Manzanas, piñas, plátanos y Curvas Elípticas

Pedro Montero

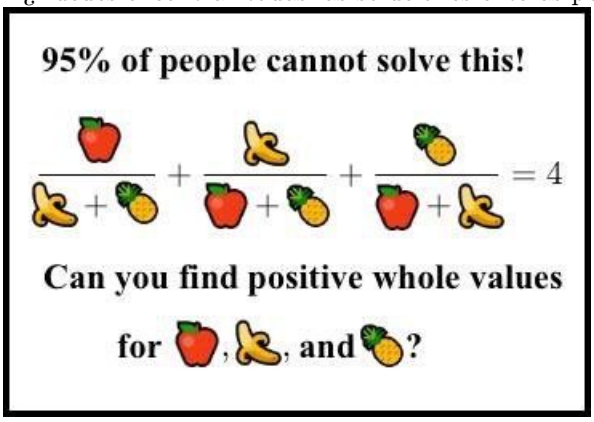
Departamento de Matemática, Universidad Técnica Federico Santa María

El objetivo de esta nota es resolver un **desafío** sobre álgebra no-lineal planteado en el curso MAT210. Tal como será expuesto más abajo, los métodos más elementales (de álgebra básica) no son suficientes para resolver el problema. De hecho, el objetivo de plantearlo en el curso fue tratar de que las y los estudiantes se den cuenta de que no se puede llegar muy lejos haciendo símplemente manipulaciones algebraicas y que es necesario usar herramientas adicionales. Así, el verdadero objetivo de esta nota es introducir y motivar algunas nociones importantes que serán necesarias para resolverlo.

Valparaíso, Agosto 2020

Consideremos el problema planteado en la siguiente imagen:

Figura 1: ¿Puedes encontrar todas las soluciones enteras positivas?



Estamos buscando soluciones enteras positivas (i.e., en $\mathbb{N}^{\geq 1}$) de la ecuación

$$\text{(E)} \qquad \frac{a}{b+c}+\frac{b}{a+c}+\frac{c}{a+b}=4$$

Este tipo de ecuaciones en varias variables donde buscamos soluciones **enteras** son llamadas **ecuaciones diofánticas**[1], y veremos enseguida que la restricción de buscar sólamente soluciones positivas hace que las cosas sean bastante más difíciles.

1. Ecuaciones homogéneas

Recordemos que un polinomio en varias variables $P(X_1,\ldots,X_n)$ es **homogéneo** de grado $d\in\mathbb{N}$ si para toda constante λ se tiene que

$$P(\lambda X_1,\ldots,\lambda X_n)=\lambda^d P(X_1,\ldots,X_n).$$

Al considerar **funciones racionales homogeneas** $f(X_1,\ldots,X_n)$ (i.e., cocientes de polinomios homogéneos) podemos extender dicha definición a $d\in\mathbb{Z}$ considerando constantes $\lambda\neq 0$. Así, la función

$$f(a,b,c)=\frac{a}{b+c}+\frac{b}{a+c}+\frac{c}{a+b}$$

es una función racional homogénea de grado 0. En particular, si (a_0,b_0,c_0) es una solución de (E), entonces $(\lambda a_0,\lambda b_0,\lambda c_0)$ también es solución de (E).

[1] En honor al matemático griego Diofantos de Alejandría (siglo III).

2. Interludio geométrico: espacio proyectivo

Consideremos el $\mathbb{Q}$-espacio vectorial $\mathbb{Q}^{n+1}$ y consideremos el **espacio proyectivo** $\mathbb{P}^n(\mathbb{Q})$ dado por el conjunto cuyos elementos son las rectas en $\mathbb{Q}^{n+1}$. Explícitamente,

$$\mathbb{P}^n(\mathbb{Q}) = \{L \subseteq \mathbb{Q}^{n+1} \mid \dim_{\mathbb{Q}}(L) = 1\}.$$

Veámos rápidamente porqué escogimos la notación $\mathbb{P}^n(\mathbb{Q})$ (en lugar de $\mathbb{P}^{n+1}(\mathbb{Q})$).

