Criptografía en curvas de Pell y generalizaciones

María Soledad Villar Lozano
Orientadores: Gonzalo Tornaría, Nathan Ryan

Licenciatura en Matemática
Facultad de Ciencias
Universidad de la República
Uruguay, 27 de setiembre de 2010

Resumen

Esta monografía explica los fundamentos de la criptografía, haciendo particular énfasis en la criptografía de clave pública. Se presentan los conceptos matemáticos detrás de los criptosistemas clásicos como son los basados en el problema de la factorización y el problema del logaritmo discreto en subgrupos del grupo multiplicativo. Luego se desarrolla la criptografía en curvas de Pell como el nexo natural entre la criptografía clásica y la criptografía basada en geometría algebraica. Por último se explica la geometría algebraica detrás de la criptografía de curvas elípticas e hiperelípticas.

Índice general

Introducción

La seguridad de la información es de gran importancia en un mundo en que la comunicación sobre redes abiertas y almacenamiento de datos en formato digital juegan un rol cotidiano. La criptografía es una técnica que provee herramientas eficientes para asegurar la información. En [Coh05] encontramos la siguiente definición de criptografía:

> *La criptografía es el estudio de técnicas matemáticas relacionadas a aspectos de la seguridad de la información tal como confidencialidad, integridad de datos, identificación de entidades y autenticación del origen de datos.* [1]

Históricamente la criptografía se centró en el estudio de métodos para transmitir información en secreto, aún si la transmisión se realiza a través de un canal inseguro como puede ser un línea telefónica, o Internet. Para lograr transmisiones seguras, el método más antiguo y rápido es la criptografía simétrica o de clave secreta.

La criptografía simétrica descansa esencialmente en un secreto compartido entre las partes que se quieren comunicar. Esta clave se utiliza tanto para cifrar el mensaje como para descifrar. Si bien estos criptosistemas permiten cifrar y descifrar rápidamente, tienen una desventaja: las claves deben compartirse entre ambas partes de antemano por un canal seguro.

A medidados de la década del 70 aparece una idea revolucionaria: la criptografía de clave pública o criptografía asimétrica. La criptografía de clave pública se basa en las funciones *one-way* con *trapdoor* que se definen en la sección 1.4.1; esencialmente son funciones cuya inversa no pueden calcularse en un tiempo razonable, a no ser que se posea una información especial conocida como *trapdoor*.

Todos los métodos conocidos de criptografía de clave pública son considerablemente más lentos que la criptografía simétrica. Es por esto que la criptografía

[1] [Coh05] p. xxix

de clave pública se usa como un complemento de la criptografía simétrica, es decir, se utiliza criptografía asimétrica para autenticación, intercambio de claves y firma de mensajes, pero una vez establecido un canal seguro, el resto de la comunicación se cifra con criptografía simétrica porque es más rápido.

Los problemas más usuales en los que se basa la criptografía de clave pública son la factorización (en el que se basa RSA) o el logaritmo discreto en grupos cíclicos de orden divisible por un primo grande. El problema del logaritmo discreto en términos informales es el siguiente:

> Si G es un grupo cíclico, $g \in G$ un generador y $h = g^k$ para algún $k \in \mathbb{Z}$. El problema del logaritmo discreto consiste en calcular k conociendo g y h.

La criptografía basada en el logaritmo discreto fue propuesta originalmente en 1976 en el protocolo de intercambio de claves de Diffie Hellman. La idea plantea el uso del problema del logaritmo discreto sobre el grupo multiplicativo de cuerpos finitos.

En 1985 Koblitz y Miller plantean independientemente la idea de utilizar grupos provenientes de la geometría algebraica, en particular el grupo de puntos de una curva elíptica sobre un cuerpo finito. Sin embargo, por distintos motivos, pasó mucho tiempo antes de que se utilizaran en la práctica estos criptosistemas.

En 1989 se propone el uso de la jacobiana de curvas hiperelípticas para criptosistemas basados en el logaritmo discreto, como una generalización natural del grupo de puntos de la curva elíptica. En el caso de las curvas hiperelípticas, el conjunto de puntos de la curva no es un grupo, sino que es necesario hacer uso de la geometría para construir el grupo sobre el cual se utilizarán primitivas basadas en el logaritmo discreto.

La distancia cronológica y conceptual, nos permite considerar dos familias de criptosistemas separadas. Por un lado el planteo clásico, con RSA y el logaritmo discreto en cuerpos finitos, y por otro lado el planteo geométrico, donde se encuentra la criptografía de curvas elípticas y jacobianas de curvas hiperelípticas.

En el mundo clásico, tanto RSA como los criptosistemas basados en el logaritmo discreto sobre cuerpos finitos, tienen ataques subexponenciales. En criptografía basada en geometría algebraica, tanto en curvas elípticas como en jacobianas de curvas hiperelípticas con género pequeño, hasta ahora no se conocen ataques subexponenciales. Esta situación trae como consecuencia que la criptografía basada en la geometría permita utilizar claves más pequeñas que en la criptografía clásica. Por ejemplo, la seguridad que ofrece el criptosistema RSA con claves de 2048 bits se estima equivalente

a la seguridad ofrecida por criptosistemas basados en curvas elípticas o hiperelípticas sobre cuerpos finitos con claves de 256 bits [Pat00]. Aunque las operaciones del grupo en curvas elípticas e hiperelípticas son más costosas de calcular que en el grupo multiplicativo, el tamaño pequeño de la clave lo compensa, en particular en ambientes restringidos como smart cards (ver 3.4).

En este contexto se centra el tema de esta monografía. Por un lado tenemos el planteo clásico basado en cuerpos finitos, y por otro lado el planteo moderno basado en geometría algebraica sobre cuerpos finitos. Es entonces donde presentamos la criptografía basada en curvas de Pell, que podría considerarse el nexo entre el mundo clásico y el mundo geométrico. La faceta geométrica radica en que las curvas de Pell son curvas algebraicas de género cero y la definición de la ley de grupo es similar a la ley de grupo en curvas elípticas. La faceta clásica la da el hecho de que el grupo de puntos es de hecho el grupo multiplicativo (a menos de un isomorfismo fácilmente computable).

El objetivo de esta monografía es explicar los fundamentos de la criptografía, haciendo especial énfasis en la criptografía clásica, para luego exponer la criptografía basada en curvas de Pell como un nexo natural entre el mundo clásico anterior, y lo que sigue: la criptografía basada en la variedad jacobiana de curvas elípticas e hiperelípticas.

Una pregunta pertinente es a qué nivel se realiza el estudio de estos sistemas criptográficos.

Los sistemas criptográficos en general se basan en un objeto matemático, sobre el cual se realiza un supuesto. A partir de este supuesto se plantea una primitiva criptográfica. Luego se diseña un protocolo de comunicación que hace uso de esas primitivas criptográficas con el afán de lograr ciertos objetivos de seguridad.

Luego de diseñado el sistema, se implementan en hardware o software el objeto matemático, las primitivas criptográficas y el protocolo, y de esta forma se obtiene un criptosistema.

Esta monografía estudia desde un punto de vista matemático las capas más abstractas de los sistemas criptográficos, como son los objetos matemáticos y las primitivas criptográficas. Desarrolla los conceptos básicos de la criptografía, y presenta distintos objetos matemáticos utilizados en la criptografía asimétrica; buscando explicitar la relación entre la concepción más clásica de la criptografía asimétrica (basada en subgrupos del grupo multiplicativo de cuerpos finitos) y el enfoque moderno basado en grupos provenientes de la geometría.

La presente monografía se desarrolla en los siguientes capítulos:

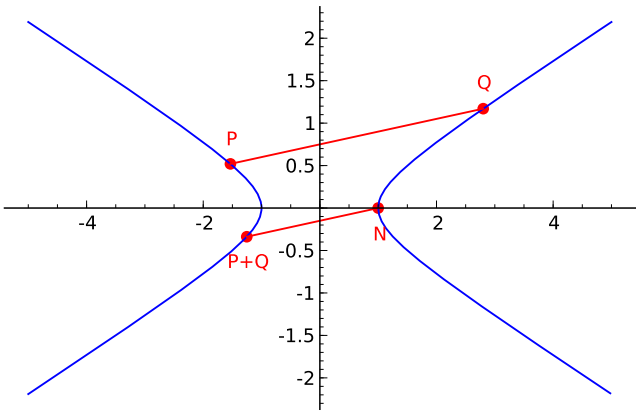


Figura 1: Interpretación geométrica de la ley de grupo para curvas de Pell.

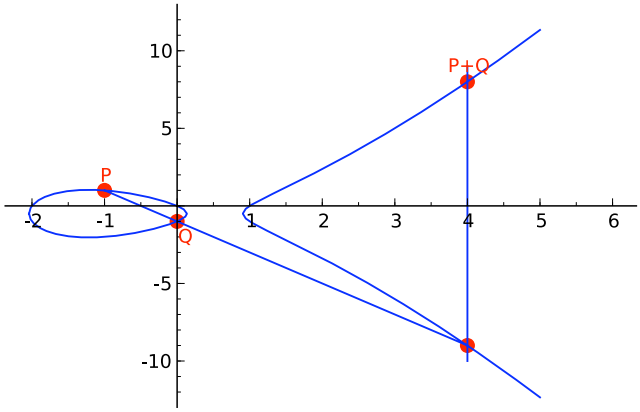


Figura 2: Ley de grupo para curvas elípticas. El grupo de puntos coincide con el grupo de Picard.

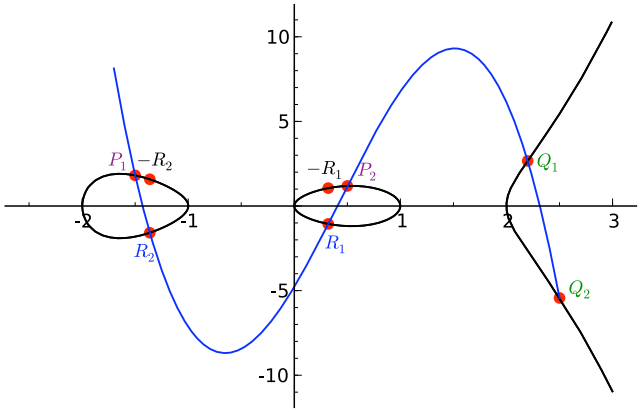
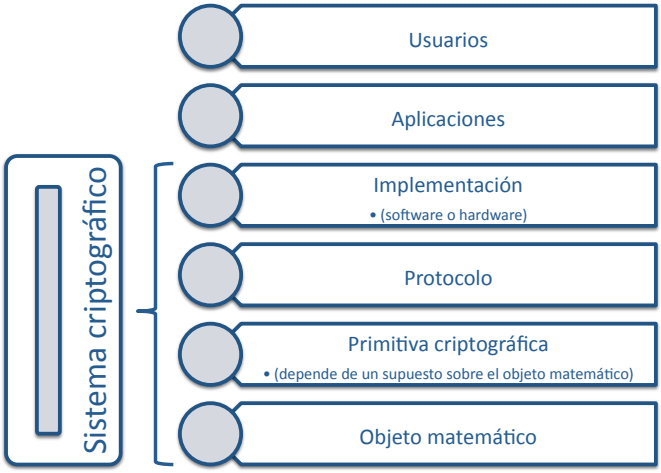


Figura 3: La ley de grupo en curvas hiperelípticas no está definida sobre el conjunto de puntos. En este caso $g = 2$ y la suma puede expresarse de a pares de puntos: $([P_1] + [P_2]) + ([Q_1] + [Q_2]) = [-R_1] + [-R_2]$



Cuadro 1: Estructura de un sistema criptográfico

En el capítulo 1 se hace una breve introducción a la criptografía, explicando cuáles son los problemas que busca resolver y cómo se formalizan. El foco de este capítulo son los objetos matemáticos y las primitivas criptográficas desde su visión clásica.

En el capítulo 2 se define una estructura de grupo sobre los puntos $\mathbb{F}_p-$racionales de una curva de Pell, que desde el punto de vista geométrico, son curvas de género cero. La ley de grupo en los puntos de las curvas de Pell se define geométricamente y resulta muy similar a la definición de la ley grupo en las curvas elípticas. El grupo de puntos obtenido se parametriza racionalmente por el grupo multiplicativo $\mathbb{F}_p^\times$ o por el grupo de elementos de norma 1 de $\mathbb{F}_{p^2}$, motivo por el cual, podemos decir que desde un punto de vista criptográfico, este ejemplo pertenece al mundo de la criptografía clásica. Sin embargo, moralmente este ejemplo descubre una relación entre la visión clásica basada en cuerpos y la geometría.

Por último, en el capítulo 3 se presentan grupos basados en la geometría de curvas algebraicas. En la sección 3.1 se explica la teoría que permite construir una variedad abeliana a partir de la geometría de una curva algebraica proyectiva. Esta variedad abeliana se conoce como variedad jacobiana de la curva. En las secciones 3.2 y 3.3 se estudian las jacobianas de las curvas elípticas e hiperelípticas respectivamente. El objetivo es explicar la estructura del grupo y cómo se utiliza con fines criptográficos.

Capítulo 1

Criptografía

El objetivo de este capítulo es introducir las visiones clásica y moderna de la criptografía; presentar una breve explicación de sus fundamentos, los problemas que busca resolver y los supuestos sobre los que descansa.

Mi intención para este capítulo es trasmitir lo mejor posible mi visión personal de esta área, que me resulta una teoría muy interesante que a veces escapa a la intuición.

1.1. Conceptos básicos

La criptografía estudia el diseño y análisis de técnicas matemáticas que permiten comunicaciones seguras en presencia de adversarios maliciosos.

El modelo es el siguiente: dos entidades Alice y Bob se comunican sobre un canal inseguro en presencia de un adversario malicioso (Eve). Los objetivos principales de la comunicación segura son:

Confidencialidad
Los mensajes enviados por Alice a Bob no deberán poder ser leidos por Eve.

Integridad de los datos
Bob deberá poder detectar si el mensaje enviado por Alice fue modificado por Eve.

Autenticación del emisor
Bob deberá poder verificar si los mensajes enviados supuestamente por Alice fueron efectivamente enviados por Alice.

Autenticación de la entidad

Bob deberá poder verificar la identidad del otro extremo de la comunicación.

No repudio
Cuando Bob recibe un mensaje de Alice, no sólo Bob puede convencerse de que el mensaje proviene de Alice, sino que también podrá convencer a una tercer parte de esto. Es decir, Alice no podrá negar haber enviado el mensaje a Bob.

Con el objetivo de modelar amenazas realistas en general se asume que el adversario Eve tiene la capacidad de leer todos los datos transmitidos sobre el canal e incluso tiene la capacidad de modificar datos enviados e introducir datos. Además Eve tiene un poder de cómputo significativo y conoce todos los protocolos y esquemas criptográficos utilizados (y sus implementaciones) a excepción de las claves secretas. El desafío de la criptografía es diseñar mecanismos que aseguren se cumplan los objetivos de seguridad en presencia de este tipo de adversarios.

Sin embargo, la criptografía no fue planteada siempre en estos términos. El concepto de criptografía existe desde las primeras civilizaciones, donde se desarrollaban distintas técnicas muy ingeniosas para enviar mensajes secretos durante las campañas militares. La criptografía clásica era una especie de arte desarrollado con objetivos exclusivamente militares y utilizado por los gobiernos.

Desde un punto de vista académico, la criptografía moderna (o criptografía matemática) surge con Claude Shannon cuando en 1949 publica el artículo *Communication Theory of Secrecy Systems* [Sha49], y años más tarde el libro *The Mathematical Theory of Communication*, con Warren Weaver [Sha63]. Estos trabajos, junto con los otros que publicó sobre la teoría de la información y la comunicación, establecieron las bases teóricas para la criptografía y el criptoanálisis.

Luego de estos trabajos, la criptografía desapareció de la escena por un tiempo para quedarse dentro de organizaciones gubernamentales secretas como la NSA (National Security Agency, Estados Unidos) y la GCHQ (Government Communications Headquarters, Gran Bretaña). Muy pocos trabajos se hicieron públicos hasta mediados de los 70.

Hoy en día, la criptografía tiene un uso generalizado en todo el mundo, principalmente para comercio electrónico e intercambio de información a través de Internet.

A través de los siglos, las prácticas fueron cambiando y las técnicas evolucionando. Uno de los cambios más paradigmáticos sucedió mucho antes de la era de la información, cuando a fines del siglo XIX el lingüista y criptógrafo holandés Auguste Kerckhoffs publicó seis principios básicos para el correcto

diseño de sistemas criptográficos [Ker83]. Uno de ellos se mantiene vigente hasta el día de hoy y se conoce como principio de Kerckhoffs:

> El método de cifrado no deberá requerir se mantenga en secreto y debe ser susceptible de caer en manos enemigas sin que esto represente un inconveniente.

Este principio plantea que el esquema criptográfico no debería mantenerse en secreto, sino que el secreto debe radicar solamente en la clave. Las ventajas que tiene la aplicación de este principio incluyen:

- Es más fácil de guardar el secreto si es únicamente la clave.
- Es más fácil de cambiar una clave que un algoritmo en caso de que el secreto sea descubierto.
- La especificación del algoritmo se puede "filtrar" y de esta forma romper el criptosistema.
- Muchas veces el algoritmo se puede reconstruir con métodos de ingeniería inversa.
- La distribución de un algoritmo secreto entre N partes suele ser una dificultad.
- La seguridad de la mayoría de los sistemas criptográficos depende de supuestos no demostrados. Si el algoritmo es público, se somete a un escrutinio más amplio, especialmente si es muy utilizado. De esta forma, si hay una comunidad científica tratando de romper el criptosistema sin éxito se puede "confiar" más en su seguridad que si no la hay.

Este principio sienta las bases del *diseño criptográfico abierto*, en contraposición al paradigma conocido como *seguridad por oscuridad* que era el paradigma criptográfico clásico y aún hoy en día se suele utilizar, a pesar de sus desventajas.

1.2. Definiciones

Sea $\mathcal{M}$ el espacio de mensajes, $\mathcal{K}$ un espacio de claves y $\mathcal{C}$ un espacio que llamaremos espacio de mensajes cifrados; todos conjuntos finitos. Un criptosistema se compone por tres funciones:

- **Gen** es una variable aleatoria sobre el espacio de claves $\mathcal{K}$ con cierta distribución de probabilidad.

