un grupo de secuencias codificantes con un promotor que presenta un gen regulador asociado con su propio promotor. Este gen regulador se transcribe continuamente, así que su producto proteico –el represor lac– siempre está presente. Fue el fruto de trabajos realizados con *Escherichia coli* (*E. coli*), una bacteria cuyo genoma actualmente ha sido secuenciado y su fisiología es bien conocida. El operón lac fue descrito por Francois Jacob y Jacques Monod en 1961 y les valió un premio Nobel en 1965. Los genes en el operón codifican las proteínas que permiten a las bacterias utilizar la lactosa (un disacárido) como fuente de energía. *E. coli* puede degradar la lactosa, aunque no es su azúcar preferido. Si la glucosa está presente en el medio, será el combustible utilizado, ya que ésta ofrece un mayor rendimiento energético que la lactosa. Sin embargo, si la lactosa es el único azúcar disponible, *E. coli* la empleará como fuente de energía. Para utilizar la lactosa, las bacterias deben expresar los genes del operón lac, que codifica enzimas claves para el consumo y el metabolismo de esta molécula. Sin embargo, con vista a una mayor eficiencia energética, *E. coli* debe expresar el operón lac solamente cuando exista lactosa en el medio y no tenga glucosa disponible.

¿Cómo se detectan los niveles de lactosa y de glucosa y cómo los cambios en estos niveles afectan la transcripción del operón lac? Dos proteínas reguladoras están involucradas:

- el represor lac, que actúa como un sensor de presencia de lactosa, y
- la proteína activadora por catabolito (CAP), que actúa como un sensor de los niveles de glucosa disponible.

Estas proteínas se fijan al ADN del operón lac y regulan su transcripción en base a los niveles de lactosa y de glucosa.

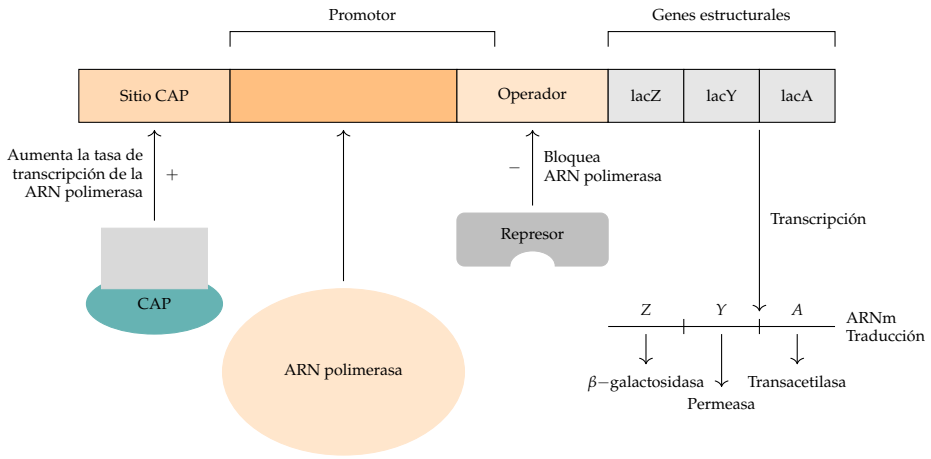


Figura 5: Operón lactosa

La región codificante del operón lac presenta tres genes: lacZ, lacY y lacA (genes estructurales). Estos genes se transcriben como un solo ARNm, bajo el control de un promotor. Los genes en el operón lac especifican las proteínas que ayudan a la célula a utilizar la lactosa. El gen lacZ codifica una enzima llamada β-galactosidasa, que es responsable de dividir la lactosa (un disacárido) en glucosa y galactosa (monosacáridos), fáciles de usar. El gen lacY codifica una proteína de membrana llamada lactosa permeasa, que permite que la célula importe la lactosa. El gen lacA codifica una enzima conocida como transacetilasa, que pega un grupo químico particular a moléculas objetivo. No es claro si esta enzima realmente juega un papel en la descomposición de la lactosa.

Además de los tres genes, el operón lac también contiene secuencias de ADN reguladoras, a las que se pueden unir proteínas reguladoras particulares y que controlan la transcripción del operón.