Importante: Esto será clave para darnos cuenta que la ecuación (E) depende realmente de 2 parámetros (o intuitivamente, define un objeto geométrico de dimensión 1: corresponde a una ecuación dentro de un espacio de dimensión 2) y **no** de 3 parámetros (las variables a, b, c):

Toda recta en $\mathbb{Q}^{n+1}$ puede ser generada por un vector director $v = (x_0, \ldots, x_n) \in \mathbb{Q}^{n+1}$, donde al menos una de las coordenadas x_i es no-nula[2]. Más aún, si $\lambda \in \mathbb{Q}$ es un escalar no-nulo, entonces v y λv generan la misma recta, y luego definen el mismo punto en $\mathbb{P}^n(\mathbb{Q})$. Esto último lo escribiremos más cómodamente de la manera siguiente:

$$[x_0, \ldots, x_n] = [\lambda x_0, \ldots, \lambda x_n] \text{ en } \mathbb{P}^n(\mathbb{Q}).$$

En otras palabras, utilizamos paréntesis cuadrados $[x_0, \ldots, x_n]$ para denotar *cualquier* vector director de la recta generada por $v = (x_0, \ldots, x_n)$. Así, podemos identificar la recta L con el conjunto de vectores directores $[x_0, \ldots, x_n]$.

Gracias a la discusión anterior, podemos pensar los puntos en $\mathbb{P}^n(\mathbb{Q})$ como elementos de la forma $[x_0, \ldots, x_n]$ donde al menos un $x_i \neq 0$ y donde $[x_0, \ldots, x_n] = [\lambda x_0, \ldots, \lambda x_n]$ para todo $\lambda \neq 0$ constante. Veamos que podemos cubrir a $\mathbb{P}^n(\mathbb{Q})$ por conjuntos U_i que pueden ser identificados con $\mathbb{Q}^n$:

Para todo índice $i \in \{0, \ldots, n\}$ definimos el conjunto

$$U_i := \{[x_0, \ldots, x_n] \in \mathbb{P}^n(\mathbb{Q}) \mid x_i \neq 0\}.$$

Por ejemplo, para $i = 0$ tenemos que un elemento de U_0 es de la forma $[x_0, \ldots, x_n] \in \mathbb{P}^n(\mathbb{Q})$ donde $x_0 \neq 0$. Es importante notar que, dado que $x_0 \neq 0$, entonces ya se cumple la condición que *al menos una* coordenada x_i es no-nula, por lo cual el resto de las coordenadas $(x_1, \ldots, x_n) \in \mathbb{Q}^n$ puede tomar *cualquier* valor. Más aún, dado que $[x_0, \ldots, x_n] = [\lambda x_0, \ldots, \lambda x_n]$ para todo $\lambda \neq 0$, podemos escoger $\lambda = \frac{1}{x_0}$ y obtenemos así que

$$[x_0, x_1, \ldots, x_n] = [1, y_1, \ldots, y_n]$$

donde $(y_1, \ldots, y_n) = (\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}) \in \mathbb{Q}^n$ es un vector arbitrario en $\mathbb{Q}^n$.

Así, podemos identificar los elementos de U_0 con vectores de $\mathbb{Q}^n$ y más generalmente (mediante el mismo cálculo) podemos encontrar biyecciones entre U_i y $\mathbb{Q}^n$ para todo $i \in \{0, \ldots, n\}$. Por lo anterior, decimos que el conjunto $\mathbb{P}^n(\mathbb{Q})$ de rectas en $\mathbb{Q}^{n+1}$ es el **espacio proyectivo de dimensión** n sobre $\mathbb{Q}$.

3. Ecuación cúbica

La conexión entre las primeras dos secciones es la siguiente: Si $P(X_1, \ldots, X_n)$ es un polinomio homogéneo de grado $d \geq 1$ en n variables, entonces el hecho que

$$P(\lambda X_1, \ldots, \lambda X_n) = \lambda^d P(X_1, \ldots, X_n)$$

para todo $\lambda \neq 0$ implica que el conjunto

$$X = \{[x_1, \ldots, x_n] \in \mathbb{P}^{n-1}(\mathbb{Q}) \mid P(x_1, \ldots, x_n) = 0\}$$

está **bien definido**. En efecto, cualquier otro vector director de la recta $[x_1, \ldots, x_n]$ será un múltiplo no-nulo de $(x_1, \ldots, x_n)$, y por ende $P(x_1, \ldots, x_n) = 0$ si y sólo si $P(\lambda x_1, \ldots, \lambda x_n) = 0$, para cualquier $\lambda \neq 0$.

[2]Notar que comenzamos la numeración desde 0, por lo que hay $n + 1$ coordenadas.