- **Enc** es una familia de funciones parametrizadas sobre el espacio de claves que realizan el cifrado de los mensajes. $\{\mathbf{Enc}_k : \mathcal{M} \to \mathcal{C}\}_{k \in \mathcal{K}}$
- **Dec** es la familia de funciones de descifrado: $\{\mathbf{Dec}_k : \mathcal{C} \to \mathcal{M}\}_{k \in \mathcal{K}}$ con la propiedad: $\mathbf{Dec}_k(\mathbf{Enc}_k(m)) = m$

1.3. Seguridad en criptografía

Existen distintos conceptos de seguridad en criptografía que explicaremos en esta sección. Algunos son incondicionales, pero en la mayoría de los casos prácticos la afirmación de seguridad depende de algún supuesto.

Es por este motivo que la criptografía moderna se basa en tres principios:

Formulación de definiciones de seguridad

Frente a la pregunta ¿qué significa que un criptosistema sea seguro? aparecen algunas respuestas que no resultan ser convenientes en la práctica. Por ejemplo:

- Ningún adversario podrá hallar la clave a partir del texto cifrado.
- Ningún adversario podrá hallar el texto claro a partir del texto cifrado.

Ambas posiciones son inconvenientes debido a que un adversario podría conocer mucho del texto claro sin obtener la clave y sin conocer el texto claro. Por ejemplo, el adversario podría descubrir a partir del texto cifrado que el texto claro corresponde a un número en determinado rango, lo que puede resultar un inconveniente para las partes honestas.

Es por esto que en general se utiliza la siguiente definición de seguridad:

Definición 1.3.1. Un esquema de encripción es seguro si ningún adversario puede computar ninguna función del texto claro a partir del texto cifrado.

De esta forma, un criptosistema se considera roto si el adversario puede conocer una función del texto claro a partir de texto cifrado. Para utilizar correctamente esta definición habrá que definir qué es lo que el adversario puede hacer, en términos de:

- Poder de cómputo
- Tipo de ataque:
  - Solo texto cifrado: El adversario conoce sólo el texto cifrado.
  - Texto claro conocido: Conoce un conjunto de pares (texto claro, texto cifrado).

- Texto claro elegido: El adversario puede elegir algunos mensajes en texto claro y conocer su respectivo texto cifrado.

Teniendo en cuenta las definiciones anteriores, una definición de seguridad tendrá la siguiente forma:

Definición 1.3.2. Un esquema criptográfico para una tarea dada se dice seguro si ningún adversario con determinado poder puede lograr romperlo de una forma específica.

La clave es no asumir conocido qué podrá hacer el adversario para intentar romper el esquema criptográfico, es decir, se asume conocida su capacidad pero no su estrategia. Esto se conoce como principio de arbitrariedad del adversario.

Establecer supuestos

Para la mayoría de los esquemas criptográficos actuales no existe una demostración de su seguridad que sea incondicional. De hecho, probar la seguridad de estos esquemas requiere responder preguntas de la teoría de la computación que aparentemente están muy lejos de ser respondidas hoy.

Entonces la seguridad en general se basa en algún supuesto que deberá definirse con precisión. De un supuesto de este tipo, se espera que sea un problema matemático de bajo nivel (no un protocolo complejo sino un problema concreto) y que eventualmente pueda ser compartido por varios esquemas criptográficos, y puesto a prueba durante mucho tiempo.

Los objetivos detrás de este principio son:

- Conocer qué problemas matemáticos se utilizan para construir esquemas criptográficos, de forma tal que estos problemas puedan ponerse a prueba.
- Poder comparar esquemas de seguridad.
- Permitir la demostración de la seguridad de esquemas criptográficos a partir de los supuestos.

Demostraciones de seguridad

La experiencia muestra que la intuición en criptografía y seguridad informática en general no suele ser muy acertada, y las demostraciones en general condicionales de seguridad son tan sólo una garantía parcial de la seguridad de un criptosistema.

Este tipo de demostraciones se realizan por reducción y tienen la siguiente forma: *Bajo el supuesto X, el esquema Y es seguro; ya que romper el esquema Y implica romper el supuesto X.*

1.3.1. Seguridad perfecta

La seguridad perfecta es un concepto introducido por Claude Shannon durante la Segunda Guerra Mundial que garantiza seguridad incluso contra adversarios con poder de cómputo infinito. Sin embargo, los sistemas que alcanzan seguridad perfecta tienen muchas limitaciones, y muchas veces la información que se desea proteger no amerita las dificultades que estos sistemas presentan.

Definición 1.3.3 (Seguridad perfecta)**.** Sea (Gen, Enc, Dec) un criptosistema, se dice que tiene seguridad perfecta si observar el texto cifrado no provee más información al atacante que la que ya se tenía:

$$\Pr[M = m|C = c] = \Pr[M = m].$$

Un criptosistema que provee seguridad perfecta es el *Criptosistema de Vernam* también conocido como *One time pad*.

Criptosistema de Vernam

El criptosistema de Vernam, introducido en 1917, se define de la siguiente manera:

- Fijado $l \geq 0$, se definen el espacio de textos claros $\mathcal{M}$ y el espacio de claves $\mathcal{K}$ como $\{0,1\}^l$.
- El esquema de generación de claves Gen será la elección de un elemento al azar de $\mathcal{K}$ de acuerdo a una distribución uniforme.
- $\text{Enc}_k(m) = k \oplus m$ donde $\oplus$ corresponde a la operación de XOR bit a bit.
- $\text{Dec}_k(m) = k \oplus m$

Teorema 1.3.1. El criptosistema de Vernam tiene seguridad perfecta.

Demostración. Fijada una distribución sobre $\mathcal{M}$ sea $m \in \mathcal{M}$ y $c \in \mathcal{C}$ entonces:

$$\begin{aligned} \Pr[M = m|C = c] &= \frac{\Pr[M = m \ , \ C = c]}{\Pr[C = c]} \\ &= \frac{\Pr[M = m \ , \ K = m \oplus c]}{\sum_{\bar{m} \in \mathcal{M}} \Pr[M = \bar{m} \ , \ K = \bar{m} \oplus c]} \end{aligned}$$

Como M y K son independientes, obtenemos:

$$\begin{aligned}\Pr[M = m|C = c] &= \frac{\Pr[M = m]\ \Pr[K = m \oplus c]}{\sum_{\bar{m}\in\mathcal{M}} \Pr[M = \bar{m}]\ \Pr[K = \bar{m} \oplus c]} \\ &= \frac{(1/|\mathcal{K}|)\Pr[M = m]}{(1/|\mathcal{K}|)\sum_{\bar{m}\in\mathcal{M}} \Pr[M = \bar{m}]} \\ &= \Pr[M = m]\end{aligned}$$

□

Este esquema, si bien provee seguridad perfecta, tiene ciertas limitaciones en la práctica. La clave debe ser tan larga como el mensaje y solamente se puede utilizar una vez (si se quiere mantener el status de seguridad perfecta). En la práctica entonces aparecen dos problemas no menores: la distribución y el manejo de claves.

La pregunta natural en este contexto sería ¿se puede sortear este problema y aún así mantener la seguridad perfecta?. La respuesta es no: estas limitaciones son inherentes a cualquier sistema que quiera lograr seguridad perfecta. Esto se justifica en el siguiente teorema:

Teorema 1.3.2. Sea (Gen, Enc, Dec) un esquema criptográfico con seguridad perfecta sobre un espacio de mensajes $\mathcal{M}$ y espacio de claves $\mathcal{K}$. Entonces $|\mathcal{K}| \geq |\mathcal{M}|$.

Demostración. Suponemos por absurdo que $|\mathcal{K}| < |\mathcal{M}|$. Consideramos $c \in \mathcal{C}$ un texto cifrado con probabilidad no nula y el conjunto

$$\mathcal{M}(c) := \{\bar{m} : \bar{m} = \mathrm{Dec}_k(c) \text{ para algún } k \in \mathcal{K}\}$$

Como $|\mathcal{M}(c)| < |\mathcal{K}|$ entonces existe $m \in \mathcal{M}$ tal que $m \notin \mathcal{M}(c)$

$$\Pr[M = m|C = c] = 0 \neq \Pr[M = m]$$

□

El Teorema de Shannon[1] caracteriza los esquemas que ofrecen seguridad perfecta, cualquiera sea la distribución de probabilidad en $\mathcal{M}$ en los siguientes términos:

Teorema 1.3.3 (Shannon)**.** Sea (Gen, Enc, Dec) un esquema criptográfico sobre un espacio de mensajes $\mathcal{M}$ donde $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$. El esquema tiene seguridad perfecta sí y sólo si:

[1] Ver por ejemplo el capítulo 2 de [Kat08]

1. La elección de las claves Gen tiene distribución uniforme sobre $\mathcal{K}$.
2. Para todo $m \in \mathcal{M}$ y todo $c \in \mathcal{C}$ existe un único $k \in \mathcal{K}$ tal que $\text{Enc}_k(m) = c$.

1.3.2. Seguridad computacional

En la práctica, los esquemas criptográficos más utilizados no ofrecen seguridad perfecta. Esto ocurre debido a la dificultad práctica que significa utilizar claves tan largas como el texto a cifrar, teniendo en cuenta que la clave deberá ser compartida entre ambas partes en la comunicación y que además debe intercambiarse a través de un canal seguro.

En cambio se utilizan criptosistemas que tienen la propiedad de poder ser rotos (se puede hallar la clave con la que se encripta) con una cantidad suficiente de tiempo y poder de cómputo. Sin embargo, bajo ciertos supuestos el tiempo de computación necesario para romper estos sistemas es de cientos de años, aún usando la supercomputadora más rápida. Para fines prácticos en general se considera que este nivel de seguridad es suficiente.

Aparece entonces el concepto de *seguridad computacional*, que es más débil que la seguridad perfecta en el siguiente sentido: la seguridad computacional no ofrece garantías contra adversarios de poder de cómputo ilimitados como lo hacen los sistemas de seguridad perfecta. Además, no está demostrado que de hecho existan criptosistemas computacionalmente seguros. La existencia de este tipo de criptosistemas depende de supuestos no demostrados[2], mientras no se necesitan supuestos para considerar un sistema criptográfico con seguridad perfecta.

La seguridad computacional supone ciertas relajaciones en la seguridad:

- La seguridad se preserva contra adversarios *eficientes* que computan en una cantidad determinada de tiempo.
- Los adversarios podrán potencialmente tener éxito en su ataque, pero con una probabilidad muy pequeña o despreciable (de forma tal que no represente una amenaza realista).

Esta idea heurística de seguridad matemática se puede definir en los siguientes términos:

[2]La existencia de este tipo de criptosistemas en particular implica que $\mathcal{P} \neq \mathcal{NP}$. Esta es una pregunta abierta en el área de la teoría de complejidad computacional que forma parte de la lista de los Problemas del Milenio del Clay Mathematics Institute http://www.claymath.org/millennium/P_vs_NP/

Sea $n \in \mathbb{N}$ un parámetro de seguridad elegido por las partes honestas (que podría ser por ejemplo, la cantidad de bits de la clave). El tiempo de ejecución de los algoritmos de cifrado y descifrado se estudiará como una función de n, así como también será una función de n el tiempo de ejecución del adversario y su probabilidad de éxito.

Definición 1.3.4. Decimos que un algoritmo es *eficiente* cuando es un algoritmo probabilístico que corre en tiempo polinomial en el parámetro n. Esto significa que para ciertas constantes a, c el algoritmo ejecuta una cantidad de instrucciones acotada por an^c.

Se requiere que las partes honestas (Alice y Bob) ejecuten en tiempo polinomial y la seguridad computacional solo provee seguridad (condicionada) contra adversarios que ejecuten en tiempo polinomial (aunque podrán ser más poderosos y ejecutar por más tiempo que las partes honestas). Las estrategias de los adversarios que requieran tiempos de ejecución de órdenes de ejecución mayores que polinomial no se consideran amenazas realistas.

Definición 1.3.5. Se dice que una familia de sucesos parametrizados en n tiene probabilidad muy pequeña o despreciable si esta probabilidad es menor que el inverso de cualquier polinomio en n. Esto implica que para toda constante c, la probabilidad del suceso es menor que n^{-c} para valores de n suficientemente grandes.

Una función con esta propiedad será llamada **negl** en esta monografía por su nombre en inglés. En general **negl** es una función que verifica que para todo polinomio p existe un entero N tal que $\forall\, n > N$ se cumple

$$\mathbf{negl}(n) < \frac{1}{p(n)}$$

Definición 1.3.6. Un esquema criptográfico es computacionalmente seguro si todo adversario que compute únicamente algoritmos probabilísticos en tiempo polinomial tiene éxito de romper el algoritmo con probabilidad despreciable.

Ataques de fuerza bruta

Sea un esquema criptográfico donde el espacio $\mathcal{K}$ es significativamente más pequeño que el espacio $\mathcal{M}$. Los siguientes ataques se aplican independientemente de cómo sea el esquema criptográfico.

Dado un texto cifrado c, un adversario puede descifrar c utilizando todas las claves $k \in \mathcal{K}$. De esta forma se construye una lista de todos los posibles mensajes para los cuales c puede corresponder. Como $|\mathcal{K}| < |\mathcal{M}|$ esto filtra información sobre el mensaje.

Si además se realiza un ataque de texto claro conocido, sabiendo que los textos cifrados $c_1, \dots c_l$ corresponden a los textos claros $m_1, \dots m_l$ el adversario puede puede descifrar cada uno de los c_i con todas las posibles claves hasta encontrar $k \in \mathcal{K}$ tal que $\text{Dec}_k(c_i) = m_i$ para todo i. Esta clave k será única con una probabilidad alta, y en tal caso, el criptosistema fue roto.

Este tipo de ataque es conocido como ataque de fuerza bruta, y permite al adversario tener éxito en su ataque con probabilidad esencialmente 1 en tiempo polinomial en $|\mathcal{K}|$

Por otro lado, el adversario podría tratar de "adivinar" la clave eligiendo $k \in \mathcal{K}$ al azar y chequeando si $\text{Dec}_k(c_i) = m_i$ para todo i. Si es así, con una probabilidad alta k es la clave que Alice y Bob estén usando.

En este ataque, el adversario ejecuta un algoritmo de tiempo de ejecución constante, con una probabilidad de éxito de $1/|\mathcal{K}|$.

De esta forma, podemos concluir que el tamaño de $\mathcal{K}$ deberá ser de orden mayor que polinomial que el parámetro de seguridad n.

1.4. Paradigmas criptográficos

En el planteo de la criptografía clásico Alice y Bob comparten una clave secreta k y las funciones compartidas Dec_k y Enc_k les permiten comunicarse. Este tipo de esquemas criptográficos se llaman sistemas criptográficos simétricos.

Los esquemas criptográficos simétricos son apropiados en muchas situaciones pero tienen algunas desventajas derivadas del hecho de que las partes tienen que conocer un secreto compartido para poder enviarse mensajes cifrados. Entre los problemas que tienen este tipo de esquemas se destacan:

1. Problema de distribución de claves.
2. Problema de manejo de claves.
3. No ofrecen posibilidad de no repudio.

El problema de distribución de claves es un tema no menor. En la década de 1970 los bancos que querían intercambiar mensajes cifrados con sus clientes importantes tenían empleados denominados *courier* que se encargaban de visitar a los clientes y entregar las claves personalmente. Esta práctica además de ser costosa, complicada y poco escalable, es de hecho una vulnerabilidad en la seguridad del sistema. No importa cuan seguro es un criptosistema en la teoría, en la práctica puede ser atacado en el proceso de distribución de claves.

Para resolver este problema, a mediados de la década de 1970 aparece un concepto que revoluciona la criptografía. Este concepto es el de criptografía asimétrica, que es el tema central de esta monografía. Antes de explicar en qué consiste la criptografía asimétrica me permito la siguiente digresión:

En [Sin01] se presenta una idea sumamente sencilla para enviar un mensaje en secreto sin la necesidad de compartir una clave de antemano.

> *Supongamos que Alice quiere mandar un mensaje secreto a Bob. Alice podría ponerlo en una caja con un candado y enviarlo por correo. Cuando le llega la caja a Bob, éste no va poder abrirla porque la llave la tiene Alice, pero lo que puede hacer es ponerle un candado suyo a la caja y enviarla nuevamente a Alice. Al recibir la caja Alice retira su candado y manda nuevamente la caja a Bob, que ahora se encuentra en condiciones de abrir la caja y leer el mensaje secreto, ya que el único candado que tiene la caja es el de Bob. De esta forma Alice y Bob logran compartir un secreto, asegurándose que en todo momento el secreto viajó seguro.*[3]

Este esquema hace uso fuertemente de la *conmutatividad de los candados* para permitir enviar el mensaje secreto.

Una idea anterior a la criptografía asimétrica, desarrollada en secreto por James Ellis para el Gobierno Británico en 1969 es la siguiente:

> *Para encriptar un mensaje telefónico, la estrategia podría ser la siguiente: el receptor oculta la información enviada por el emisor introduciendo ruido aleatorio pregrabado en la linea telefónica. Luego el receptor podrá substraer el ruido (que solamente él conoce).*[4]

Ninguna de estas ideas tiene como objetivo ser implementadas en la práctica, sino que son ejemplos que muestran que el problema de la distribución de claves tiene solución, es decir, es posible enviar un mensaje secreto sin la necesidad de compartir un secreto de antemano si el receptor participa en forma activa en el proceso de encriptación. Este es el cambio de paradigma entre el modelo de criptografía simétrica clásico y el modelo de criptografía asimétrico.

[3] [Sin01] p. 193
[4] [Sin01] p. 213

1.4.1. Criptografía de clave pública

La noción de criptografía de clave pública fue introducida en 1975 por Diffie, Hellman y Merkle. Para explicarla es necesario definir lo que en inglés se conoce como *trapdoor one-way function*.

Definición 1.4.1 (Función *one-way*)**.** Una función *one-way* $f : \mathcal{M} \to \mathcal{C}$ es una función invertible que verifica las siguientes propiedades:

1. Computable eficientemente: Existe un algoritmo de tiempo polinomial M_f que permite calcular f, es decir $M_f(m) = f(m)$ para todo $m \in \mathcal{M}$.
2. Difícil de invertir: Para todo algoritmo probabilístico que corra en tiempo polinomial $\mathcal{A}$, existe una función negl que verifica que:

   $$\Pr_{m \in \mathcal{M}}[\mathcal{A}(f(m)) = m] \leq \text{negl}(|\mathcal{M}|)$$

   donde negl es una función que verifica que para todo polinomio p existe un entero N tal que $\forall\, n > N$ se cumple

   $$\text{negl}(n) < \frac{1}{p(n)}$$

La existencia de este tipo de funciones implica $\mathcal{P} \neq \mathcal{NP}$, que como mencionamos antes es un problema abierto. Sin embargo, la criptografía (tanto simétrica como asimétrica) descansan en el supuesto de la existencia de este tipo de funciones.