- El promotor es el sitio de unión de la ARN polimerasa, la enzima que realiza la transcripción.
- El operador es un sitio regulador negativo al que se une la proteína represora lac. Como el operador se encuentra antes del promotor, cuando se le ha unido el represor lac, éste impide que la ARN polimerasa se una al promotor y se inhibe la transcripción.
- El sitio de unión de CAP es un sitio regulador positivo al que se une la proteína CAP. Cuando se une CAP a este sitio, promueve la transcripción al ayudar a la ARN polimerasa a aumentar su tasa de transcripción.

Cuando la lactosa no está disponible, el represor lac se une al operador, y así evita la transcripción por la ARN polimerasa. Sin embargo, cuando la lactosa está presente en el medio, el represor lac pierde su capacidad de unirse al ADN. Se separa del operador, y despeja el camino para que la ARN polimerasa transcriba el operón.

Este cambio en el represor lac lo causa la alolactosa, un isómero de la lactosa. Cuando se dispone de lactosa, la β-galactosidasa genera algunas moléculas de alolactosa dentro de la bacteria. La alolactosa se une al represor lac y origina un cambio conformacional del represor provocando la separación del operador, lo cual permite que el ARN polimerasa transcriba el operón. La alolactosa actúa como inductor, una molécula que provoca la expresión de un gen o de un operón. El operón lac se considera un operón inducible porque generalmente está apagado (reprimido), pero se puede encender (activar) en presencia del inductor alolactosa. Resulta que la ARN polimerasa sola presenta una tasa de transcripción pobre. Sin ambargo, si CAP se une a la región del ADN (región CAP) localizada justo antes del promotor del operón lac, provocará un aumento muy notable de dicha tasa de transcripción.

CAP no siempre es activa (capaz de unirse al ADN). Su actividad la regula una molécula pequeña llamada AMP cíclico (AMPc). AMPc es una "señal de hambre" que fabrica *E. coli* cuando los niveles de glucosa son bajos. AMPc se une a CAP, cambia su forma y la hace capaz de unirse al ADN y provocará la transcripción. Sin AMPc, CAP no puede unirse al ADN y es inactiva. CAP solamente está activa cuando los niveles de glucosa son bajos, que inducirá niveles de AMPc altos. Así, el operón lac solo puede transcribirse en altos niveles cuando no hay glucosa. Esta estrategia asegura que las bacterias solamente enciendan el operón lac y empiecen a usar la lactosa después de que hayan utilizado toda la fuente de energía preferida, la glucosa.

En definitiva, el operón lac se expresará en niveles altos si se cumplen dos condiciones:

- La lactosa debe estar disponible: provoca la inhibición del represor y que el operón se transcriba.
- La glucosa no debe estar disponible: promueve que CAP estimule la transcripción.

Ambas condiciones llevan a una elevada tasa de transcripción del operón lac y a la producción de las enzimas necesarias para la utilización de la lactosa.

En resumen, se tiene:

- Glucosa presente, lactosa ausente: no ocurre la transcripción del operón lac. Eso es porque el represor lac permanece unido al operador e impide la transcripción por la ARN polimerasa. Además, los niveles de AMPc son bajos porque los niveles de glucosa son altos, así que CAP está inactiva y no puede unirse al ADN.
- Glucosa presente, lactosa presente: se da la transcripción del operón lac a un nivel bajo. El represor lac es liberado del operador porque el inductor (alolactosa) está presente. Los niveles de AMPc, sin embargo, son bajos porque hay glucosa. Entonces, CAP permanece inactiva y no puede unirse al ADN, así que la transcripción del operón solo ocurre a un nivel bajo.
- Glucosa ausente, lactosa ausente: no ocurre transcripción del operón lac. Los niveles de AMPc son altos porque los niveles de glucosa son bajos, así que CAP está activa y

estará unida al ADN. Sin embargo, el represor lac también estará unido al operador (debido a la ausencia de alolactosa), y actúa como barrera a la ARN polimerasa y previene la transcripción.

- Glucosa ausente, lactosa presente: ocurre una fuerte transcripción del operón lac. El represor lac es liberado del operador porque el inductor (alolactosa) está presente. Los niveles de AMPc son altos porque no hay glucosa, así que CAP está activa y unida al ADN. CAP provoca una intensificación de la tasa de transcripción por parte de la ARN polimerasa.