Volvamos a nuestra ecuación original

$$\text{(E)} \qquad \frac{a}{b+c} + \frac{b}{a+c} + \frac{c}{a+b} = 4,$$

la cual ya vimos que estaba definida por una función racional $f(a,b,c)$ homogénea de grado 0. Veamos que también podemos pensarla como una ecuación homogénea polinomial de grado 3:

Dado que estamos buscando soluciones (a,b,c) enteras positivas, podemos multiplicar la ecuación (E) por el denominador común $(a+b)(a+c)(b+c)$ para obtener la ecuación (E') siguiente:

$$\text{(E}') \qquad a(a+b)(a+c) + b(b+a)(b+c) + c(c+a)(c+b) = 4(a+b)(a+c)(b+c).$$

Así, expandiendo y reordenando los términos, obtenemos el polinomio

$$P(a,b,c) = a^3 + b^3 + c^3 - 3(a^2b + ab^2 + a^2c + ac^2 + b^2c + bc^2) - 5abc,$$

que es homogéneo de grado 3 en las variables a, b, c. Así, la ecuación (E) se reduce a estudiar la **ecuación cúbica** $P(a,b,c) = 0$, o equivalentemente estudiar el conjunto

$$\mathcal{E} = \{[a,b,c] \in \mathbb{P}^2(\mathbb{Q}) \mid P(a,b,c) = 0\}.$$

4. Curvas elípticas

En palabras simples, una **curva elíptica** es (el conjunto de soluciones racionales de) una ecuación de la forma

$$(C) \qquad y^2 = x^3 + Ax + B,$$

donde $A, B \in \mathbb{Q}$ son racionales, y donde usualmente pedimos que (C) posea al menos una solución racional $(x,y) \in \mathbb{Q}^2$ y además pedimos que C sea **suave**.

El hecho que la curva elíptica C sea suave puede verificarse gráficamente dibujando las soluciones de la ecuación $y^2 = x^3 + Ax + B$ en $\mathbb{R}^2$, y observando que la curva obtenida no tenga *puntas* (o más formalmente, que cada punto $p \in C$ sólo posea una recta tangente afín T_pC pasando por p).

Una forma más algebraica de determinar si una curva definida por una ecuación cúbica es suave es utilizar el **discriminante**:

Por definición, el **discriminante** de un polinomio $P(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0$ de grado n con coeficientes complejos[3] es un polinomio $\Delta(a_0, \ldots, a_n)$ en los coeficientes de P tal que

$$\Delta = 0 \Leftrightarrow P \text{ tiene raíces múltiples.}$$

Por ejemplo, sabemos que si $P(X) = aX^2 + bX + c$, entonces $\Delta = b^2 - 4ac$. En general[4], el discriminante puede ser escrito como el determinante de cierta matriz de tamaño $(2n-1) \times (2n-1)$.

El discriminante de un polinomio cúbico $P(X) = aX^3 + bX^2 + cX + d$ está dado por

$$\Delta = b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd.$$

En particular, el discriminante del polinomio $P(X) = X^3 + AX + B$ es $\Delta = -4A^3 - 27B^2$. Así, se puede probar[5] que C es suave si y sólo si $\Delta \neq 0$.

Finalmente, notamos que si restringimos nuestro conjunto

$$\mathcal{E} = \{[a,b,c] \in \mathbb{P}^2(\mathbb{Q}) \mid P(a,b,c) = 0\}.$$

[3] Necesitamos mirar raíces, por lo que es útil usar el hecho que $\mathbb{C}$ es algebraicamente cerrado.

[4] Ver por ejemplo Wikipedia.

[5] En aplicaciones más avanzadas de curvas elípticas usualmente se define $\Delta_C = -16(4A^3 + 27B^2)$, aunque este factor no afecta el hecho que C sea suave o no.

al sub-conjunto $U_2 = \{[a, b, c] \in \mathbb{P}^2(\mathbb{Q}) \mid c \neq 0\}$ entonces siempre podemos suponer que $c = 1$ (reescalando si fuera necesario) y reducirnos a estudiar

$$\{(x, y) \in \mathbb{Q}^2 \mid P(x, y, 1) = 0\},$$

donde $P(x, y, 1) := Q(x, y) = 1 + x^3 + y^3 - 3x - 3x^2 - 3y - 3y^2 - 3x^2y - 3xy^2 - 5xy$ polinomio cúbico. De manera similar para $U_0 \subseteq \mathbb{P}^2(\mathbb{Q})$ y $U_1 \subseteq \mathbb{P}^2(\mathbb{Q})$.

Hecho (sin demostración): Existen algoritmos[6] para transformar una ecuación cúbica en una ecuación de la forma $y^2 = x^3 + Ax + B$, llamada la **forma de Weierstrass** del polinomio cúbico. Del mismo modo, una ecuación cúbica puede ser llevada a la forma $y^2 = x^2 + Ax^2 + Bx + C$, que es llamada la **forma de Weierstrass larga** del polinomio cúbico, y que en ocasiones es más simple que la forma de Weierstrass.