Definición 1.4.2 (Función *one-way* con *trapdoor*)**.** Una función *one-way* $f : \mathcal{M} \to \mathcal{C}$ se dice que tiene *trapdoor* (o puerta trasera) si existe alguna información extra con la cual f puede invertirse en tiempo polinomial. Esa información adicional se denomina *trapdoor*.

Para construir un criptosistema de clave pública se parte de una familia de funciones *one-way* con *trapdoor* $\{f_k : \mathcal{M} \to \mathcal{C}\}_{k \in \mathcal{K}}$. Esta familia deberá tener la propiedad de que para cada $k \in \mathcal{K}$ el *trapdoor* $t(k)$ deberá ser computable eficientemente (en el sentido de las definiciones anteriores). Además, para cada $k \in K$ deberá conocerse un algoritmo eficiente para computar f_k y deberá ser computacionalmente difícil (y prácticamente imposible) obtener $t(k)$ y por ende k a partir de f_k.

Dada esta familia, Alice elige al azar $a \in \mathcal{K}$ y publica el algoritmo E_a para calcular f_a. E_a será la clave pública de Alice, mientras el *trapdoor* $t(a)$ que se utiliza para invertir f_a es la clave privada de Alice.

Para enviar un mensaje $m \in \mathcal{M}$ a Alice, Bob envía $f_a(m)$ a Alice. Como Alice es la única persona que tiene la habilidad de invertir f_a, solo Alice

podrá recuperar el mensaje m. De esta forma se solucionan los problema de distribución y manejo de claves (debido a que hay un solo par de claves asociado a cada entidad y no es necesario un canal seguro para distribuir las claves).

Para solucionar los problemas de autenticación y no repudio, Alice puede firmar digitalmente los mensajes que envía utilizando la función f_a. Supongamos que $\mathcal{M} = \mathcal{C}$. Si Alice quiere mandar un mensaje m firmado a Bob, puede enviar a Bob el par $(m, s = f_a^{-1}(m))$. De esta forma cualquiera puede verificar que $m = f_a(s)$ usando la clave pública E_a, pero solamente Alice pudo haber computado s.

1.4.2. Criptografía basada en grupos

Sea G un grupo multiplicativo de orden n. Si existe un algoritmo polinomial para multiplicar en G, entonces la exponenciación también se puede calcular polinomialmente a través del algoritmo de exponenciación binaria.

Existen dos posibles enfoques que asumen dos supuestos distintos:

- Utilizar el orden del grupo como *trapdoor*.
- Utilizar la exponenciación como función *one-way* con *trapdoor*.

Orden del grupo como *trapdoor*

Para este primer enfoque se considera un grupo G cuya ley de grupo es computable mediante un algoritmo polinomial, pero su orden n es computacionalmente difícil (en términos de la definición anterior). Entonces se construye un criptosistema de clave pública de la siguiente manera:

Alice elige un grupo G del cual conoce $|G| = n$, y un entero al azar e tal que $1 < e < n$ y $\gcd(n, e) = 1$. Utilizando el algoritmo de Euclides Alice calcula d tal que $de \equiv 1 \pmod{n}$. La clave pública de Alice será (G, e), mientras la clave privada será d.

De esta forma, $\mathcal{M} = G$, $\mathcal{C} = G$ y la función *one-way* con *trapdoor* será

$$\begin{aligned} f(m) &= m^e \\ f^{-1}(c) &= c^d \end{aligned}$$

De esta forma si Bob quiere enviar el mensaje m cifrado a Alice deberá enviar $c = m^e$ y Alice para recuperar el mensaje original calcula

$$c^d = (m^e)^d = m^{de} = m$$

Para que este criptosistema sea seguro, tomar raices e-ésimas en G deberá ser un problema computacionalmente difícil.

Ejemplo (RSA)**.** Introducido por Rivest, Shamir y Adleman en 1977 fue el primer criptosistema de clave pública y el más utilizado hoy en día.

Se construye: Alice elige dos primos grandes p y q; considera $N = p\,q$ y G el grupo multiplicativo $\mathbb{Z}_N^\times$. Entonces $|G| = \phi(N) = (p-1)(q-1)$. Como Alice conoce la factorización de N puede calcular el orden del grupo y así aplicar el sistema definido con el orden del grupo como *trapdoor*.

Alice	Eve	Bob
elige dos primos grandes p, q		
calcula $N = p\,q$		
elige e coprimo con $\phi(N)$		
clave pública: (N, e)		
	$\xrightarrow{(N,e)}$	
		codifica su mensaje a un elemento $m \in \mathbb{Z}_N^\times$
		encripta el mensaje como $c = m^e \pmod N$
	$\xleftarrow{c}$	
conociendo que		
$d\,e \equiv 1 \pmod{\phi(N)}$		
computa $c^d = (m^e)^d = m$		
obteniendo el mensaje cifrado		

Cuadro 1.1: Algoritmo RSA.

El problema de la factorización se puede definir en los siguientes términos:

Sea GenMódulo un algoritmo polinomial que recibe como entrada n y devuelve (N, p, q) donde $N = p\,q$ y p, q son números de n bits primos excepto en un caso de probabilidad despreciable en n. Se considera el experimento para todo algoritmo $\mathcal{A}$

Experimento de factorización $\text{Factor}_{\mathcal{A},\text{GenMódulo}}(n)$

1. (N, p, q)=GenMódulo(n)
2. (p', q')=$\mathcal{A}(n)$ donde $p', q' > 1$
3. Devolver 1 si $p'q' = N$, 0 si no

Decir que *factorizar es un problema difícil* formalmente significa que existe un algoritmo GenMódulo polinomial en n, y una función negl despreciable en n

tal que para todo algoritmo probabilístico de tiempo de ejecución polinomial $\mathcal{A}$ se verifica que

$$\Pr[\text{Factor}_{\mathcal{A},\text{GenMódulo}}(n) = 1] \leq \text{negl}(n)$$

Decir que factorizar es un problema difícil es equivalente a decir que calcular el orden del grupo es díficil, ya que calcular $\phi(N)$ dado N permite obtener la factorización de N en tiempo polinomial:

$$\begin{aligned} p+q &= N+1-\phi(N) \\ p\,q &= N \\ \Rightarrow p,q &= \frac{N+1-\phi(N) \pm \sqrt{(N+1-\phi(N))^2-4N}}{2} \end{aligned}$$

Es claro que de existir un algoritmo eficiente para factorizar (o equivalentemente, hallar $\phi(N)$) RSA no sería un criptosistema seguro. De hecho, la seguridad de RSA depende de que el siguiente problema sea computacionalmente difícil:

Problema RSA
Dados: N, $e > 0$ tal que $\gcd(e, \phi(N)) = 1$, $y \in \mathbb{Z}_N^\times$
Hallar: $x \in \mathbb{Z}_N^\times$ tal que $x^e \equiv y \mod N$

Se cree que el problema de la factorización y el problema RSA son equivalentes, pero no está demostrado que resolver el segundo implique resolver el primero. Por eso se dice que la seguridad de RSA *se basa* en la dificultad de la factorización.

Logaritmo discreto

Los sistemas que utilizan la exponenciación como función *one-way* con *trapdoor* se basan en el problema del logaritmo discreto:

Sea GenGrupo un algoritmo de tiempo de ejecución polinomial que recibe como entrada un entero n y devuelve un grupo cíclico G, su orden q (de n bits) y un generador $g \in G$. Se deberá conocer para G una forma de computar la operación de grupo eficiente. Consideremos entonces:

Experimento del logaritmo discreto $\text{LogDiscreto}_{\mathcal{A},\text{GenGrupo}}(n)$

1. $(G, q, g) = \text{GenGrupo}(n)$
2. Elegir $x \in \mathbb{Z}_q$ y sea $h = g^x$

3. $\mathcal{A}(G, q, g, h)$ devuelve $x' \in \mathbb{Z}_q$
4. Devolver 1 si $g^{x'} = h$, 0 si no.

Se dice que el problema del logaritmo discreto es difícil para GenGrupo si para todo algoritmo $\mathcal{A}$ que ejecute en tiempo polinomial existe una función negl, despreciable en n tal que

$$\Pr[\text{LogDiscreto}_{\mathcal{A},\text{GenGrupo}}(n) = 1] \leq \text{negl}(n)$$

Sea G un grupo de orden n en el cual el problema del logaritmo discreto se considera computacionalmente difícil. El siguiente algoritmo fue propuesto por T. ElGamal en el año 1985 y se basa en la exponenciación discreta:

1. Se elige un grupo cíclico y finito G. Sea $g \in G$ un generador.
2. Alice elige al azar un entero a tal que $1 < a < |G|$. g^a será su clave pública y a su clave privada.
3. Bob para enviar un mensaje m cifrado a Alice deberá elegir k al azar tal que $1 < k < |G|$ y calcular g^k.
4. Bob calcula $(g^a)^k$ y envía (g^k, mg^{ak})
5. Alice, para recuperar el mensaje m calcula $(g^k)^a$ y determina $m = mg^{ak}(g^{ka})^{-1}$

El cuadro 1.2 muestra el algoritmo y exactamente a qué datos tiene acceso cada una de las partes. Este algoritmo en principio se puede implementar para cualquier grupo, pero el requerimiento para que tenga utilidad criptográfica es que la operación de grupo pueda calcularse eficientemente y el logaritmo discreto sea difícil.

Algunos grupos que se utilizan para este tipo de criptosistemas son:

- Grupo multiplicativo del cuerpo finito $\mathbb{Z}_p$.
- Grupo multiplicativo del cuerpo finito $\mathbb{F}_{2^k}$.
- Grupo multiplicativo $\mathbb{Z}_N^\times$.
- Grupo de matrices invertibles sobre cuerpos finitos.
- Grupo de clases de un cuerpo cuadrático imaginario.

Alice	Eve	Bob
elige a tal que $1 < a < \|G\|$ si g es un generador del grupo G, g^a es la clave pública;		
	$\xrightarrow{g^a}$	
		codifica su mensaje a un elemento $m \in G$; elige k tal que $1 < k < \|G\|$ al azar y calcula g^k; encripta el mensaje como $c = mg^{ak}$;
	$\xleftarrow{g^k, c}$	
calcula $(g^k)^a$ y obtiene $m = mg^{ak}(g^{ka})^{-1}$		

Cuadro 1.2: Algoritmo ElGamal para un grupo genérico.

A los efectos de la presente monografía, es de particular interés el estudio de los grupos provenientes de la geometría algebraica, como son:

- Grupo de puntos racionales en curvas de Pell sobre cuerpos finitos.
- Grupo de puntos racionales en curvas elípticas sobre cuerpos finitos.
- Jacobiano de una curva hiperelíptica.

Capítulo 2

Criptografía en curvas de Pell

2.1. Introducción

La ecuación de Pell es una ecuación diofántica de la forma $x^2 - Dy^2 = 1$, donde $D > 0$ no es un cuadrado perfecto.

Existen varios enfoques posibles para estudiar la ecuación de Pell. Por ejemplo, desde un punto de vista aritmético las soluciones de la ecuación de Pell realizan una aproximación de $\sqrt{D}$. De hecho, las soluciones de la ecuación de Pell sobre $\mathbb{Z}$ se pueden calcular completamente a partir de la fracción continua de $\sqrt{D}$. Un enfoque más relacionado con la teoría de números algebraica, estudia la relación de la ecuación de Pell con las extensiones cuadráticas de $\mathbb{Q}$. En particular nuestro objetivo es enfocarnos en las posibles aplicaciones criptográficas de la ecuación de Pell.

Un requerimiento básico para diseñar un criptosistema asimétrico es contar con un grupo que tenga una cantidad grande pero finita de elementos. Como se explica en el capítulo anterior, existen dos posibles enfoques: utilizar grupos para los cuales calcular el orden es un problema computacionalmente difícil, o utilizar grupos donde el logaritmo discreto es computacionalmente intratable. En cualquiera de los dos casos los mensajes a cifrar se codificarán como elementos de tal grupo.

Los enfoques criptográficos de la ecuación de Pell se basan en las siguientes ideas:

- Utilizar el grupo de clases de $\mathbb{Q}[\sqrt{D}]$.
- Estudiar el grupo de puntos solución de la ecuación de Pell módulo p.

En esta monografía estudiamos el segundo enfoque.

2.2. Grupo de puntos

Si K es un cuerpo, definimos $\mathcal{P}(K) = \{(x, y) \in K \times K : x^2 - Dy^2 = 1\}$ como el conjunto de puntos de la curva de Pell $\mathcal{P}$ con coordenadas en K. En particular se definirá una ley de grupo para $\mathcal{P}(K)$ y se utilizará $\mathcal{P}(\mathbb{F}_p)$ con fines criptográficos.

2.2.1. Definición de la ley de grupo

La definición de la ley de grupo tiene un origen esencialmente geométrico. Al trabajar en cuerpos finitos la intuición geométrica muchas veces deja de funcionar, pero las definiciones se aplican de la misma manera.

Para definir la ley de grupo haremos uso de la siguiente observación:

Observación. Sea r una recta en $K \times K$. Dependiendo de la pendiente de r, el conjunto de puntos de r en K podrá ser de uno de dos casos posibles:

Caso 1 $r(K) = \{(x, y) \in K \times K : y = ax + b\}$ con $a, b \in K$

Caso 2 $r(K) = \{(x, y) \in K \times K : x = c\}$ con $c \in K$

En el primer caso, la cantidad de puntos de intersección entre la curva de Pell y la recta coincidirá con la cantidad las raíces en K de la ecuación

$$x^2 - D(ax + b)^2 = 1 \tag{2.1}$$

pudiendo ocurrir:

- La ecuación 2.1 no tiene raíces en K.
- La ecuación 2.1 tiene dos raíces distintas en K.
- La ecuación 2.1 tiene una raíz doble en K.
  En este caso decimos que la recta es tangente a la curva.
- La ecuación 2.1 tiene una única raíz simple en K.
  En este caso ocurre que toda recta paralela a r (con la misma pendiente que r) cortará en a lo sumo un punto a $\mathcal{P}(K)$ y no será tangente.

  Demostración. Si 2.1 tiene un única raíz simple en K entonces es una ecuación de grado 1, o equivalentemente, $Da^2 = 1$. Una recta r' paralela a r será de la forma $r' : y = ax + b'$, y por lo tanto la ecuación de la intersección $x^2 - D(ax + b')^2 = 1$ será también de grado 1. □

En el segundo caso, la cantidad de puntos de intersección coincidirá con la cantidad de raíces en K de la ecuación:

$$c^2 - Dy^2 = 1 \tag{2.2}$$

Estas ecuación puede o bien no tener raíces en K, o bien tener dos raíces distintas, o bien una raíz doble. En este último caso la recta se dice tangente a la curva.

Corolario. *Dada una recta r por dos puntos de $\mathcal{P}(K)$, la paralela a r por un punto Q de $\mathcal{P}(K)$ verifica una de las dos propiedades:*

- *Es tangente a $\mathcal{P}$*
- *Corta a $\mathcal{P}$ en dos puntos de $\mathcal{P}(K)$.*

Definición 2.2.1 (Ley de grupo en $\mathcal{P}(K)$)**.** Sea $N = (1, 0) \in \mathcal{P}(K)$

- Si $P \neq Q \in \mathcal{P}(K)$, se considera la recta paralela a $\overline{PQ}$ por N y se define $P + Q$ como el segundo punto de intersección con la cónica. Si la recta paralela es tangente en N se define $P + Q = N$.
- Para definir $2P = P + P$ se considera la tangente a $\mathcal{P}$ por P y la paralela por N, siendo $2P$ el segundo punto de intersección.

El corolario anterior justifica que la definición es correcta.

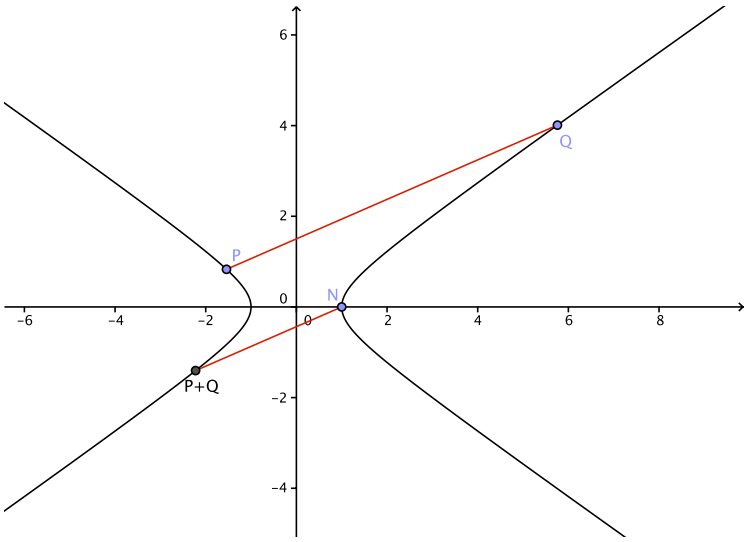


Figura 2.1: Interpretación geométrica de la ley de grupo cuando $K = \mathbb{R}$

Para demostrar que $(\mathcal{P}(K), +)$ es un grupo abeliano será necesario observar:

- Conmutatividad de $+$
- Asociatividad de $+$
- Existencia de un elemento neutro (N)
- Existencia del opuesto

La asociatividad es la única propiedad que a priori no se verifica trivialmente. Es posible demostrarla a través de la expresión algebraica de la suma de puntos, utilizando la ecuación de la recta y calculando intersecciones explícitamente con la cónica. Sin embargo, elegimos otra demostración que si bien es en cierto sentido más sofisticada, ya que utiliza técnicas de geometría algebraica, permite entender más globalmente la operación.

Un pilar fundamental para esta demostración es el teorema clásico de Pascal. Este teorema, originalmente de la geometría euclídea, tiene demostraciones elementales para $\mathbb{R}$. Sin embargo, la motivación criptográfica requiere el estudio de estas curvas sobre cuerpos más generales, en particular sobre cuerpos finitos. Si bien la demostración es sencilla, se basa en conceptos de geometría algebraica que se desarrollan en el capítulo 3, es por esto que la demostración de este teorema se realiza en el apéndice A

Teorema 2.2.1 (Pascal)**.** Sean A, B, C, D, E, F seis puntos sobre una cónica. Si $\overline{AB} \parallel \overline{DE}$ y $\overline{BC} \parallel \overline{EF}$, entonces $\overline{CD} \parallel \overline{FA}$.

Demostración. Ver apéndice A. □

Teorema 2.2.2 (Asociatividad de $+$)**.** Sean $P, Q, R \in \mathcal{P}(\mathbb{F}_p)$, entonces $(P + Q) + R = P + (Q + R)$

Demostración. Basta considerar los puntos $P, Q, R, P + Q, N, Q + R$ sobre la curva $\mathcal{P}(\mathbb{F}_p)$. Para ver que estamos en las hipótesis del teorema de Pascal basta observar:

- $\overline{PQ} \parallel \overline{(P + Q)N}$
- $\overline{QR} \parallel \overline{N(Q + R)}$

El teorema de Pascal afirma que $\overline{P(Q + R)} \parallel \overline{(P + Q)R}$, lo que implica que las paralelas por N a ambas rectas coinciden, y por lo tanto

$$(P + Q) + R = P + (Q + R)$$

□

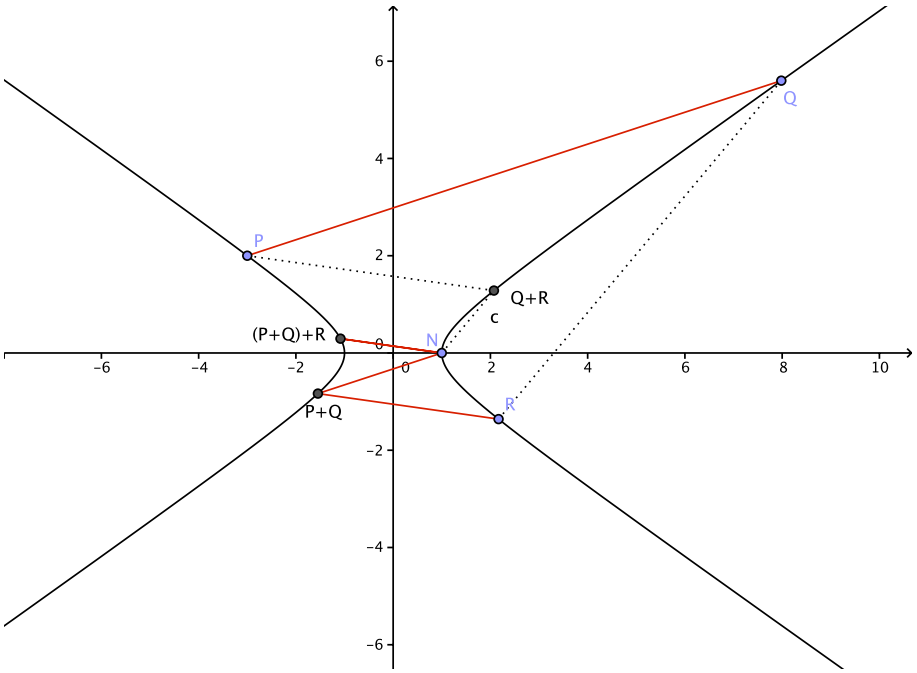


Figura 2.2: Interpretación geométrica de la asociatividad cuando $K = \mathbb{R}$

2.2.2. Descripción algebraica de la ley de grupo

Proposición. *La operación definida anteriormente es equivalente a esta otra definición:*
Si $P = (r, s)$ *y* $Q = (t, u)$ *ambos puntos en* $\mathcal{P}(K)$*, entonces definimos* $P + Q = (rt + Dsu, ru + st)$

Demostración. Observar que $P + Q \in \mathcal{P}(K)$ (verifica la ecuación de Pell) y que si $N = (1, 0)$, la pendiente de $\overline{PQ}$ coincide con la pendiente de $\overline{N(P + Q)}$. □

2.2.3. Estructura del grupo $\mathcal{P}(\mathbb{F}_p)$

En esta sección caracterizamos la estructura del grupo de puntos de una cónica de Pell sobre $\mathbb{F}_p$, con la finalidad de utilizar este grupo con objetivos criptográficos en la siguiente sección.

Proposición (Cantidad de puntos de $\mathcal{P}(\mathbb{F}_p)$)**.** *Sea* p *un primo impar. Consideramos la cónica de Pell* $\mathcal{P} : x^2 - Dy^2 = 1$*, donde* $p \nmid D$*. Entonces:*

$$\#\mathcal{P}(\mathbb{F}_p) = \begin{cases} p + 1, & si \left(\frac{D}{p}\right) = -1 \\ p - 1, & si \left(\frac{D}{p}\right) = 1 \end{cases}$$

Demostración. Consideramos rectas por $P = (1, 0)$. Cada una de estas rectas cortará a la cónica en a lo sumo 2 puntos: P y P_m, dependiendo de la pendiente de la recta: m. Las rectas por P son $L : y = m(x-1)$ y $x = 1$. Para hallar los puntos de intersección entre la recta y la cónica sustituimos la ecuación de la recta en la ecuación de Pell, obteniendo:

$$\begin{aligned} 0 &= x^2 - Dy^2 - 1 \\ &= x^2 - Dm^2(x-1)^2 - 1 \\ &= (x-1)(x+1-(x-1)Dm^2) \end{aligned}$$

Obtenemos de esta forma dos puntos en la intersección de la recta y la cónica de Pell: $P = (1, 0)$ y $P_m = \left(\frac{Dm^2+1}{Dm^2-1}, \frac{2m}{Dm^2-1}\right)$.

Es claro que de esta forma alcanzamos todos los puntos de $\mathcal{P}$, ya que dado $Q \in \mathcal{P}(\mathbb{F}_p)\backslash\{P\}$ basta considerar m la pendiente de $\overline{PQ}$.

Recíprocamente, todo valor de $m \in \mathbb{F}_p$ resulta un punto de la curva excepto si el denominador $Dm^2 - 1 = 0$. Entonces:

- Si $\left(\frac{D}{p}\right) = -1$ no ocurre que $Dm^2 - 1 = 0$, entonces para todo $m \in \mathbb{F}_p$ la recta L y la curva $\mathcal{P}(\mathbb{F}_p)$ se cortan en 2 puntos distintos. El total son $p + 1$ puntos.
- Si $\left(\frac{D}{p}\right) = 1$ hay dos valores de m que cumplen que $Dm^2 - 1 = 0$, entonces podemos encontrar $p - 2$ distintos de P, siendo en total $p - 1$ puntos.

□

Teorema 2.2.3. El grupo de puntos $(\mathcal{C}_p, +) := (\mathcal{P}(\mathbb{F}_p), +)$ es un grupo cíclico

Demostración. Consideramos dos casos:

- Si $D = a^2 \mod p$, consideramos el siguiente isomorfismo entre el grupo de puntos de la curva, y el grupo multiplicativo $\mathbb{F}_p^\times$:

$$\begin{aligned} \psi : (\mathcal{C}_p, +) &\to (\mathbb{F}_p^\times, \cdot) \\ (1, 0) &\mapsto 1 \\ (x, y) &\mapsto x - ay \mod p \end{aligned}$$

  Para verificar que efectivamente es un homomorfismo basta observar que $\psi(P + Q) = \psi(P) \cdot \psi(Q)$.

Por otro lado, para observar que el homomorfismo definido anteriormente es de hecho un isomorfismo, basta definir ψ^{-1}:

$$\begin{aligned} \psi^{-1} : (\mathbb{F}_p^{\times}, \cdot\,) &\to (\mathcal{C}_p, +) \\ u &\mapsto \Big(\frac{u+u^{-1}}{2}, \frac{u-u^{-1}}{2a}\Big) \end{aligned}$$

Entonces concluimos que $(\mathcal{C}_p, +) \simeq (\mathbb{F}_p^{\times}, \cdot\,)$ y por lo tanto el grupo de puntos de la cónica de Pell con $D = a^2 \mod p$ es un grupo cíclico de orden $p-1$.

- Si D no es un cuadrado módulo p, consideramos $\mathbb{F}_{p^2} = \mathbb{F}_p[\sqrt{D}]$ y consideramos el siguiente homomorfismo:

$$\begin{aligned} \psi : (\mathcal{C}_p, +) &\to (\mathbb{F}_{p^2}^{\times}, \cdot\,) \\ (x, y) &\mapsto x + y\sqrt{D} \mod p \end{aligned}$$

En este caso, ψ no es sobreyectivo. De hecho, la imagen de ψ es $\{x + y\sqrt{D} : x^2 - Dy^2 = 1\}$, es decir los elementos de norma 1. En particular es un grupo cíclico, por ser un subgrupo finito del grupo multiplicativo de un cuerpo.

□

2.2.4. Estructura del grupo $\mathcal{P}(\mathbb{Z}_N)$

Con el objetivo de definir el criptosistema RSA en una curva de Pell, trabajaremos con $\mathcal{P}(\mathbb{Z}_N)$ donde $N = p\,q$ con p, q primos distintos. Hasta ahora la ley de grupo está definida para $\mathcal{P}(K)$ con K un cuerpo, pero la estructura de $\mathbb{Z}_N$ nos permitirá definir en forma natural una estructura de grupo para el conjunto de puntos sobre este anillo.

$$\mathcal{P}(\mathbb{Z}_N) = \{(x, y) \in \mathbb{Z}_N \times \mathbb{Z}_N : x^2 - Dy^2 = 1\}$$

Consideramos $\mathbb{Z}_N \simeq \mathbb{F}_p \oplus \mathbb{F}_q = \{(r, s) : r \in \mathbb{F}_p,\, s \in \mathbb{F}_q\}$ y la biyección inducida entre $\mathcal{P}(\mathbb{Z}_N)$ y $\mathcal{P}(\mathbb{F}_p \oplus \mathbb{F}_q)$.

Sea $(x, y) \in \mathcal{P}(\mathbb{F}_p \oplus \mathbb{F}_q)$ entonces $x = (r_1, s_1)$, $y = (r_2, s_2)$ verifican

$$(r_1, s_1)^2 - d(r_2, s_2)^2 = (1, 1)$$

o equivalentemente

$$\begin{aligned} r_1^2 - dr_2^2 &= 1 \\ s_1^2 - ds_2^2 &= 1 \end{aligned}$$

lo cual es condición necesaria y suficiente para que $(r_1, r_2) \in \mathcal{P}(\mathbb{F}_p)$ y $(s_1, s_2) \in \mathcal{P}(\mathbb{F}_q)$.

Teorema 2.2.4. Sea $\mathcal{P} : x^2 - Dy^2 = 1$. Si R y S son dos anillos con unidad tal que $\mathcal{P}(R)$ y $\mathcal{P}(S)$ son grupos abelianos. Entonces $\mathcal{P}(R \oplus S)$ es un grupo y

$$\mathcal{P}(R \oplus S) \simeq \mathcal{P}(R) \oplus \mathcal{P}(S)$$

Demostración. Consideramos

$$\begin{array}{rcl} \psi : \mathcal{P}(R) \oplus \mathcal{P}(S) & \to & \mathcal{P}(R \oplus S) \\ \Big((r_1, r_2), (s_1, s_2)\Big) & \mapsto & \Big((r_1, s_1), (r_2, s_2)\Big) \end{array}$$

Es claro que ψ es una biyección. Veamos que la estructura de grupo inducida por ψ en $\mathcal{P}(R \oplus S)$ coincide con la definición algebraica de la suma de la sección 2.2.2, y por lo tanto, coincide con la definición geométrica.

Sean

$$\begin{array}{ll} P_1 = (r_1, r_2), \in \mathcal{P}(R) & P_2 = (s_1, s_2) \in \mathcal{P}(S) \\ Q_1 = (r_1', r_2') \in \mathcal{P}(R) & Q_2 = (s_1', s_2') \in \mathcal{P}(S) \end{array}$$

Calculamos $(P_1, P_2) + (Q_1, Q_2) \in \mathcal{P}(R) \oplus \mathcal{P}(S)$ según la definición algebraica de la sección 2.2.2:

$$(P_1, P_2) + (Q_1, Q_2) = \Big((r_1 r_1' + D r_2 r_2', r_1 r_2' + r_2 r_1'), (s_1 s_1' + D s_2 s_2', s_1 s_2' + s_2 s_1')\Big)$$

Aplicando ψ obtenemos $\psi\Big((P_1, P_2) + (Q_1, Q_2)\Big)$:

$$\Big((r_1 r_1' + D r_2 r_2', s_1 s_1' + D s_2 s_2'), (r_1 r_2' + r_2 r_1', s_1 s_2' + s_2 s_1')\Big)$$

Por otro lado,

$$\begin{array}{c} \psi\Big((P_1, P_2)\Big) + \psi\Big((Q_1, Q_2)\Big) = \Big((r_1, s_1), (r_2, s_2)\Big) + \Big((r_1', s_1'), (r_2', s_2')\Big) = \\ \Big((r_1, s_1)(r_1', s_1') + D(r_2, s_2)(r_2', s_2'), (r_2, s_2)(r_1', s_1') + (r_1, s_1)(r_2', s_2')\Big) = \\ \Big((r_1 r_1' + D r_2 r_2', s_1 s_1' + D s_2 s_2'), (r_1 r_2' + r_2 r_1', s_1 s_2' + s_2 s_1')\Big) \end{array}$$

Lo que prueba el teorema. □

Corolario (Estructura de $\mathcal{P}(\mathbb{Z}_N)$)**.** *Si $N = pq$ con p, q primos impares distintos, tal que p ni q divide a D, entonces $\mathcal{P}(\mathbb{Z}_N) \simeq \mathcal{P}(\mathbb{F}_p) \oplus \mathcal{P}(\mathbb{F}_q)$ es un grupo de orden $(p - \left(\frac{D}{p}\right))(q - \left(\frac{D}{q}\right))$*

Alice	Eve	Bob
elige dos primos grandes p y q; calcula $N = pq$; elige e coprimo con $\Phi(N)$ clave pública: (N, e, D);		
	$\xrightarrow{(N,e,D)}$	
		codifica su mensaje a un punto de la curva: $m \mapsto P$; encripta el mensaje como $C = eP$;
	$\xleftarrow{C}$	
conociendo que $de \equiv 1 \mod \Phi(N)$ computa $P = dC$ y decodifica $P \mapsto m$		

Cuadro 2.1: Aplicación de RSA para curvas de Pell

2.3. Criptografía en cónicas de Pell

En esta sección hacemos referencia a posibles aplicaciones criptográficas de las cónicas de Pell.

2.3.1. RSA sobre curvas de Pell

Alice elige un entero D y dos primos $p \neq q$. Si $N = pq$ el orden del grupo de puntos de la curva módulo N será $\Phi(N) := (p - \left(\frac{D}{p}\right))(q - \left(\frac{D}{q}\right))$. De esta forma se construye un grupo cuyo orden se presume difícil de calcular (ya que calcular $\Phi(N)$ permite calcular la factorización de N en tiempo polinomial y viceversa). Esta construcción permite desarrollar un algoritmo que utilice el orden del grupo como *trapdoor*, análogo a RSA, que se describe en el cuadro 2.1.

La idea de usar el grupo de puntos de una curva en un esquema análogo a RSA fue propuesta originalmente para curvas elípticas en el criptosistema KMOV [Koy91]. Sin embargo, el criptosistema KMOV es vulnerable a ciertos ataques que RSA no lo es [Ble97].

A priori la utilización de este esquema en curvas de Pell no parece ser

razonable en la práctica, ya que para enviar un punto de la cónica (x, y) es necesario utilizar el doble de bits que trabajando con enteros módulo N, sin incrementarse la seguridad del sistema. Este problema se puede contrarrestar con cierto compromiso entre cómputo y ancho de banda:

Si $(x_0, y_0) \in \mathcal{P}(Z_N)$, conociendo x_0 se puede calcular en tiempo polinomial $\{y \in \mathbb{Z}_N : x_0^2 - Dy^2 = 1\} = \{y_0, -y_0\}$. De esta forma se podría evitar enviar el doble de bits, enviando x_0 y un bit extra que determine si el punto corresponde a (x_0, y_0) o a $(x_0, -y_0)$. Esta solución tiene como desventaja que el receptor tendrá el costo de calcular y_0.

Por otro lado, una posible ventaja que podría estar ofreciendo este esquema con respecto al esquema clásico de RSA es que la implementación de la ley de grupo sea considerablemente rápida, según plantean [Pad02] y [Che07]. Una discusión sobre la eficiencia de un algoritmo frente a otro debe tener en cuenta distintos aspectos, como por ejemplo las características de posibles implementaciones por hardware y por software; y las posibles optimizaciones particulares, que podrán depender la curva de Pell utilizada.

2.3.2. ElGamal sobre curvas de Pell

El sistema criptográfico ElGamal basa su seguridad en la dificultad de resolver el logaritmo discreto en un grupo, clásicamente $F_p^\times$. En esta aplicación particular, el grupo consiste en el conjunto de puntos de la cónica de Pell sobre F_p con la operación definida anteriormente.

En la sección 2.2.3 se estudió la estructura del grupo $\mathcal{P}(F_p)$, diferenciándose dos casos:

- Si $(\frac{D}{p}) = 1$ el grupo es isomorfo al grupo multiplicativo $F_p^\times$, y además el isomorfismo se calcula en tiempo polinomial, de acuerdo al teorema 2.2.4. En este caso, el criptosistema sería esencialmente el esquema clásico de ElGamal.

  Desde el punto de vista de la teoría, el algoritmo ElGamal sobre cuerpos finitos supone una visión clásica de la criptografía, mientras esta visión análoga con curvas de Pell supone una perspectiva geométrica del mismo grupo. En la práctica, la diferencia entre ambos criptosistemas radica en la implementación de la operación de grupo únicamente.

- Si $(\frac{D}{p}) = -1$ el grupo es isomorfo al subgrupo multiplicativo de orden $p + 1$ de F_{p^2}.

  Utilizar criptosistemas basados en logaritmo discreto en extensiones cuadráticas F_{q^2} fue sugerido por el mismo ElGamal cuando propuso el problema del logaritmo discreto en criptografía [ElG85]. Entre los crip-

tosistemas que utilizan el logaritmo discreto en extensiones cuadráticas de cuerpos primos, hay uno llamado LUC que trabaja con ciertas funciones llamadas *funciones de Lucas* en extensiones cudráticas de cuerpos [Smi93].

Según Lenstra y Verheul [Len00], los subgrupos multiplicativos de $\mathbb{F}_{p^n}$ se supone que proveen una seguridad comparable a la seguridad de $\mathbb{F}_{p^n}$ pero en lugar de requerirse n elementos de $\mathbb{F}_p$ para representarlos, en algunos subgrupos basta con $\phi(n)$ elementos [Rub03]. Retomaremos este tema en la siguiente sección.

En el año 2000 [Len00] planteó un criptosistema donde se sugería esta forma compacta de representar los elementos de subgrupos multiplicativos de cuerpos finitos. Este criptosistema llamado XTR por *Efficient and Compact Subgroup Trace Representation* estaba planteado específicamente para $\mathbb{F}_{p^6}$.