Por ejemplo, al aplicar dichos algoritmos a nuestra ecuación cúbica obtenemos la forma de Weierstrass dada por $y^2 = x^3 + A_0x + B_0$ con $A_0 = -11209/48$ y $B_0 = 1185157/864$. Sin embargo, las transformaciones *birracionales*[7] siguientes

$$x = \frac{-28(a + b + 2c)}{6a + 6b - c} \quad y = \frac{364(a - b)}{6a + 6b - c}$$

transforman el polinomio P a su forma de Weierstrass larga

$$y^2 = x^3 + 109x^2 + 224x,$$

que tiene la ventaja de tener coeficientes enteros. Luego, debemos estudiar

$$C = \{(x, y) \in \mathbb{Q}^2 \mid y^2 = x^3 + 109x^2 + 224x\}.$$

Primero que todo, notamos que el discriminante de $P(X) = X^3 + 109X^2 + 224X$ es $\Delta = 551183360 \neq 0$. Equivalentemente, al graficar

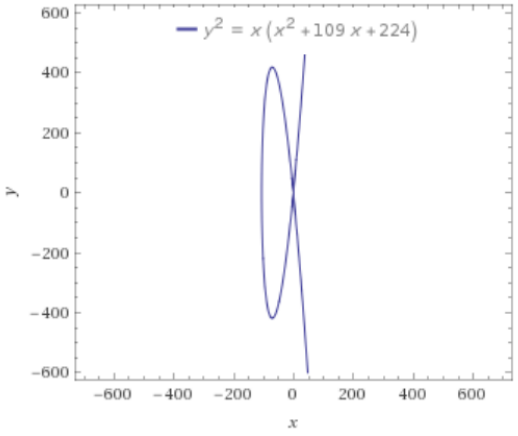


Figura 2: Curva elíptica C.

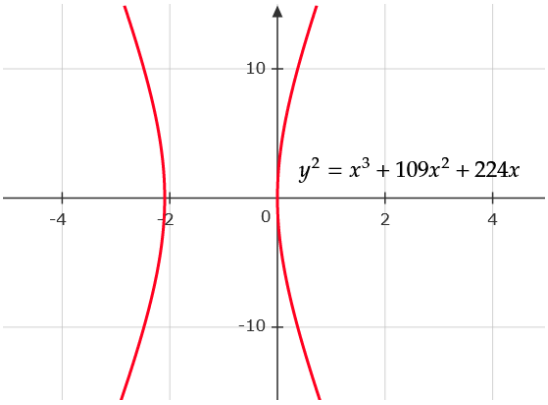


Figura 3: Curva elíptica C cerca del origen.

obtenemos curvas *sin puntas*, por lo que C es una curva suave. Notemos también que nuestra ecuación cúbica (E') admite al menos una solución racional: si $b = 1$ y $c = 0$ entonces obtenemos

$$a^3 - 3a^2 - 3a + 1 = 0,$$

de donde $a = -1$ es una raíz. Así, C es una curva elíptica.

Finalmente, observamos que dada cualquier solución $(x, y) \in \mathbb{Q}^2$ de la ecuación $y^2 = x^3 + 109x^2 + 224x$, podemos obtener una solución $(a, b, c) \in \mathbb{Q}^3$ de la ecuación (E') mediante el cambio de variable *birracional*

$$a = \frac{56 - x + y}{56 - 14x} \quad b = \frac{56 - x - y}{56 - 14x} \quad c = \frac{-28 - 6x}{28 - 7x}.$$

[6] Ver aquí, usando SAGE.

[7] Un concepto central en geometría, cuyo nombre proviene del hecho que son funciones racionales en cada variable.

5. Un poco de grupos abelianos y criptografía

Una de las propiedades más importantes de las curvas elípticas es que ellas forman un **grupo abeliano**. Recordemos que un grupo es un conjunto no-vacío G junto con una operación interna

$$G \times G \to G,\ (g_1, g_2) \mapsto g_1 \star g_2$$

que es asociativa, que admite un elemento neutro e, y que todo elemento admite un único inverso g^{-1} tal que $g \star g^{-1} = g^{-1} \star g = e$. Además, decimos que es **abeliano** si $g_1 \star g_2 = g_2 \star g_1$ para todos $g_1, g_2 \in G$. Típicamente, cuando consideramos grupos abelianos escribirmos la operación $g_1 \star g_2$ como una suma $g_1 + g_2$, el neutro es denotado por 0, y el inverso de g se denota $-g$.