En el año 2003 Karl Rubin y Alice Silverberg propusieron una familia de grupos algebraicos con interés criptográfico que tiene a LUC, a XTR, y al grupo de puntos de las cónicas de Pell como caso particular [Rub03]. Estos grupos, llamados *toros algebraicos* serán explicados en la siguiente sección.

Alice	Eve	Bob
elige $a \in \mathbb{F}_p^\times$; si G es un generador del grupo de puntos $A = aG$ es la clave pública;		
	$\xrightarrow{A}$	
		codifica su mensaje a un punto de la curva: $m \mapsto P$; elige $b \in \mathbb{F}_p^\times$ al azar $B = bG$ encripta el mensaje como $C = P + bA$;
	$\xleftarrow{B, C}$	
$P = C + (p - a)B \mod p$ $P \mapsto m$		

Cuadro 2.2: Aplicación de ElGamal para curvas de Pell

2.4. Criptografía basada en toros algebraicos

Basamos esta exposición en el artículo Torus-Based Cryptography, en el que Rubin y Silverberg introducen el concepto de la criptografía basada en toros algebraicos [Rub03].

La filosofía de la criptografía basada en toros algebraicos es *obtener toda la seguridad de F_{p^n} requiriendo transmitir $\phi(n)$ elementos de $\mathbb{F}_p$*. Esta idea se vuelve criptográficamente interesante cuando el cociente $n/\phi(n)$ es "grande" (como $n = 2, 6, 30, 210$) y existe una forma de representar los elementos del toro con $\phi(n)$ elementos de $\mathbb{F}_p$. Sin embargo, según se explica más adelante, para que tal representación exista, el toro algebraico debe ser racional, y en general no se sabe si el toro algebraico de dimensión $\phi(n)$ es racional para todo n.

Sea $\mathbb{G}_m$ el grupo multiplicativo: $G_m(K) = \{(x, y) \in K^2 : xy = 1\} \simeq K^\times$. Los toros algebraicos son generalizaciones de $\mathbb{G}_m$ en el sentido de la siguiente definición:

Definición 2.4.1 (Toro algebraico)**.** Un toro algebraico T sobre $\mathbb{F}_q$ es un grupo algebraico definido sobre $\mathbb{F}_q$ que sobre alguna extensión finita es isomorfo a $(\mathbb{G}_m)^d$, donde d es la dimensión de T como grupo algebraico. Si T es isomorfo a $(\mathbb{G}_m)^d$ sobre $\mathbb{F}_{q^n}$ se dice que $\mathbb{F}_{q^n}$ descompone T.

Para definir los toros algebraicos utilizados con fines criptográficos, es necesario definir la norma relativa a una extensión de cuerpos. En particular, en este caso basta definirla para extensiones de Galois como son $\mathbb{F}_{q^n}|\mathbb{F}_q$.

Definición 2.4.2. Si $L|K$ es una extensión de Galois, la norma $N_{L|K}$ de un elemento α de L es el producto de todos los conjugados a α por los automorfismos de $G_{L|K}$, es decir

$$N_{L|K}(\alpha) = \prod_{g \in G_{L|K}} g(\alpha)$$

Teorema 2.4.1. Existe un toro algebraico de dimensión $\phi(n)$ cuyos puntos F_q-racionales coinciden con el conjunto

$$T_n(\mathbb{F}_q) = \{\alpha \in \mathbb{F}_{q^n} : N_{L|K}(\alpha) = 1 \text{ para todo } K \text{ tal que } \mathbb{F}_q \subset K \subsetneq \mathbb{F}_{q^n}\}$$

La demostración de este teorema implica trabajar con la restricción de escalares de Weil y escapa al alcance de esta monografía. A partir del teorema anterior, en [Rub03] se prueba el siguiente lema:

Lema. *$T_n(\mathbb{F}_q)$ es isomorfo al subgrupo multiplicativo de $\mathbb{F}_{q^n}^\times$ de orden $\Phi_n(q)$ donde Φ_n es en n-ésimo polinomio ciclotómico.*

El n-ésimo polinomio ciclotómico es el polinomio $\Phi_n(x) = \prod_{\xi}(x-\xi)$ donde ξ son las raíces primitivas n-ésimas de la unidad del cuerpo. Como en un grupo multiplicativo existen $\phi(d)$ elementos de orden d para todo $d|n$, el grado del polinomio Φ_n es $\phi(n)$.

Observación. En particular, el lema implica que $\#T_n(\mathbb{F}_q) = \Phi_n(q)$, por ejemplo, para $n = 2, \quad \Phi_2(q) = q + 1$. En la sección 2.2.3 caracterizamos la estructura del grupo $\mathcal{P}(\mathbb{F}_p)$ y si D no es un cuadrado en $F_q^\times$ observamos que este grupo coincide con $T_2(\mathbb{F}_q)$.

Definición 2.4.3. Sea T un toro algebraico sobre F_q de dimensión d. Se dice que T es racional sí y sólo si existe un mapa birracional $\rho : T \to \mathbb{A}^d$ definido sobre $\mathbb{F}_q$. En otras palabras, T es racional sí y sólo si dada una inmersión de T en un espacio afín $\mathbb{A}^t$, existen abiertos de Zariski $W \subset T$ y $U \subset \mathbb{A}^d$ y funciones racionales $\rho_1, \ldots, \rho_d \in \mathbb{F}_q(x_1, \ldots, x_t)$ y $\psi_1, \ldots, \psi_t \in \mathbb{F}_q(y_1, \ldots, y_d)$ tal que $\rho = (\rho_1, \ldots, \rho_d) : W \to U$ y $\psi = (\psi_1, \ldots, \psi_t) : U \to W$ son isomorfismos inversos.

En este caso se dice que ρ es una parametrización racional de T.

La parametrización racional da una representación compacta de $T(\mathbb{F}_q)$, donde todo elemento de $W(\mathbb{F}_q)$ se puede representar con d elementos de $\mathbb{F}_q$ (correspondientes a sus coordenadas en $\mathbb{A}^d$).

Voskresenskii conjeturó que $T_n(\mathbb{F}_q)$ es racional para todo n. Sin embargo, esta conjetura está demostrada únicamente para n potencia de primo o producto de dos primos, por lo que por ahora no es posible utilizar los toros algebraicos $T_{30}(\mathbb{F}_q)$ ni $T_{210}(\mathbb{F}_q)$ con fines criptográficos de esta forma. Sí existen parametrizaciones racionales para $T_2(\mathbb{F}_q)$ y para $T_6(\mathbb{F}_q)$. En la sección 2.4.1 se muestra una parametrización explícita de $T_2(\mathbb{F}_q)$. La parametrización de $T_6(\mathbb{F}_q)$ es calculada en [Rub03].

2.4.1. Parametrización explícita de T_2

Sea $q \neq 2^k$, consideramos $F_{q^2} = \mathbb{F}_q(\sqrt{d})$ tal que $d \in \mathbb{F}_q^\times$ no es un cuadrado. Sea σ el automorfismo no trivial de $\mathbb{F}_{q^2}|F_q$, entonces $\sigma(\sqrt{d}) = -\sqrt{d}$.

Sea $\psi : \mathbb{A}^1(\mathbb{F}_q) \to T_2(\mathbb{F}_q)$ definida tal que:

$$\psi(a) = \frac{a+\sqrt{d}}{a-\sqrt{d}} = \frac{a^2+d}{a^2-d} + \frac{2a}{a^2-d}\sqrt{d}$$

.

Si $\beta = \beta_1 + \beta_2\sqrt{d} \in T_2(\mathbb{F}_q)$ tal que $\beta \neq \pm 1$, es decir, $\beta_2 \neq 0$, entonces

$$\beta = \frac{1+\beta}{1+\sigma(\beta)} = \psi\left(\frac{1+\beta_1}{\beta_2}\right)$$

Por lo tanto, si definimos $\rho : \mathbb{A}^1 - \{0\} \to T_2 - \{\pm 1\}$ tal que

$$\rho(\beta) = \frac{1 + \beta_1}{\beta_2}$$

Resulta que ρ y ϕ son isomorfismos inversos, y ρ es un parametrización racional de $T_2(\mathbb{F}_q)$.

Estos mapas se extienden a un isomorfismo $T_2(\mathbb{F}_q) \to \mathbb{F}_q \cup \{\infty\}$ donde $\psi(1) = \infty$, $\psi(-1) = 0$ donde

$$\text{Si } a, b \in \mathbb{F}_q, a \neq -b, \text{ entonces } \psi(a)\psi(b) = \psi\left(\frac{ab + d}{a + b}\right)$$

De esta forma, la multiplicación de T_2 se puede implementar a través del mapa $(a, b) \mapsto \dfrac{ab + d}{a + b}$, y utilizar los algoritmos clásicos descritos en el capítulo 1.

Capítulo 3

Curvas algebraicas y su aplicación a la criptografía

El objetivo de este capítulo es profundizar en la dirección geométrica del capítulo anterior. Se introducirán objetos de la de geometría algebraica de los que será posible obtener otros grupos con aplicaciones a la criptografía, que resultan, en cierto sentido, generalizaciones de las curvas de Pell.

En la sección 3.1 se introduce la definición de variedad algebraica, siendo las curvas algebraicas, variedades algebraicas de dimensión 1. A partir de las curvas algebraicas se construye el grupo de Picard, al cual se le da una estructura de variedad abeliana llamada jacobiana de la curva. Los temas de esta sección están desarrollados en [Mil09] y [Ivo08].

Una variedad abeliana es una variedad algebraica proyectiva y un grupo al mismo tiempo. Las variedades abelianas sobre cuerpos finitos son particularmente interesantes para la criptografía. El por qué del interés criptográfico sobre este tipo de grupos se resume en una cita al artículo [Rub03], mencionado en el capítulo anterior, donde se introduce la criptografía basada en toros algebraicos. En la introducción de este artículo se plantea la siguiente idea:

> *Lo que hace que los criptosistemas basados en el logaritmo discreto funcionen es que se basan en la matemática de grupos algebraicos. Un grupo algebraico es un grupo y una variedad algebraica al mismo tiempo. La estructura de grupo permite la multiplicación y exponenciación. La estructura de variedad permite expresar todos los elementos y operaciones en términos de polinomios, y por lo tanto, en una forma que puede ser eficientemente manejada por una computadora.*

En particular, las variedades abelianas son grupos algebraicos; y si bien la cita apunta a criptosistemas de una naturaleza más algebraica y menos geométrica (como la criptografía basada en toros del capítulo anterior), la filosofía se aplica también en este contexto.

En las secciones 3.2 y 3.3 se estudia la variedad jacobiana correspondiente a curvas elípticas y curvas hiperelípticas respectivamente. Referencias en estos temas son [Ivo10] y [Mil06] para curvas elípticas y [Sch03] para curvas hiperelípticas.

Por último, en la sección 3.4 se desarrollan las características generales de la criptografía basada en curvas elípticas y curvas hiperelípticas. El estudio en profundidad de las aplicaciones criptográficas es un área muy grande con problemas de ingeniería muy diversos. El objetivo de esta sección es presentar un pantallazo general de un área que tiene muchas direcciones en las que se puede profundizar.

3.1. Curvas algebraicas y variedad Jacobiana

En esta sección se definen variedades proyectivas y afines sobre un cuerpo perfecto K, es decir, un cuerpo donde toda extensión finita es separable, como es el caso de cuerpos con característica cero o cuerpos finitos. Denotaremos $\overline{K}$ a la clausura algebraica de K y si L es una extensión de K tal que $K \subseteq L \subseteq \overline{K}$, el grupo de Galois $G_{\overline{K}|L}$ se denotará G_L.

Con el objetivo de definir el espacio proyectivo de dimensión n sobre un cuerpo perfecto K será necesario definirlo sobre su clausura algebraica y utilizar la teoría de Galois para obtener un conjunto con estructura K-racional.

3.1.1. Espacio proyectivo

Definición 3.1.1 (Espacio proyectivo $\mathbb{P}^n(\overline{K})$)**.** Sea $\overline{K}$ un cuerpo algebraicamente cerrado, se define el espacio proyectivo de dimensión n:

$$\mathbb{P}^n(\overline{K}) := \{(X_0 : X_1 : \ldots : X_n) | X_i \in \overline{K}, \text{ al menos un } X_i \neq 0\} / \sim$$

con $\sim$ la relación de equivalencia:

$$(X_0 : X_1 : \ldots : X_n) \sim (Y_0 : Y_1 : \ldots : Y_n) \Leftrightarrow \exists \lambda \in \overline{K} \text{ tal que } X_i = \lambda Y_i \ \ \forall i$$

Cada una de estas clases de equivalencia son los puntos del espacio proyectivo.

Definición 3.1.2 (Espacio proyectivo $\mathbb{P}^n(L)$)**.** Sea K un cuerpo perfecto, $\overline{K}$ su clausura algebraica, y L tal que $K \subseteq L \subseteq \overline{K}$. El grupo de Galois G_L opera sobre $\mathbb{P}^n(\overline{K})$ preservando las clases de equivalencia de $\sim$. El conjunto de los

puntos L-racionales se define como el subconjunto de $\mathbb{P}^n(\overline{K})$ que queda fijo por G_L. En términos de coordenadas es:

$$\mathbb{P}^n(L) := \{(X_0 : X_1 : \ldots : X_n) \in \mathbb{P}^n(\overline{K}) | \exists \lambda \in \overline{K}, \text{ tal que } \forall i : \lambda X_i \in L\}$$

Observación. En esta definición un punto L-racional no verifica que sus coordenadas $X_i \in L$ pero si $X_j \neq 0$ entonces ocurre que $\forall i : X_i/X_j \in L$.

El papel del cuerpo K queda más claro cuando al conjunto $\mathbb{P}^n$ (sobre cualquier cuerpo extensión de K) se lo considera con la estructura topológica dada por la topología de Zariski que se define a continuación.

Definición 3.1.3 (Topología de Zariski)**.** Un polinomio $f \in K[X_0, \ldots, X_n]$ se dice homogéneo de grado d si es la suma de monomios de grado d. Esta definición es equivalente a pedir que $f(\lambda X_0, \ldots, \lambda X_n) = \lambda^d f(X_0, \ldots, X_n)$ para todo $\lambda \in \overline{K}$. De esta forma tiene sentido la siguiente definición:

$$D_f(L) := \{P \in \mathbb{P}^n(L) : f(P) \neq 0\}.$$

Se denota $D_f := D_f(\overline{K})$

Se define la topología de Zariski en $\mathbb{P}^n(\overline{K})$ a través de la base $\{D_f\}_f$

Para describir los conjuntos cerrados en esta topología, si $I \subsetneq K[X_0, \ldots, X_n]$ es un ideal generado por polinomios homogéneos definimos

$$V_I := \{P \in \mathbb{P}^n(\overline{K}) : f(P) = 0 \ \ \forall f \in I\}$$

De esta forma un conjunto es cerrado bajo la topología de Zariski del espacio proyectivo sobre K si es cero simultáneo de polinomios homogéneos en $K[X_0, \ldots, X_n]$.

Si L es un cuerpo tal que $K \subset L \subset \overline{K}$ consideramos la topología de Zariski en L como la topología relativa, siendo $V_I(L) = V_I \cap \mathbb{P}^n(L)$.

3.1.2. Espacio afín

Definición 3.1.4 (Espacio afín $\mathbb{A}^n(\overline{K})$)**.** Si $\overline{K}$ un cuerpo algebraicamente cerrado, se define el espacio afín de dimensión n sobre K:

$$\mathbb{A}^n := \{(x_1, \ldots, x_n) : x_i \in \overline{K}\}$$

El conjunto de los puntos L-racionales, es decir, aquellos puntos de $\mathbb{A}^n(\overline{K})$ invariantes bajo la acción de automorfismos de Galois sobre $\overline{K}|L$ en cada coordenada coincide con el conjunto $\mathbb{A}^n(L) := \{(x_1, \ldots, x_n) : x_i \in L\}$ según el teorema de correspondencia de Galois.

De forma análoga al caso proyectivo se considera la topología de Zariski sobre $\mathbb{A}^n$, siendo en este caso los conjuntos cerrados V_I para $I \subsetneq K[x_1, \ldots, x_n]$ un ideal.

Observación. El teorema de la base de Hilbert afirma que todo ideal en el anillo de polinomios multivariados con coeficientes en un cuerpo es finitamente generado. Esto implica que todo conjunto cerrado bajo la topología de Zariski en $\mathbb{A}^n$ puede ser descrito como el conjunto de ceros comunes de un conjunto finito de polinomios.

3.1.3. Variedades algebraicas

Un subconjunto S de un espacio topológico se dice irreducible si no puede ser escrito como la unión de dos conjuntos cerrados ambos no vacíos. Esta noción de irreducibilidad permite definir variedad proyectiva y variedad afín de acuerdo a la topología de cada uno de los espacios de la siguiente manera:

Definición 3.1.5 (Variedad proyectiva y variedad afín)**.** V es una variedad afín si es un conjunto cerrado afín irreducible. Análogamente V es una variedad proyectiva si es un conjunto cerrado proyectivo irreducible. Se dice que una variedad está definida sobre un cuerpo K cuando los cerrados de la topología están descritos por polinomios con coeficientes en K.

Los puntos de L-racionales de una variedad V se denotarán $V(L)$.

Definición 3.1.6 (Variedad algebraica absolutamente irreducible)**.** Una variedad V afín o proyectiva sobre K se dice absolutamente irreducible si es irreducible como conjunto cerrado con respecto a la topología de Zariski sobre $\overline{K}$.

Observación. Esta última definición es equivalente a decir que V es una variedad absolutamente irreducible si $I(V)$ es un ideal primo en $\overline{K}$. Por ejemplo el ideal generado por $(x_1^2 - 2x_2^2)$ es primo en $\mathbb{Q}[x_1, x_2]$ pero no lo es en su clausura algebraica. En el resto de la exposición vamos a considerar variedades absolutamente irreducibles.

Para definir la dimensión de una variedad algebraica, observamos que tanto el espacio proyectivo como el espacio afín son Noetherianos, es decir, cualquier sucesión de conjuntos cerrados $S_1 \supseteq S_2 \supseteq \ldots$ eventualmente se estaciona. Esto ocurre porque los conjuntos cerrados corresponden a ideales de polinomios, que por el teorema de la base de Hilbert son ideales finitamente generados. Esta propiedad permite la siguiente definición de dimensión para una variedad algebraica:

Definición 3.1.7 (Dimensión de una variedad algebraica)**.** Si V es una variedad algebraica (proyectiva o afín). La dimensión $\dim(V)$ se define como

el supremo en las longitudes de todas las cadenas $S_1 \supsetneq S_2 \supsetneq \ldots \supsetneq S_n$ de subespacios cerrados absolutamente irreducibles tal que $S_i \subseteq V$.

En particular una variedad es llamada curva si es una variedad algebraica de dimensión 1.

Definición 3.1.8 (Curva plana proyectiva)**.** Una variedad algebraica se llama curva plana proyectiva si está definida por un polinomio homogéneo absolutamente irreducible $F(X, Y, Z) = 0$.

Para los fines criptográficos perseguidos, las curvas algebraicas consideradas serán no singulares. La definición de variedad no singular, en términos del ideal de polinomios que define a la variedad es la siguiente:

Definición 3.1.9 (Variedad no singular (afín))**.** Sea V una variedad afín, $P \in V$ y $f_1, \ldots, f_m \in \overline{K}[x_1, \ldots, x_n]$ un conjunto de generadores de $I(V)$. Entonces se dice que V es no singular en P si la matriz $m \times n$

$$\left(\frac{\partial f_i}{\partial X_j}(P)\right)_{i,j}$$

tiene rango $n - \dim(V)$. Si V es no singular en todo punto, entonces decimos que V es no singular.

En el caso particular de las curvas ocurre para que la curva sea no singular en un punto $P = [x : y : z]$ una condición necesaria y suficiente es que:

$$\left(\frac{\partial F}{\partial X}(P), \frac{\partial F}{\partial Y}(P), \frac{\partial F}{\partial Z}(P)\right) \neq (0, 0, 0)$$

Una variedad es no singular si lo es en todos sus puntos.

A continuación haremos una construcción que permitirá "cubrir" una variedad proyectiva con variedades afines. Permitiendo una definición natural de punto no singular de una variedad en función de las partes afines de la misma.

3.1.4. Relación entre espacios proyectivo y afín

Con el objetivo de encontrar una relación entre variedades proyectivas y afines observamos que las topologías de $\mathbb{P}^n$ y $\mathbb{A}^n$ son compatibles en el siguiente sentido:

Si $F \in K[X_0, X_1, \ldots, X_n]$ es un polinomio homogéneo de grado d. Consideramos la dehomogenización de F con respecto a la variable X_i al proceso de remplazar $F(X_0, \ldots, X_n)$ por $F_i := F(x_0, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n) \in$

$K[x_1, \ldots, x_n]$. Recíprocamente, dado un polinomio $f \in K[x_1, \ldots, x_n]$ de grado d consideramos el proceso de homogenización del polinomio con respecto a la variable X_i obteniendo

$$f_i = X_i^d f(X_0/X_i, \ldots, X_{i-1}/X_i, X_{i+1}/X_i, \ldots, X_n/X_i)$$

De esta forma existe una biyección entre ideales primos homogéneos de $K[X_0, \ldots, X_n]$ e ideales primos de $K[x_1, \ldots, x_n]$ y por lo tanto existe una compatibilidad topológica entre los espacios proyectivo y afín. De hecho, el espacio proyectivo se puede cubrir con espacios afines de la siguiente manera:

Consideramos los conjuntos abiertos $U_i = D_{X_i} \subset \mathbb{P}^n$. La función continua y abierta $\psi_i : U_i \to \mathbb{A}^n$ se construye a través de la dehomogenización del polinomio X_i con respecto a la variable X_i. La función inversa está dada por:

$$\begin{array}{rcl} \phi_i : \mathbb{A}^n & \to & U_i \\ (x_1, \ldots, x_n) & \mapsto & (x_1 : \ldots : x_i : 1 : x_{i+1} : \ldots : x_n) \end{array}$$

Entonces para cada $i = 0, \ldots, n$ existe una biyección canónica entre U_i y $\mathbb{A}^n$ que es un homeomorfismo.

Los conjuntos $U_0, \ldots, U_n$ son el cubrimiento estándar de $\mathbb{P}^n$, y los mapas ϕ_i pueden verse como la inclusión $\mathbb{A}^n \subset \mathbb{P}^n$.

Esta misma construcción permite cubrir una variedad proyectiva con variedades afines. Para ello consideramos V un conjunto cerrado proyectivo tal que $V = V_{I(V)}$ donde $I(V)$ es un ideal homogéneo en $\overline{K}[X_0, \ldots, X_n]$. Sea $V_i = \phi^{-1}(V \cap U_i)$. Este conjunto es un cerrado afín cuyo ideal es obtenido mediante la dehomogenización de todos los polinomios de $I(V)$ respecto a la variable X_i. Entonces $\phi_i(V_i) : \ i = 0, \ldots, n$ es un cubrimiento por cerrados afines de V.

El proceso inverso implica considerar un conjunto cerrado afín $V_I \subset \mathbb{A}^n$ y el encaje $V_i \subset \mathbb{A}^n \overset{\phi_i}{\to} \mathbb{P}^n$. La clausura proyectiva $\overline{V}_I$ de V_I es el conjunto cerrado proyectivo definido por el ideal $\overline{I}$ generado por la homogenización de los polinomios $\{f_i : f \in I\}$. Los puntos agregados para obtener la clausura proyectiva son llamados "puntos del infinito".

Ambas construcciones se resumen en el siguiente lema:

Lema. *Sea $\phi_i : \mathbb{A}^n \to \mathbb{P}^n$ un encaje canónico donde identificamos $\mathbb{A}^n$ con su imagen por ϕ_i. Sea $V \subseteq \mathbb{A}^n$ una variedad afín. Entonces $\overline{V}$ es una variedad proyectiva y $V = \overline{V} \cap \mathbb{A}^n$.*

Por otro lado, si $V \subseteq \mathbb{P}^n$ es una variedad proyectiva, entonces $V \cap \mathbb{A}^n$ es una variedad afín y o bien $V \cap \mathbb{A}^n = \emptyset$ o bien $V = \overline{V \cap \mathbb{A}^n}$. Si V es una variedad proyectiva definida sobre K, entonces $V \cap \mathbb{A}^n$ es vacío o una variedad

afín definida sobre K. Siempre existe al menos una coordenada i tal que $V_{(i)} := V \cap \phi_i(\mathbb{A}^n)$ es no vacío. En ese caso decimos que $V_{(i)}$ es una parte afín de V.

En virtud del lema anterior, podremos trabajar indistintamente con variedades proyectivas o afines, teniendo en cuenta las identificaciones construidas anteriormente.

3.1.5. Cuerpos de funciones en variedades algebraicas

Definición 3.1.10 (Cuerpo de funciones $K(V)$)**.** Sea V una variedad afín sobre K en $\mathbb{A}^n$ definida mediante cierto ideal primo I. Definimos el anillo coordenado de V:

$$K[V] := K[x_1, \ldots, x_n]/I$$

y definimos el cuerpo de funciones de V como su cuerpo de fracciones:

$$K(V) := \mathrm{Frac}(K[V]).$$

Si V es una variedad proyectiva sobre K, consideramos $V_a \subseteq \mathbb{A}^n$ una parte afín no vacía de V. Entonces el cuerpo de funciones $K(V)$ se define como $K(V_a)$.

Los elementos $f \in K(V)$ se pueden representar como fracciones de polinomios $f = g/h$ donde $g, h \in K[x_1, \ldots, x_n]$ o como fracciones de polinomios homogéneos del mismo grado $f = g/h$ donde $g, h \in K[X_0, \ldots, X_n]$. Las funciones $f_1 = g_1/h_1$ y $f_2 = g_2/h_2$ son iguales si $g_1h_2 - g_2h_1 \in I(V)$.

3.1.6. Variedades abelianas

Definición 3.1.11. Un grupo algebraico G absolutamente irreducible sobre un cuerpo K es una variedad algebraica absolutamente irreducible definida sobre K (afín o proyectiva) dotado de una operación definida sobre K bajo la cual es un grupo:

(I) Suma: un morfismo $\oplus : G \times G \to G$ asociativo definido sobre K.

(II) Elemento neutro: $0_G \in G(K)$ tal que para todo $P \in G$ se cumple que $P \oplus 0_G = 0_G \oplus P = P$

(III) Inverso: un morfismo $- : G \to G$ definido sobre K tal que para todo $P \in G$ satisface $P \oplus (-P) = (-P) \oplus P = 0_G$

Si L es una extensión del cuerpo K y $G(L)$ el conjunto de puntos L-racionales, el conjunto $G(L)$ es un grupo cuya suma e inverso se calculan a través de la evaluación de morfismos definidos sobre K.

Si G es una variedad proyectiva, resulta que la ley de grupo es necesariamente conmutativa, dando lugar a la siguiente definición:

Definición 3.1.12 (Variedad abeliana)**.** Una variedad abeliana es un grupo algebraico proyectivo.

3.1.7. Aritmética de curvas

Consideramos C una curva proyectiva no singular. En las siguientes secciones construiremos una variedad abeliana a partir de la curva C.

Divisores

Con el objetivo de conseguir un grupo a partir del conjunto de puntos de una curva primero consideramos el grupo abeliano libre generado por el conjunto de puntos de una curva algebraica C sobre un cuerpo algebraicamente cerrado $\overline{K}$ para luego obtener un conjunto de L-racional donde $K \subset L \subset \overline{K}$. Para ello partimos de las siguiente definiciones:

Definición 3.1.13 (Grupo de divisores)**.** Llamamos divisor en $C(\overline{K})$ una suma formal de la forma

$$\sum_{P \in C(\overline{K})} n_P \cdot [P]$$

donde $n_P \in \mathbb{Z}$ y n_P es distinto de cero en una cantidad finita de puntos. El grupo abeliano generado por los puntos de C se denomina grupo de divisores en $C(\overline{K})$ y se denota como $\mathrm{Div}_{\overline{K}}(C)$.

Definición 3.1.14 (Grado de un divisor)**.** El grado de un divisor D se define como el número entero:

$$\deg(D) = \sum_{P \in C(\overline{K})} n_P$$

Los divisores de grado cero forman un subgrupo en $\mathrm{Div}_{\overline{K}}(C)$ y se denota como $\mathrm{Div}^0_{\overline{K}}(C) = \{D \in \mathrm{Div}_{\overline{K}}(C) : \deg(D) = 0\}$.

El soporte de un divisor D se define como el conjunto de puntos P donde n_P es distinto de cero: $\mathrm{sop}(D) = \{P \in C(\overline{K}) : n_P \neq 0\}$. En virtud de las definiciones anteriores, el soporte de un divisor es siempre finito.

Decimos que dos divisores D y D' son coprimos si tienen soporte disjunto.

Por otro lado, un divisor se dice efectivo si $n_P \geq 0$ para todo P punto de C; y se denota $D \geq D'$ cuando el divisor $D - D'$ es efectivo.

Sea G_K el grupo de automorfismos de Galois de $\overline{K}$ sobre K. Si $\sigma \in G_K$ y $P = [X_0 : \ldots : X_n]$ consideramos $P^\sigma = [X_0^\sigma : \ldots : X_n^\sigma]$. Si $D \in \mathrm{Div}_{\overline{K}}(C)$ definimos:

$$D^\sigma = \sum_{P \in C} n_P \cdot [P^\sigma]$$

Entonces decimos que un divisor D es L-racional si $D^\sigma = D$ para todo σ en G_L.

El grupo de divisores K-racionales se denota como $\mathrm{Div}_K(C)$ y análogamente $\mathrm{Div}_K^0(C)$ es el grupo de divisores de grado cero K-racionales

Observación. A diferencia de los puntos K-racionales, los divisores K-racionales pueden ser suma de puntos cuyas coordenadas no necesariamente están en K. Por ejemplo, si consideramos $P_i = (i : 1 : 1)$, $P_{-i} = (-i : 1 : 1)$, el divisor $D = [P_i] + [P_{-i}]$ es $\mathbb{Q}$-racional, a pesar de que los puntos P_i y P_{-i} no lo son.

Valuaciones discretas

Definición 3.1.15 (Valuación discreta)**.** Una valuación discreta sobre un cuerpo F es un mapa sobreyectivo $v : F \to \mathbb{Z} \cup \{\infty\}$ tal que:

1. $v(a) = \infty$ sí y sólo si $a = 0$.
2. $v(a \cdot b) = v(a) + v(b)$.
3. $v(a + b) \geq \mathrm{mín}(v(a), v(b))$.

Ahora consideramos $F = K(C)$ el cuerpo de funciones C sobre K. A cada punto P de $C(K)$ le asociamos una valuación discreta ord_P que indica si una función $f \in K(C)$ tiene un cero o un polo en P y también indicará la multiplicidad.

Proposición. *Sea C una curva algebraica definida sobre K y sea $P \in C(\overline{K})$ un punto no singular. Entonces la función*

$$\mathrm{ord}_P : K(C) \to \mathbb{Z} \cup \{\infty\}$$

tal que

$$\mathrm{ord}_P(f) = \begin{cases} \infty & \text{si } f \equiv 0 \\ \text{el orden de anulación de } f \text{ en } P & \text{si } f \not\equiv 0 \end{cases}$$

es una valuación discreta.

Divisores principales

A cada función $f \in K(C)^*$ se le asocia un divisor definido como la suma formal

$$\operatorname{div}(f) := \sum_{P \in C(\overline{K})} \operatorname{ord}_P(f) \cdot [P]$$

que resulta ser un divisor debido a la siguiente proposición:

Proposición. *Sea C una curva algebraica no singular definida sobre K y $f \in K(C)$. Entonces existe una cantidad finita de puntos de $C(\overline{K})$ en los cuales f tiene un cero o un polo. Además si f no tiene polos, entonces f es constante.*

Definición 3.1.16 (Divisor principal)**.** Un divisor $D \in \operatorname{Div}_K(C)$ se dice principal si existe $f \in K(C)^*$ tal que $D = \operatorname{div}(f)$

Proposición (Propiedades de los divisores principales)**.** *Si C es una curva algebraica no singular definida sobre K, y $f, g \in K(C)^*$ entonces:*

1. $\operatorname{div}(f) = 0$ *sí y sólo si $f \in K^*$ (es decir, sí y sólo si f es constante).*
2. $\deg(\operatorname{div}(f)) = 0$. *Es decir, los divisores principales tienen grado cero.*
3. $\operatorname{div}(f \cdot g) = \operatorname{div}(f) + \operatorname{div}(g)$
4. $\operatorname{div}(\frac{f}{g}) = \operatorname{div}(f) - \operatorname{div}(g)$

Sea $\operatorname{Princ}_K(C) = \{D \in \operatorname{Div}_K(C) : D \text{ es principal}\}$. Entonces la proposición anterior, en particular implica que $\operatorname{Princ}_K(C)$ es un subgrupo de $\operatorname{Div}^0_K(C)$. De hecho, los puntos 1 y 4 de la proposición anterior implican que si dos funciones $f, g \in K(C)$ verifican que $\operatorname{div}(f) = \operatorname{div}(g) = D$ entonces existe $c \in K^*$ tal que $f = c \cdot g$.

El paso previo a construir la variedad Jacobiana de una curva es considerar la siguiente relación de equivalencia.

Definición 3.1.17. Sean $D_1, D_2 \in \operatorname{Div}_K(C)$. Decimos que D_1 y D_2 son linealmente equivalentes ($D_1 \sim D_2$) si $D_1 - D_2$ es un divisor principal.

Teorema de Riemann-Roch

Definición 3.1.18. Sea $D \in \operatorname{Div}_{\overline{K}}(C)$ un divisor $\overline{K}$-racional. Definimos

$$\mathcal{L}(D) = \{f \in \overline{K}(C)^* : \operatorname{div}(f) \geq -D\} \cup \{0\}$$

Proposición. *El conjunto $\mathcal{L}(D) \subseteq \overline{K}(C)$ es un $\overline{K}$-espacio vectorial de dimensión finita.*

Definición 3.1.19. Definimos $l(D)$ como la dimensión de $\mathcal{L}(D)$ como K-espacio vectorial.

El Teorema de Riemann-Roch descubre una conexión importante entre $deg(D)$ y $l(D)$. A continuación se enuncia una versión simplificada de este teorema, que resulta suficiente para el objetivo dentro de esta exposición. La versión completa involucra trabajar con formas diferenciales en curvas algebraicas y divisores canónicos.

Teorema 3.1.1 (Riemann-Roch)**.** Sea C una curva algebraica irreducible y no singular sobre un cuerpo perfecto K. Entonces existe un entero $g \leq 0$ tal que para todo divisor $D \in \mathrm{Div}(C)$ se cumple que

$$l(D) \geq \deg(D) - g + 1.$$

Además, para todo $D \in \mathrm{Div}(C)$ tal que $\deg(D) > 2g - 2$ se cumple la igualdad $l(D) = deg(D) - g + 1$.

Este teorema garantiza la existencia de funciones con ceros y polos dados, en caso de que la cantidad de ceros (con multiplicidad) sea superior a la cantidad de polos en $2g - 2$.

Grupo de Picard

En esta sección se construirá un grupo abeliano a partir del grupo de divisores, obteniendo una relación entre la aritmética de curvas y variedades abelianas.

Definición 3.1.20. Sea C una curva absolutamente irreducible no singular definida sobre un cuerpo K. Si $\mathrm{Div}^0_K(C)$ es el grupo de divisores K-racionales de grado 0, definimos el grupo de clases de divisores de grado 0 (grupo de Picard) como

$$\mathrm{Pic}^0_{\overline{K}}(C) := \frac{\mathrm{Div}^0_{\overline{K}}(C)}{\mathrm{Princ}_{\overline{K}}(C)}$$

Los elementos de $\mathrm{Pic}^0_K(C)$ son los elementos de $\mathrm{Pic}^0_{\overline{K}}(C)$ fijos por G_K, es decir $\mathrm{Pic}^0_K(C) = \left(\mathrm{Pic}^0_{\overline{K}}(C)\right)^{G_K}$. Sus elementos son clases de divisores que se denotan $[D]$.

El siguiente teorema nos indica que el grupo de Picard resulta isomorfo a una variedad abeliana:

Teorema 3.1.2. Sea C una curva algebraica absolutamente irreducible y no singular de género g definida sobre K. Entonces existe una variedad abeliana J_C definida sobre K de dimensión g y un isomorfismo de grupos

$$\varphi : \mathrm{Pic}^0_{\overline{K}}(C) \to J_C(\overline{K})$$

que preserva la acción de Galois, es decir, para todo cuerpo L tal que $K \subset L \subset \overline{K}$

$$\varphi(\mathrm{Pic}_L(C)) = J_C(L))$$

La variedad J_C es llamada Jacobiana de C. $J_C(L)$ es un subgrupo de $J_C(\overline{K})$ y sus puntos son los puntos L-racionales de J_C.