Los grupos abelianos son muy importantes en **criptografía**, por lo que antes de hablar de curvas elípticas veamos el caso del **algoritmo RSA**[8]:

Imaginemos que tenemos una carta con un mensaje secreto, por ejemplo "`CAFE`", que queremos enviar a nuestra mejor amiga (que vive muy lejos) y que no queremos que nadie sepa[9]. Una primera posibilidad es que tu amiga te envíe una **caja** y una **llave**, para que tu puedas cerrarla.

El problema es que si nuestra amiga tuviera muchas cajas iguales con llaves iguales, y se las diera a sus amigos para guardar secretos, entonces una persona podría robar nuestra caja y usar su llave para abrir la nuestra, y describir nuestro mensaje `CAFE`. Para resolver esto, lo que hace nuestra mejor amiga es darnos **cajas**, pero en vez de darnos la llave nos da un **candado** para cerrarla. Así, una vez que cerramos nuestra caja (con nuestro mensaje adentro) con nuestro candado, seremos incapacez de abrirla.

Veamos un ejemplo numérico para ilustrar cómo funciona en la práctica:

Nuestra mejor amiga muestra a todo el mundo dos números **públicos**: por ejemplo, $e = 3$ y $n = 10$. Nuestra amiga además posee un número **secreto** N, que sólo conoce ella. La idea ahora será convertir nuestro mensaje `CAFE` en números:

La forma más simple de hacerlo es seguir el orden alfabético[10]: `A` corresponde a 1, `B` corresponde a 2,..., `Z` corresponde a 26. Luego, nuestro mensaje `CAFE` corresponde la secuencia de cinco números 3 1 6 5, que escribiremos como las coordenadas de un vector $(3, 1, 6, 5)$ por comodidad.

La idea ahora será esconder este mensaje: algunos códigos van a adicionar a cada componente del mensaje $(3, 1, 6, 5)$ una constante (e.g. sumar 3), otros van a multiplicar por una constante (e.g. multiplicar por 3), etc. Lo que haremos en este ejemplo (RSA) es usar el número público (exponente) $e = 3$ para elevar a la potencia 3 y obtener de este modo $(27, 1, 216, 125)$.

La segunda parte del proceso es usar el número público $n = 10$ y considerar enteros *módulo 10*, es decir, dividimos cada coordenada por 10 sucesivamente hasta obtener un resto entre 0 y 9. En nuestro ejemplo, obtenemos

$$(7, 1, 6, 5).$$

Finalmente, nuestra mejor amiga posee además su número **secreto** N, que sólo ella conoce (y que en este ejemplo, ella calcula que $N = 3$). Así, cuando nuestra amiga reciba el mensaje encriptado $(7, 1, 6, 5)$ ella va a elevar a la potencia $N = 3$ cada componente para obtener $(343, 1, 216, 125)$, y luego considerara el vector módulo $n = 10$ para obtener

$$(3, 1, 6, 5)$$

y recibir de este modo el mensaje: `CAFE`.

La idea detrás del algoritmo es escoger el número público $n = 10 = 2 \cdot 5$ como el *producto de dos números* primos (en este caso, 2 y 5, pero en la práctica se usan primos gigantes). Por otro lado, una etapa fundamental en el proceso de descifrado (i.e., cómo nuestra amiga calcula N) es el **pequeño teorema de Fermat**, que afirma para todo número primo p y $x \in \mathbb{Z}$ se tiene que

$$x^p \equiv x \mod p,$$

o equivalentemente $x^p - x$ es un múltiplo del número primo p, así como el **teorema chino del resto**. Estos tópicos forman parte de los contenidos del curso de Estructuras Algebraicas (MAT214).

[8] En honor a Rivest, Shamir y Adleman.

[9] Para la gente menos amistosa, pueden pensar también en una transacción privada que quieren comunicarle a su banco.

[10] Usaremos el alfabeto inglés de 26 letras para simplificar.

6. Las curvas elípticas son grupos abelianos

Tal como adelantemos en la sección anterior, una de las principales propiedades de las curvas elípticas es el hecho que (los conjuntos de soluciones racionales) forman un grupo abeliano. Sin embargo, la forma de describir la operación interna es más complicada que en el caso de los enteros $\mathbb{Z}$ o que $\mathbb{Z}/n\mathbb{Z}$, los enteros módulo n.

La idea geométrica es la siguiente: si tenemos dos soluciones $P = (x_1, y_1) \in \mathbb{Q}^2$ y $Q = (x_2, y_2) \in \mathbb{Q}^2$ de nuestra ecuación $y^2 = x^3 + Ax^2 + Bx + C$ entonces calculamos $P + Q$ de la manera siguiente:

1. Si $P \neq Q$ son puntos distintos, entonces trazamos la recta que pasa por P y Q. Dado que nuestra curva está dada por una ecuación de grado 3, esta recta intersecta la curva elíptica en un tercer punto. Luego, **definimos** la suma $P+Q$ como el punto obtenido al considerar la simetría ortogonal de este último punto respecto al eje x.

2. Si $P = Q$ son el mismo punto, entonces consideramos la recta tangente[11] a la curva elíptica pasando por el punto P. Tal como antes, la recta tangente intersecta la curva elíptica en un tercer punto. Luego, **definimos** la suma $P+P := 2P$ como el punto obtenido al considerar la simetría ortogonal de este último punto respecto al eje x.

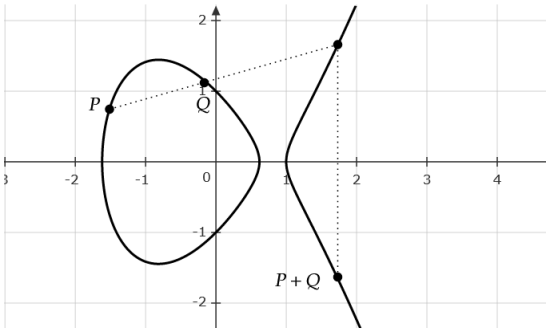


Figura 4: Suma de dos puntos distintos.

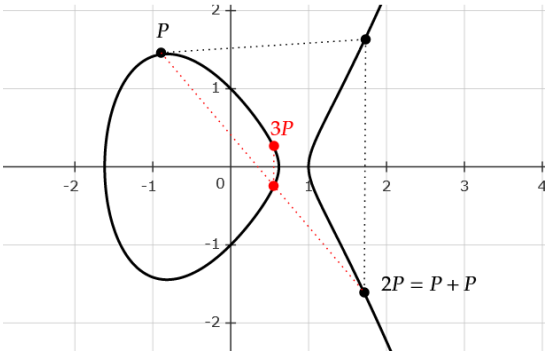


Figura 5: Suma de un mismo punto.

Usando esta construcción geométrica, no es muy difícil obtener fórmulas explícitas[12] para obtener las coordenadas de $P + Q$. El neutro respecto a esta ley de grupo es un poco más complicado de ver, pero ya hemos introducido el lenguaje necesario: la idea es que la curva $y^2 = x^3 + Ax^2 + Bx + C$ en $\mathbb{Q}^2$ puede verse como un sub-conjunto de

$$U_2 = \{[x, y, z] \in \mathbb{P}^2(\mathbb{Q}) \mid z \neq 0\}$$

en el plano proyectivo $\mathbb{P}^2(\mathbb{Q})$. Así, si **homogeneizamos** dicha ecuación, agregando la variable z, obtenemos la ecuación (E") siguiente:

$$y^2 z = x^3 + Ax^2 z + Bxz^2 + Cz^3,$$

una ecuación cúbica homogénea que define una curva en $\mathbb{P}^2(\mathbb{Q})$, y que al restringirse a U_2 se reduce a nuestra ecuación original (haciendo $z = 1$). Finalmente, el punto que hace las veces del neutro aditivo es un **punto al infinito**[13] que se obtiene al considerar $z = 0$ en $\mathbb{P}^2(\mathbb{Q})$: la ecuación (E") se reduce a $x = 0$ y luego $\mathcal{O} = [0 : 1 : 0]$ es el neutro de la ley de grupo.

La criptografía basada en curvas elípticas utiliza esta estructura de grupo abeliano. Debido a que la ley de grupo es mucho más sofisticada, los mensajes privados son mucho más difíciles de descifrar y es por esto que las curvas elípticas son muy usadas en transacciones bancarias[14] (e.g. bitcoins[15]).

[11] Intuitivamente, si consideramos el límite cuando Q tiende a P entonces la recta secante se convierte en tangente.

[12] Ver aquí las fórmulas explícitas y aquí la implementación en SAGE.

[13] Por definición, dicho punto no intersecta U_2. Luego, es un punto que agregamos a la curva vista en $\mathbb{Q}^2$.