3.2. Jacobiana de una curva elíptica

La criptografía basada en curvas algebraicas tiene su representante más importante en las curvas elípticas. Una posible razón para esto es que la variedad jacobiana es la propia curva elíptica; lo que facilita la teoría, y la implementación de las operaciones del grupo. En esta sección se define curva elíptica y se demuestra que la variedad jacobiana y la curva elíptica son isomorfas como variedades algebraicas. En la sección 3.4 se mencionan sus aplicaciones criptográficas.

Definición 3.2.1 (Curva elíptica)**.** Una curva elíptica $E|K$ es una curva proyectiva irreducible no singular de género 1 definida sobre K, con al menos un punto K -racional.

Teorema 3.2.1 (Forma de Weierstrass)**.** Toda curva elíptica $E|K$ es isomorfa a una curva algebraica plana definida por una *ecuación de Weierstrass* de la forma:

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_6Z^3 \tag{3.1}$$

con $a_1,\ a_2,\ a_3,\ a_6 \in K$

Demostración. Sea $\mathcal{O}$ un punto K- racional de E. Sea $K(E)$ su cuerpo de funciones. El Teorema de Riemann-Roch 3.1.1 aplicado a curvas de genéro 1 dice que $l(D) = \deg(D)$ para todo divisor D de grado mayor o igual que 1.

Consideraremos los divisores $[\mathcal{O}]$, $2[\mathcal{O}]$, $3[\mathcal{O}]$, $4[\mathcal{O}]$, $5[\mathcal{O}]$, $6[\mathcal{O}]$, que nos permitirán extraer conclusiones sobre las curvas elípticas como curvas algebraicas.

A partir del teorema de Riemann Roch ocurre que:

- $l([\mathcal{O}]) = 1$ entonces $\mathcal{L}([\mathcal{O}]) = \overline{K}^*$ las funciones constantes no nulas.
- $l(2[\mathcal{O}]) = 2$ entonces existe una función $x \in K(E)$ tal que $\mathcal{L}(2[\mathcal{O}])$ es generado como $\overline{K}$-espacio vectorial por $\{1, x\}$.
- $l(3[\mathcal{O}]) = 3$ entonces existe una función $y \in K(E)$ tal que $\{1, x, y\}$ es base de $\mathcal{L}(3[\mathcal{O}])$ sobre $\overline{K}$.
- $l(4[\mathcal{O}]) = 4$ y como $x^2 \in \mathcal{L}(4[\mathcal{O}]) - \mathcal{L}(3[\mathcal{O}])$ entonces $\{1, x, y, x^2\}$ es base de $\mathcal{L}(4[\mathcal{O}])$.

- Análogamente $\{1, x, y, x^2, xy\}$ es base de $\mathcal{L}(5[\mathcal{O}])$.
- Por último $\langle\{1, x, y, x^2, xy, x^3, y^2\}\rangle \subset \mathcal{L}(6[\mathcal{O}])$, pero $l(6[\mathcal{O}]) = 6$ por lo tanto existe una dependencia lineal donde x^3 y y^2 aparecen con coeficientes distintos de cero.

Entonces obtenemos la siguiente dependencia lineal, con $a_6, a_7 \neq 0$:

$$A_1 + A_2x + A_3y + A_4x^2 + A_5xy + A_6x^3 + A_7y^2 = 0 \quad A_i \in K$$

La cual podemos llevar a la forma:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_6 \tag{3.2}$$

con el cambio de variable $y \mapsto y/A_7$ y multiplicar la ecuación por A_7, y luego el cambio de variable $x \mapsto x/A_6$, $y \mapsto y/A_6$ y multiplicar la ecuación por A_7^2.

La ecuación 3.2 corresponde a una curva plana afín e irreducible que resulta ser no singular ya que el cuerpo de funciones de esta curva plana coincide con el cuerpo de funciones de E que es no singular.

Ahora consideremos una función $\phi : E \to \mathbb{P}^2$ tal que $\phi(P) = [x(P) : y(P) : 1]$. Entonces ϕ define una correspondencia biyectiva y racional entre E y la curva plana C definida en la ecuación 3.1.

El único punto en donde la función podría no estar definida es en $\mathcal{O}$, pero $\phi(P) = [x(P) : y(P) : 1] = [(x/y)(P) : 1 : (1/x)(P)]$ y en particular, $\phi(\mathcal{O}) = [0 : 1 : 0]$. □

Teorema 3.2.2. Sea $E|K$ una curva elíptica. El grupo de Picard $\mathrm{Pic}^0_K(E)$ induce una estructura de grupo abeliano en $E(K)$.

Demostración. Sea $\mathcal{O}$ un punto K-racional de E. Sea

$$\begin{aligned} \phi : E &\to \mathrm{Pic}^0_K(E) \\ P &\mapsto [P] - [\mathcal{O}] \end{aligned}$$

- ϕ es sobreyectivo
  Sea D divisor de grado 0, entonces $D + [\mathcal{O}]$ es de grado 1, y por Riemann-Roch $l(D + [\mathcal{O}]) = 1$. Sea $f \in \mathcal{L}(D + [\mathcal{O}])$ un generador. Entonces
  $$\mathrm{div}(f) = \sum_P v_P(f) \cdot [P]$$
  Por definición de f ocurre que $\mathrm{div}(f) \leq D + [\mathcal{O}]$, entonces $\mathrm{div}(f) + D + [\mathcal{O}]$ es un divisor efectivo de grado 1 que entonces puedo escribir como $[P]$.
  Entonces D es equivalente a $[P] - [\mathcal{O}]$

- ϕ es inyectivo
  Sean P, Q puntos de E tal que existe una función f tal que

  $$\text{div}(f) + [P] - [\mathcal{O}] = [Q] - [\mathcal{O}]$$

  Entonces $[Q] = [P] + \text{div}(f)$ es un divisor efectivo de grado 1, entonces $div(f)$ es un divisor efectivo de grado 0, es decir, f es constante no nula. Entonces $v_R(f) = 0$ para todo $R \in E$, entonces $P = Q$.

Como ϕ es una biyección entre $E(K)$ y Pic^0_E, esta misma biyección induce una estructura de grupo abeliano en E. □

La siguiente observación justifica la definición geométrica usual de la ley grupo en una curva elíptica.

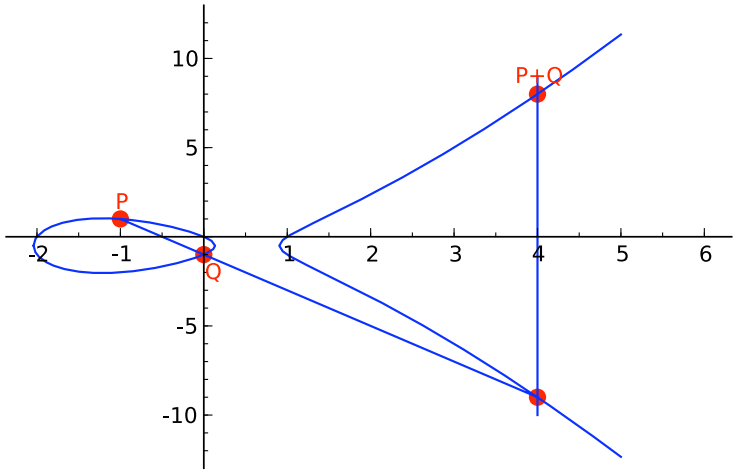


Figura 3.1: Interpretación geométrica de la ley de grupo en $E(\mathbb{R})$. El punto $\mathcal{O} = [0 : 1 : 0]$ en el plano afín se identifica con la dirección vertical

Observación (Estructura del grupo de puntos $E(K)$). Si P, Q son puntos de E, y l la recta de $\mathbb{P}^2$ por P y Q, l corta necesariamente en un tercer punto E que llamaremos R. R puede coincidir con P o Q (o no), pero necesariamente es K-racional.

Sea l' la recta de $\mathbb{P}^2$ por R y $\mathcal{O}$. $l' \cap E = \{R, \mathcal{O}, S\}$ (S podrá coincidir con R o con $\mathcal{O}$ en caso de que la recta sea tangente a la curva).

$(P - \mathcal{O}) + (Q - \mathcal{O})$ y $S - \mathcal{O}$ son divisores linealmente equivalentes:

Sea $f = \alpha X + \beta Y + \gamma Z$ que determina a l.
Sea $f' = \alpha' X + \beta' Y + \gamma' Z$ que determina l'.
Entonces f/f' es una función racional en E.

El divisor de f/f' es $(P+Q+R)-(R+S+\mathcal{O}) = (P-\mathcal{O})+(Q-\mathcal{O})-(S-\mathcal{O})$. Entonces

$$(P-\mathcal{O})+(Q-\mathcal{O}) = S-\mathcal{O}$$

□

3.3. Jacobiana de una curva hiperelíptica

En esta sección se estudia la estructura de grupo de la variedad jacobiana de la curva hiperelíptica, con el objetivo de explicar la idea detrás del algoritmo para calcular la ley de grupo.

Definición 3.3.1 (Curva hiperelíptica)**.** Una curva algebraica afín definida sobre un cuerpo K perfecto se dice hiperelíptica de género g si está definida por un polinomio de la forma:

$$y^2 + h(x)y = f(x) \quad h, f \in K[x, y] \tag{3.3}$$

tal que $\deg(h) < g$, $\deg(f) = 2g+1$, y no tiene puntos singulares en $\overline{K} \times \overline{K}$.

Observación. El polinomio 3.3 es absolutamente irreducible.

Demostración. Si existe una factorización para $y^2 + h(x)y - f(x)$, entonces es de la forma $(y - a(x))(y - b(x))$ pero $\deg_x(a\,b) = \deg_x(f) = 2g+1$, y $\deg_x(a+b) = \deg_x(h) \leq g$. Por lo tanto, tal factorización no es posible. □

Teorema 3.3.1. Sea C una curva hiperelíptica sobre K definida por el polinomio 3.3. Entonces:

1. Si $h(x) = 0$, entonces $\mathrm{car}(K) \neq 2$.
2. Si $\mathrm{car}(K) \neq 2$ existe un cambio de variables $x \mapsto x$, $y \mapsto (y - h(x)/2)$ que transforma C a una ecuación de la forma $y^2 = f(x)$ donde $\deg_x(f) = 2g+1$.
3. Sea C una ecuación como en 3.3 con $h(x) = 0$ y $\mathrm{car}(K) \neq 2$. Entonces C es una curva hiperelíptica sí y sólo si $f(x)$ no tiene raíces repetidas en $\overline{K}$.

Demostración. 1. Si $h(x) = 0$ y $\mathrm{car}(K) = 2$. Entonces la condición de no singularidad se reduce a $f'(x) \neq 0$. Sea $u \in \overline{K}$ tal que $f'(u) = 0$, y sea $v \in \overline{K}$ una raíz de la ecuación $y^2 = f(u)$. Entonces el punto (u, v) es un punto singular de C.

2. Aplicando el cambio de variable

$$(y - h(x)/2)^2 + h(x)(y - h(x)/2) = f(x)$$

que se simplifica a $y^2 = f(x) + h(x)^2/4$, y como $\deg_x(h) \leq g$ entonces $\deg_x(f + h^2/4) = 2g + 1$.

3. Un punto singular $(u, v) \in C$ debe satisfacer $v^2 = f(u)$, $2v = 0$ y $f'(u) = 0$. Entonces $v = 0$ y u es una raíz doble del polinomio $f(x)$.

□

Observación. La intersección de la curva con la recta proyectiva $Z = 0$ es un único punto "∞" que llamaremos $\mathcal{O}$. Si $g \geq 2$, el punto $\mathcal{O}$ es singular.

Definición 3.3.2. Sea $P = (u, v)$ un punto en una curva hiperelíptica C. El opuesto de P es el punto $-P = (u, -v - h(u)) \in C$. Se define $-\mathcal{O} = \mathcal{O}$.

Si P es tal que $-P = P$ se dice que P es un punto de Weierstrass. En caso contrario se dice que P es un punto ordinario.

Definición 3.3.3 (Divisor semirreducido)**.** Un divisor semirreducido es un divisor de la forma

$$D = \sum m_i[P_i] - (\sum m_i)[\mathcal{O}]$$

donde cada $m_i \geq 0$ y los P_i son una cantidad finita de puntos tal que si $P_i \in sop(D)$ entonces $-P_i \notin sop(D)$, a no ser que P_i sea un punto de Weierstrass, en cuyo caso $m_i \in \{0, 1\}$.

Observación. Si $P = (x_0, y_0)$ es un punto ordinario, entonces $\mathrm{div}(x - x_0) = [P] + [-P] - 2[\mathcal{O}]$. Si $P = (x_0, y_0)$ es un punto de Weierstrass, entonces $\mathrm{div}(x - x_0) = 2[P] - 2[\mathcal{O}]$.

Teorema 3.3.2. Para cada divisor D de grado 0, existe un divisor semirreducido D_1 tal que $D \sim D_1$

Demostración. Si $D = \sum m_p[P]$ consideramos C_0 el conjunto de los puntos de Weierstrass de C; y (C_1, C_2) una partición de los puntos ordinarios de C con la siguiente propiedad:

1. $P \in C_1 \Leftrightarrow -P \in C_2$
2. $P \in C_1 \Rightarrow m_P \geq m_{-P}$

Entoces:

$$D = \sum_{P \in C_1} m_P[P] + \sum_{P \in C_2} m_P[P] + \sum_{P \in C_0} m_P[P] - m[\mathcal{O}]$$

Sea

$$D_1 = D - \sum_{P=(x_0,y_0)\in C_2} m_P \mathrm{div}(x - x_0) - \sum_{P=(x_0,y_0)\in C_0} \left\lfloor \frac{m_P}{2} \right\rfloor \mathrm{div}(x - x_0)$$

Como $D - D_1$ es principal, entonces $D \sim D_1$. Además, por la observación anterior ocurre que:

$$D_1 = \sum_{P\in C_1} (m_P - m_{-P})[P] + \sum_{P\in C_0} (m_P - \left\lfloor \frac{m_P}{2} \right\rfloor)[P] - m_1[\mathcal{O}]$$

para algún $m_1 \in \mathbb{Z}$, y por lo tanto D_1 es semirreducido. □

Definición 3.3.4 (Divisor reducido)**.** Un divisor semirreducido $D = \sum m_i[P_i] - (\sum m_i)[\mathcal{O}]$ se dice reducido si $\sum m_i \leq g$ donde g es el género de C.

Teorema 3.3.3. Para cada divisor D de grado 0, existe un único divisor reducido D_1 tal que $D \sim D_1$.

La suma en el grupo $\mathrm{Pic}^0_K(C)$, no es difícil de implementar, de hecho es muy sencilla, debido a que es la suma en un grupo abeliano libre. Lo que a priori no es trivial es la reducción de los puntos módulo las clases de equivalencia.

Los divisores reducidos permiten una identificación unívoca de los elementos del grupo $\mathrm{Pic}^0_K(C)$.

3.3.1. Representación de Mumford

Los divisores reducidos K-racionales admiten una representación por polinomios con coeficientes en K, conocida como representación de Mumford. Esta representación permite la implementación de la reducción de los divisores módulo las clases de equivalencia de $\mathrm{Pic}^0_K(H)$, aplicando la idea que permite la reducción del teorema 3.3.2, repetidas veces.

Definición 3.3.5 (Representación de Mumford)**.** Dado $D = \sum m_i[P_i] - (\sum m_i)[\mathcal{O}]$, D se representa como $(u(x), v(x))$, donde los polinomios $u(x)$ y $v(x)$ se definen:

Si $P_i = (x_i, y_i)$ entonces $u(x) = \prod_i (x - x_i)^{m_i}$ y $v(x)$ es tal que $v(x_i) = y_i$ y su grado es mínimo.

Observación. Si D es reducido entonces $\deg v < \deg u \leq g$. Además, la condición de la definición $(x_i, v(x_i)) \in C$ es equivalente a pedir $u|(v^2 + vh - f)$.

Para entender la relación entre D y su representación de Mumford, se define el máximo común divisor (que se anota gcd por sus siglas en inglés) entre divisores.

Definición 3.3.6 (gcd de divisores)**.** Sean $D_1 = \sum m_P[P]$ y $D_2 = \sum n_P[P]$ dos divisores. El máximo común divisor de D_1 y D_2 se define como el divisor de grado cero:

$$\gcd(D_1, D_2) = \sum \text{mín}(m_P, n_P)[P] - \Big(\sum \text{mín}(m_P, n_P)\Big)[\mathcal{O}]$$

Observación. Si D es un divisor reducido y $(u(x), v(x))$ su representación de Mumford, entonces $D = \gcd\big(\text{div}(u(x)), \text{div}(v(x) - y)\big)$

El algoritmo de la suma fue descrito por David Cantor en 1987. La idea detrás de este algoritmo es aprovechar la representación de Mumford, y el método de reducción utilizado en la demostración del teorema 3.3.2.

Algoritmo 1 Cantor

Entrada: Dos elementos de $\text{Pic}^0_K(C)$: $D_1 = [(u_1, v_1)]$ y $D_2 = [(u_1, v_1)]$ en la curva $C = y^2 + h(x)y = f(x)$.
Salida: El único divisor reducido D tal que $D = D_1 + D_2$ en $\text{Pic}^0_K(C)$.

1: $d_1 \leftarrow \gcd(u_1, u_2)$ {Algoritmo de Euclides} $[d_1 = e_1u_1 + e_2u_2]$
2: $d \leftarrow \gcd(d_1, v_1 + v_2 + h)$ $[d = c_1d_1 + c_2(v_1 + v_2 + h)]$
3: $s_1 \leftarrow c_1e_1, \quad s_2 \leftarrow c_1e_2, \quad s_3 \leftarrow c_2$ tal que

$$d = s_1u_1 + s_2u_2 + s_3(v_1 + v_2 + h)$$

4: $u \leftarrow \dfrac{u_1u_2}{d^2}$ y $v = \dfrac{s_1u_1v_2 + s_2u_2v_1 + s_3(v_1v_2 + f)}{d} \pmod{u}$
5: **repetir**
6: $\quad u' \leftarrow \dfrac{f - vh - v^2}{u}$ y $v' \leftarrow (-h - u) \pmod{u'}$
7: $\quad u \leftarrow u'$ y $v \leftarrow v'$
8: **hasta** $\deg u \leq g$
9: u mónico
10: **devolver** $[u, v]$

3.3.2. Interpretación geométrica de la ley de grupo en $g = 2$

Consideremos C una curva hiperelíptica de género 2 en un cuerpo de característica distinta de 2. Según el teorema 3.3.1, podemos considerar que la curva está dada por el polinomio $y^2 = f(x)$ con $f \in K[x]$ de grado 5.