[14] Algunas empresas y bancos usan la curva elíptica NIST P-256 dada por $y^2 = x^3 - 3x + b$. Además, para aumentar la complejidad de descifrado, consideran soluciones en $\mathbb{Z}/p\mathbb{Z}$ para $p = 2^{256} - 2^{224} + 2^{192} + 2^{96} - 1$. Aquí, el parámetro b está dado por $b = 41058363725152142129326129780047268409114441015993725554835256314039467401291$.

[15] Que usan la curva elíptica Secp256k1 dada por $y^2 = x^3 + 7$, y donde se consideran soluciones módulo $p = 2^{256} - 2^{32} - 977$.

7. Soluciones particulares

El punto de partida es considerar un punto racional $P_0 = (x_0, y_0) \in \mathbb{Q}^2$ en la curva C. En este caso, notamos que $P_0 = (-100, 260)$ es un punto de C. En efecto, basta reemplazar y ver que la ecuación $y^2 = x^3 + 109x^2 + 224x$ se satisface. Notamos que los valores $(a, b, c) \in \mathbb{Q}^3$ correspondientes son

$$a = \frac{2}{7} \quad b = -\frac{1}{14} \quad c = \frac{11}{14},$$

y que, al multiplicar por el denominador común, obtenemos los enteros

$$a = 4 \quad b = -1 \quad c = 11.$$

Notar que $\frac{4}{-1+11} + \frac{-1}{4+11} + \frac{11}{4-1} = 4$, por lo que $(a_0, b_0, c_0) = (4, -1, 11)$ es una solución entera de nuestra ecuación diofántica original (E). Sin embargo, no es una solución entera **positiva**.

Veremos en la próxima sección que dicho punto P_0 no es fácil de encontrar a priori, pero por ahora enfoquémonos en producir más soluciones a partir de la solución P_0:

Una vez que tenemos el punto P_0 en la curva elíptica C podemos comenzar a considerar **iteraciones** del punto P_0 utilizando la ley de grupo o, en palabras simples, podemos calcular mP_0 para $m \in \mathbb{N}$ y producir así nuevas soluciones:

$$2P_0 = P_0 + P_0 = \left(\frac{8836}{25}, -\frac{950716}{125}\right) \in \mathbb{Q}^2,$$

cuya solución de (E) correspondiente es $(a, b, c) = (9499, -8784, 5165)$, la cual **no** es una solución entera positiva. Si continuamos de este modo podemos seguir produciendo soluciones enteras $3P_0$, $4P_0$,...,$8P_0$ que **no** son positivas. Sin embargo, tenemos que $9P_0 = \left(\frac{p}{q}, \frac{r}{s}\right) \in \mathbb{Q}^2$ con

$$p = -662023684042295852648424098838788747074536766450382 25,$$
$$q = 135144002927162885120709079450029433526925780004069 21,$$
$$r = 588008351573080833073767517273471813300856728502967303518717487133079887006112 10,$$
$$s = 157106866859797843455636470729189626883808694543003132219675439042028040734646 9,$$

tiene asociada la solución entera (a, b, c) de (E) dada por

$$a = 154476802108746166441951315019919837485664325669565431700026634898253202035277999,$$
$$b = 368751317941299998271978115652254748254929799689719709962831374716372246340555 79,$$
$$c = 437361267792869725786125260237139015281653755816161361862143799337842346777203 6,$$

que es efectivamente una solución **entera positiva**. De hecho, es posible probar (utilizando la *teoría de alturas* en curvas elípticas) que es la solución entera positiva *más pequeña* de la ecuación (E).

8. ¿Cómo encontrar soluciones?

Volvamos a hablar un poco sobre la teoría general de curvas elípticas:

A priori, podria ocurrir que $mP_0 = 0$ para cierto $m \in \mathbb{N}^{\geq 1}$ entero positivo[16]. Los puntos de C para los cuales existe m tal que $mP = 0$ son llamados **puntos de torsión** y se denotan $T(C)$. Un resultado importante de Mazur (1977) dice que el cardinal de $T(C)$ puede ser 0,1,2,..., 10 o 12 (no puede ser 11).

Por otra parte, tenemos los puntos tales que mP es distinto de 0 para todo m positivo, en cuyo caso decimos que P pertence a $F(C)$, la parte **libre de torsion** de C. Un teorema importante de Mordell (1922) afirma que la parte libre de torsión $F(C)$ está generada por r puntos $P_1, \ldots, P_r$ independientes entre si. Así, que cualquier punto P en la parte libre de torsión de C puede escribirse como una combinación lineal entera

$$P = a_1P_1 + \ldots + a_rP_r,$$

[16] Pensar en los enteros módulo n: n tiene resto 0 al dividir por n, por lo que $mn = 0$ en $\mathbb{Z}/n\mathbb{Z}$ para todo m positivo.

donde $a_1, \ldots, a_r \in \mathbb{Z}$. En otras palabras, $F(C) \cong \mathbb{Z}^r$ para cierto $r \in \mathbb{N}$. El entero r es llamado el **rango** de la curva elíptica C y es bastante misterioso en general: se conjetura que rangos arbitrariamente grandes deberian poder alcanzarse, sin embargo el rango más grande que se ha calculado hasta el momento es $r = 28$. Este último record fue obtenido el año 2006 por Elkies al considerar la curva elíptica

$$\begin{aligned} y^2 + xy + y = \; & x^2 - x^2 \\ & - 2006776241557552658503320820933854275093023031217895650 2x \\ & + 3448161179503055646703298569039072037485594435931918036126600829 \\ & \quad 6291939448732243429. \end{aligned}$$

En el otro extremo, la curva elíptica más simple posible $y^2 = x^3 - x$ posee sólamente 4 soluciones (contando el punto al infinito) $\{\mathcal{O}, (0,0), (1,0), (-1,0)\}$, por lo que en este caso $|T(C)| = 4$ y $r = 0$.

Volviendo a nuestra curva eliptica C dada por $y^2 = x^3 + 109x^2 + 224x$, las preguntas que tenemos que considerar ahora son: ¿cuál es el rango de C?, ¿posee puntos de torsión?. Si la curva no tiene puntos de torsión entonces iterando nuestro punto P_0 nunca llegaremos al neutro $\mathcal{O}$, por lo que tenemos chances en tal caso de producir infinitas soluciones de (E). Por otro lado, si determinamos el rango y los generadores de la parte libre de torsión $F(C)$ tendremos también una forma de construir soluciones generales a partir de las soluciones particulares $P_1, \ldots, P_r$ (los generadores).

Los algorítmos para determinar el rango, generadores y la torsión de una curva elíptica son no-triviales y hay mucha teoría detrás. Sin embargo, muchos de ellos están implementados en SAGE[17] y en la práctica podemos hacer cálculos para curvas elípticas explícitas. En nuestro caso, obtenemos que la parte libre de torsión $F(C) \cong \mathbb{Z}$ es de rango 1, generada por el punto $P_0 = (-100, 260)$ que mencionamos en la sección anterior, y se tiene que la parte de torsión está dada por los 6 puntos

$$T(C) = \{\mathcal{O}, (0,0), (4,-52), (4,52), (56,-728), (56,728)\}.$$

Así, obtenemos que **todas** las soluciones de la ecuación (E) pueden obtenerse utilizando múltiplos del punto P_0.

Comentario final: Este texto está basado en el artículo "*An unusual cubic representation problem*", publicado en 2014 por A. Bremner y A. MacLeod en *Annales Mathematicae et Informaticae*. Tal como exponen los autores, si cambiamos el número 4 en la ecuación (E) por 178 y miramos las soluciones enteras positivas de

$$\frac{a}{b+c} + \frac{b}{a+c} + \frac{c}{a+b} = 178,$$

entonces, en lugar de obtener una solución minimal de *sólo* 80 dígitos, en este caso se obtienen soluciones de 398.605.460 de dígitos. Por otro lado, si reemplazamos 4 por 896, entonces se obtienen soluciones de más de un millón de millones de dígitos.

Esto último tiene una razón profunda detrás y es el hecho que la respuesta al **décimo problema de Hilbert**[18], que pregunta sobre si es posible o no determinar un algoritmo para encontrar las soluciones enteras de una ecuación diofántica, es *negativa*. Esto último fue probado por Yu. Matiyásevich en 1970.

Lectura recomendada: Las y los estudiantes que se interesen en continuar explorando el uso de curvas elípticas para resolver ecuaciones diofánticas, así como leer sobre algunos problemas abiertos, pueden leer (aquí) el artículo de A. MacLeod "*Elliptic Curves in Recreational Number Theory*" (2016).

[17] En este caso, utilizamos el comando `C=EllipticCurve([0,109,0,224,0])` para definir nuestra curva elíptica C, y luego los comandos `C.rank()`, `E.gens()` y `E.torsion_points()` entregan el rango, generadores y la torsión, respectivamente. Otra alternativa, es chequear si nuestra curva elíptica estará entre las 3.064.705 posibles curvas elípticas con coeficientes en $\mathbb{Q}$ que se encuentran en la base de datos LMFBD.

[18] Ver Wikipedia.