Los elementos de Pic^0_K pueden representarse por pares de puntos. Por ejemplo supongamos que $D_1 = [P_1] + [P_2] - 2[\mathcal{O}]$ y $D_2 = [Q_1] + [Q_2] - 2[\mathcal{O}]$, todos distintos, ordinarios y ninguno opuesto de otro.

Estos cuatro puntos determinan una cúbica $y = g(x)$. La intersección de esta cúbica con C está dada por las soluciones de la ecuación $h(x)^2 - f(x) = 0$. Como h^2 tiene grado 6 y f grado 5 hay seis puntos de intersección: $\{P_1, P_2, Q_1, Q_2, R_1, R_2\}$.

Entonces $D_1 + D_2 = [-R_1] + [-R_2]$. Cuando los puntos R_1 y R_2 están definidos sobre K estamos en el caso representado en 3.2. Pero puede ocurrir que los puntos estén definidos en una extensión de K, de todos modos, en ese caso $[-R_1] + [-R_2]$ es un divisor invariante por los automorfismos de G_K, y por lo tanto es un divisor K-racional.

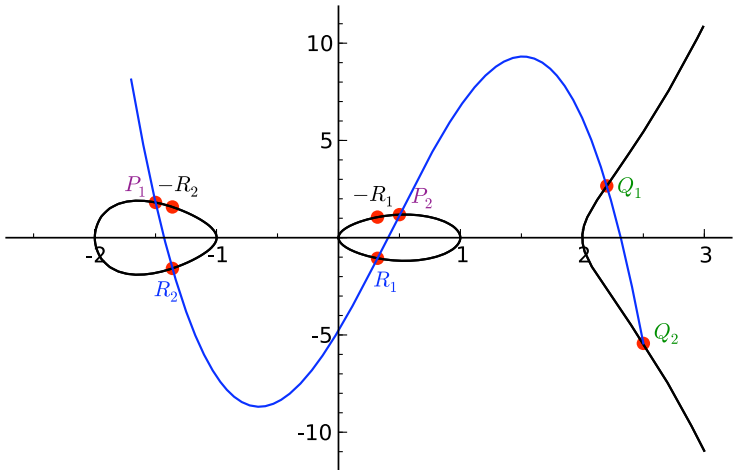


Figura 3.2: Interpretación geométrica de la ley de grupo en $\mathrm{Pic}^0_{\mathbb{R}}(C)$

3.4. Aplicaciones criptográficas

Luego de elegido el grupo y una forma de representar sus elementos, es posible utilizar primitivas criptográficas basadas en el problema del logaritmo discreto en estos grupos.

En general, existen muchos factores de seguridad a tener en cuenta, más allá del problema matemático en el que se basa un criptosistema. Como se muestra en el cuadro , un criptosistema se basa en un objeto matemático, y a partir de este objeto se realizan construcciones como ser primitivas criptográficas, protocolos de comunicación y la implementación concreta en software o hardware. Todas estas capas del modelo son relevantes e introducen dificultades nuevas a resolver. Es por esto que la implementación concreta de los criptosistemas del capítulo 1 difieren un poco de la idea allí presentada. La mayoría de estas diferencias tienen su origen en la resolución de

vulnerabilidades que no radican en el problema matemático de fondo, sino en el protocolo utilizado o su implementación.

La implementación concreta de los criptosistemas teóricos es un problema complejo de ingeniería. La representación de la curva, la codificación de los elementos del grupo y la implementación de las operaciones son problemas muy estudiados y para nada triviales.

El diseño de estos algoritmos deberá contemplar:

- Eficiencia
- Seguridad

La seguridad es un problema que va mucho más allá de la complejidad que la complejidad computacional de los ataques al algoritmo de cifrado. Existen muchos otros factores en los que un criptosistema puede ser vulnerable. Un criptosistema puede ser vulnerable a nivel de la primitiva criptográfica, a nivel del protocolo o a nivel de la implementación.

Por ejemplo, los llamados ataques de canales laterales (side-channel attacks en inglés) no atacan al algoritmo ni al protocolo sino a su implementación en el hardware. Las smart cards[1] son particularmente vulnerables a los ataques de canales laterales. Los ataques de canales laterales explotan el comportamiento del chip mientras computa. La idea detrás de este tipo de ataques es analizar el consumo de energía, emisiones electromagnéticas, tiempo de cálculo o reacciones frente a perturbaciones o errores introducidos artificialmente. De esta forma es posible revelar información sobre las claves presentes en memoria.

Aunque la idea de ataques de canales laterales pueda sonar un poco alejada de la realidad, existen muchísimos ejemplos de ataques exitosos por este medio. Por ejemplo el presentado por IBM en [Rao02] que extrae la clave utilizada por el algoritmo COMP128 en una tarjeta SIM, utilizando un ataque que necesita encriptar tan solo 8 textos elegidos.

Este tema de los ataques de canales laterales es muy vasto e interesante. En el capítulo 28 de [Coh05] se introduce el tema, y en el 29 se presentan contramedidas contra estos ataques.

Otras vulnerabilidades usuales son a nivel del protocolo. Por ejemplo, puede ser que el criptosistema utilice primitivas criptográficas muy seguras, pero si las partes honestas no se autentican de una forma segura, es posible que un adversario malicioso realice una ataque del tipo *man in the middle*. En un

[1] Las smart cards son pequeños chips con circuitos integrados que se usan con distintos fines: desde la tarjeta SIM del celular hasta la tarjeta del ómnibus. En general una smart card es un dispositivo complejo con muchas restricciones de cómputo y almacenamiento.

ataque *man in the middle* el atacante se comunica con las partes honestas, interceptando los mensajes y cambiándolos. Las partes honestas piensan que se comunican directamente en una comunicación privada, cuando en realidad el atacante es el que controla la comunicación. Para poder realizar este ataque, el atacante deberá poder interceptar todos los mensajes e introducir nuevos. La protección contra este tipo de ataques es tarea del protocolo.

Con respecto al tema de este capítulo, la criptografía de curvas elípticas es de hecho muy utilizada, en particular en dispositivos móviles ya que ofrece la misma seguridad que RSA con claves mucho más pequeñas. Por ejemplo según [Pat00] 256 bits de clave en ECC ofrecen la misma seguridad que 2048 bits de clave RSA.

La criptografía de curvas hiperelípticas fue propuesta porque aparentaba proveer la misma seguridad con claves más pequeñas que curvas elípticas, bajo el supuesto de que las curvas de género más grande no eran atacables por algoritmos subexponenciales. Sin embargo, resultó que no era así, para $g > 2$ existen ataques subexponenciales al problema del logaritmo discreto en curvas hiperelípticas [The03]. En caso de género 2, las curvas hiperelípticas se cree ofrecen la misma seguridad que las curvas elípticas pero con claves más pequeñas. De todos modos los cálculos de la ley de grupo en curvas hiperelípticas son más costosos, así como también es más costoso generar la curva inicial.

3.4.1. Generalizaciones

Una de las referencias más importantes en la que se basa esta monografía es la tesis de doctorado de Isabelle Dechene [Dec05]. En su tesis ella sugiere el uso de jacobianas generalizadas con fines criptográficos y plantea un criptosistema concreto bajo este modelo. La idea detrás de las jacobianas generalizadas es utilizar una relación de equivalencia más fina que la equivalencia lineal. De esta forma, el cociente es un grupo más grande que el grupo de Picard, pero que en general no resulta ser una variedad abeliana.

Apéndice A

Teorema de Pascal

El Teorema de Pascal conocido como Teorema del Hexágono Místico, es un teorema de geometría proyectiva que dice que si un hexágono está inscripto en una cónica, las rectas determinadas por lados opuestos se cortan en 3 puntos alineados de $\mathbb{P}^2$. En particular, cuando la recta determinada por estos 3 puntos es la recta impropia (Z=0), estamos en el caso del teorema utilizado en 2.2.1 para probar la asociatividad de la suma en curvas de Pell.

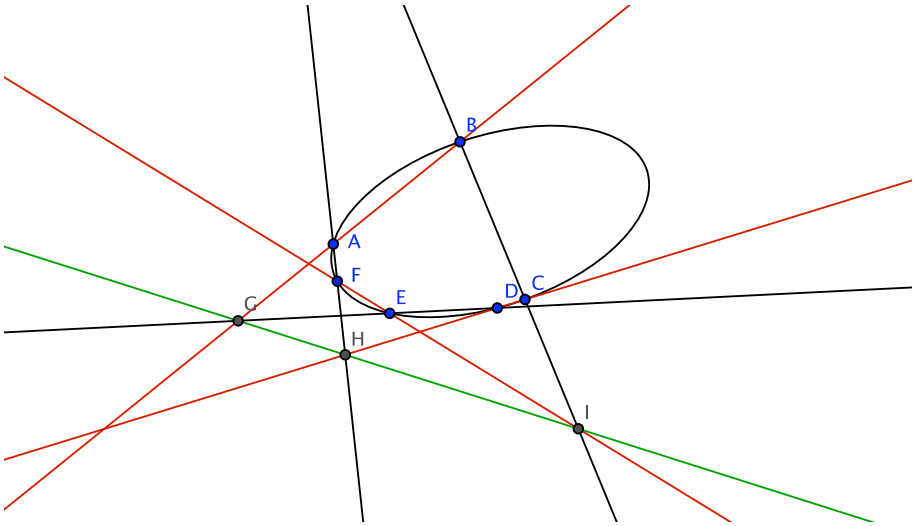


Figura A.1: Teorema de Pascal sobre $\mathbb{R}$

Para demostrar el teorema de Pascal utilizaremos un teorema clásico de la geometría algebraica: el Teorema de Bézout. Por simplicidad enunciamos la versión débil del teorema, ya que es suficiente para demostrar el teorema de Pascal y no requiere que el cuerpo sea algebraicamente cerrado ni trabajar con números de intersección.

Teorema A.0.1 (Bézout)**.** Sean C, D curvas algebraicas proyectivas sobre un cuerpo $\mathbb{K}$, tal que C se define a través del polinomio $P(X, Y, Z)$ homogéneo de grado n y D se define a través del polinomio $Q(X, Y, Z)$ homogéneo de grado m. Si P y Q no tienen factores comunes entonces la cantidad de puntos de intersección de las curvas en $\mathbb{P}^2(\mathbb{K})$ es finito. De hecho $\#C \cap D \leq n\,m$.

Lema. *Sean C y D dos curvas proyectivas definidas por los polinomios $P(X, Y, Z)$ y $Q(X, Y, Z)$ ambos de grados n, sin factores comunes. Supongamos que C y D se cortan en exactamente n^2 puntos, de los cuales $n\,r$ se encuentran en una curva irreducible E definida por el polinomio irreducible $R(X, Y, Z)$ de grado r. Entonces los restantes $n(n-r)$ puntos se encuentran en una curva de grado menor o igual a $n - r$.*

Demostración. Sea $(X_0 : Y_0 : Z_0) \in E$ tal que $(X_0 : Y_0 : Z_0) \notin C \cap D$ (para garantizar que este punto exista podemos suponer que el cuerpo es algebraicamente cerrado, ya que en caso contrario E podría tener sólo los $n\,r$ puntos de $C \cap D$).

Consideremos la curva F definida por el polinomio

$$S(X, Y, Z) = P(X_0, Y_0, Z_0)Q(X, Y, Z) - Q(X_0, Y_0, Z_0)P(X, Y, Z)$$

$S(X, Y, Z)$ es un polinomio homogéneo de grado n, pero la intersección $E \cap F$ tiene $n\,r + 1$ puntos. Por lo tanto, no estamos en las hipótesis del teorema de Bézout, es decir, los polinomios $R(X, Y, Z)$ y $S(X, Y, Z)$ tienen un factor en común. Pero como $R(X, Y, Z)$ es irreducible por hipótesis ocurre que

$$S(X, Y, Z) = R(X, Y, Z)T(X, Y, Z)$$

Entonces los otros $n(n-r)$ puntos que son ceros de $S(X, Y, Z)$ pero no de $R(X, Y, Z)$ se encuentran en la curva definida por $T(X, Y, Z)$ que resulta ser una curva algebraica afín de grado de grado menor o igual que $n - r$. □

Teorema A.0.2 (Hexágono Místico de Pascal)**.** Sean A, B, C, D, E, F seis puntos distintos sobre una cónica en $\mathbb{P}^2$. Si $\overline{AB} \cap \overline{DE} = \{G\}$, $\overline{BC} \cap \overline{EF} = \{H\}$ y $\overline{CD} \cap \overline{FA} = \{I\}$, entonces los puntos G,H,I están alineados.

Demostración. Sean l_{AB}, l_{BC}, l_{CD}, l_{DE}, l_{EF} los polinomios homogéneos de grado 1 que definen a las rectas proyectivas $\overline{AB}, \overline{BC}, \overline{CD}, \overline{DE}, \overline{EF}, \overline{FA}$ respectivamente.

Consideramos las curvas:

- $\mathcal{C}$ definida por el polinomio $l_{AB}l_{CD}l_{EF}$ de grado 3.
- $\mathcal{D}$ definida por el polinomio $l_{BC}l_{DE}l_{FA}$ de grado 3.

Como 6 puntos de la intersección $\mathcal{C} \cap \mathcal{D}$ se encuentran en una curva absolutamente irreducible de grado 2, estamos en las hipótesis del lema anterior, y por lo tanto los $n^2 - r\,n = 9 - 6 = 3$ puntos restantes se encuentran en una curva de grado $n - r = 1$, es decir, alineados. (ver dibujo A.1).

□

Corolario (Teorema 2.2.1)**.** *Sean A, B, C, D, E, F seis puntos sobre una cónica. Si $\overline{AB} \parallel \overline{DE}$ y $\overline{BC} \parallel \overline{EF}$, entonces $\overline{CD} \parallel \overline{FA}$.*

Demostración. Basta observar que si $\overline{AB} \parallel \overline{DE}$ y $\overline{BC} \parallel \overline{EF}$, entonces proyectivamente $\overline{AB} \cap \overline{DE}$, $\overline{CD} \cap \overline{FA} \in \mathcal{R}_i$, y son puntos distintos ($\mathcal{R}_i$ es la recta impropia y está definida por el polinomio Z=0).

Entonces $\overline{CD} \cap \overline{FA}$ también pertenece a la recta impropia, por lo que en el plano afín ocurre que $\overline{CD} \parallel \overline{FA}$.

□

Bibliografía

[Ble97] D. Bleichenbacher, *On the Security of the KMOV Public Key Cryptosystem* Proceedings of the 17th Annual international Cryptology Conference on Advances in Cryptology. Lecture Notes In Computer Science, vol. 1294. Springer-Verlag, London, pp 235-248, 1997.

[Che07] Z. Chen, X. Song, *A public key cryptosystem scheme on conic curves over* $\mathbb{Z}_n$ Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 2007

[Coh05] H. Cohen, G. Frey. *Handbook of Elliptic and Hyperelliptic Curve Cryptography* Chapman and Hall/CRC, Boca Raton, 2005.

[Dec05] I. Dechene, *Generalized Jacobians in Cryptography* Ph.D thesis, McGill University, 2005.

[ElG85] T. ElGamal, *A Public Key Cryptosystem and a Signature scheme Based on Discrete Logarithms* IEEE Transactions on Information Theory 31(4), pp. 469-472, 1985.

[Han04] D. Hankerson, A. Menezes, S. Vanstone, *Guide to Elliptic Curve Cryptography* Springer, 2004.

[Ivo08] C. Ivorra, *Geometría algebraica* disponible en http://www.uv.es/ivorra/Libros/Libros.htm, 2008.

[Ivo10] C. Ivorra, *Curvas elípticas* disponible en http://www.uv.es/ivorra/Libros/Libros.htm, 2010.

[Kat08] J. Katz, Y. Lindell, *Introduction to modern cryptography* Chapman & Hall, 2008.

[Ker83] A. Kerckhoffs, *La cryptographie militaire* Journal des sciences militaires, vol. IX, pp. 5–83, pp. 161–191, Feb. 1883.

[Koy91] K. Koyama, U. Maurer, T. Okamoto, S. Vanstone, *New public key schemes based on elliptic curves over the ring* Z_n Crypto'91 pp 252-266, 1991.

[Len00] A. Lenstra, E. Verheul, *The XTR public key system* Advances in cryptology CRYPTO 2000, Lecture Notes in Computer Science 1880, Springer, pp 1–19, 2000

[Men02] A. Menezes, *Elliptic curve public key cryptosystems* Kluwer Academic Publishers, 2002.

[Mil06] J. Milne *Elliptic curves* BookSurge Publishers, 2006.

[Mil09] J. Milne, *Course notes. Algebraic Geometry* disponilbe en http://www.jmilne.org/math/CourseNotes/ag.html, 1996.

[Pad02] S. Padhye, *A public key cryptosystem based on Pell equation*

[Pat00] V. Patel *Key sizes selection in cryptography security comparison between ECC and RSA.*, http://teal.gmu.edu/courses/ECE543/project/slides_2000/patel.pdf, 2000.

[Rao02] J. Rao, P. Rohatgi, H. Scherzer, S. Tinguely *Partitioning Attacks: Or How to Rapidly Clone Some GSM Cards* IEEE Symposium on Security and Privacy, Oakland, Mayo 2002 http://www.research.ibm.com/intsec/gsm.ps

[Rub03] K. Rubin, A. Silverberg, *Torus-based cryptography* Advances in Cryptography CRYPTO 2003, Lecture Notes in Computer Science, pp 349-365, Springer, 2003.

[Sch03] J. Scholten, F. Vercauteren, *An Introduction to Elliptic and Hyperelliptic Curve Cryptography and the NTRU*, disponible en http://www.math.uni-bonn.de/~saxena/courses/WS2010-ref4.pdf, 2003.

[Sha49] C. Shannon *Communication Theory of Secrecy Systems* Bell System Technical Journal, vol.28(4), pp 656–715, 1949.

[Sha63] C. Shannon, W. Weaver. *The Mathematical Theory of Communication* Urbana: University of Illinois Press, 1963

[Sin01] S. Singh *The code book: the science of secrecy from ancient Egypt to quantum cryptography* Delacorte Press, 2001.

[Smi93] P. Smith, M. Lennon, *LUC: A New Public Key System* Proceedings of the IFIP TC11 Ninth International Conference on Information Security, pp 103-117, 1993.

[Sti06] D. Stinson *Cryptography theory and Practice* Tercera edición, Chapman & Hall, 2006.

[The03] N. Theriault *Index calculus attack for hyperelliptic curves of small genus* ASIACRYPT 2003 Lecture Notes in Computer Science 2894, pp 75–92, 2003.