Sección de Matemáticas
Universidad de La Laguna

Manuel Betancort Pérez

Estructura de retículos y su aplicación en la criptografía

Lattices structure and their aplication to cryptography

Trabajo Fin de Grado
Grado en Matemáticas
La Laguna, mayo de 2023

Dirigido por
Irene Márquez Corbella
Luis José Santana Sánchez

Irene Márquez Corbella
Departamento de Matemáticas, Estadística e Investigación Operativa
Universidad de La Laguna
38200 La Laguna, Tenerife

Luis José Santana Sánchez
Departamento de Álgebra, Ánalisis Matemático, Geometría y Topología
Universidad de Valladolid
47011 Valladolid

Agradecimientos

Quiero agradecer a mi familia por acompañarme en este largo camino, a María por aguantar mis agobios constantes y motivarme a seguir, y sobre todo a mis tutores, pues su implicación y ánimos a la hora de afrontar este trabajo han sido esenciales.

Manuel Betancort Pérez
La Laguna, 22 de mayo de 2023

Resumen · Abstract

Resumen

En esta memoria, hacemos una introducción a la teoría de retículos. Se presentan resultados importantes sobre los mismos, y se estudian problemas que son computacionalmente complejos de resolver. Se destacan los conocidos por sus siglas SVP y CVP, siendo estos fundamentales en la criptografía de retículos. Finalmente se propone un sistema sencillo de criptografía basado en los problemas mencionados.

Palabras clave: *Retículos – Criptografía post-cuántica – CVP – SVP – Algoritmo LLL.*

Abstract

In this memory, we make an introduction to lattice theory. We present important results about them and study problems which are computationally hard to solve. We highlight the so called SVP and CVP, which are fundamental in lattice cryptography. Finally, we propose a simple cryptography system based in the aforementioned problems.

Keywords: *Lattice – Post-quantum cryptography – CVP – SVP – LLL Algorithm .*

Contenido

Introducción

Un retículo está compuesto por puntos ordenados en el espacio y separados entre sí. A pesar de su aparente simplicidad, esta estructura geométrica presenta propiedades muy interesantes que han atraído la atención de grandes matemáticos a lo largo de los dos últimos siglos. De hecho, ya hacia finales del siglo XVIII y principios del XIX, matemáticos como Lagrange y Gauss utilizaron los retículos para demostrar resultados de la teoría de números como la ley de reprocidad cuadrática o el teorema de los cuatro cuadrados. Pocas décadas después, es el trabajo de Minkowski el que asienta las bases sobre retículos en su trabajo titulado la *Geometría de los Números* (ver [7]). No es hasta un siglo más tarde, cuando la teoría de retículos cobra gran importancia en la creación de sistemas de *criptografía*.

Entendemos por criptografía el arte de escribir un mensaje de manera secreta, de forma que sólo el receptor deseado sea capaz de entenderlo. En contraparte, tenemos el criptoanálisis, que busca romper el secreto y obtener el mensaje original que se transmite. La confidencialidad de nuestras comunicaciones es un problema que surge desde la Antigüedad, por lo que nos tenemos que remontar al año 400 a.C. para encontrar el primer uso de la criptografía, la *escítala* utilizada por los espartanos durante la guerra de Atenas y Esparta. Otro ejemplo clásico de sistema criptográfico corresponde al cifrado de César que fue utilizado por Julio César en el siglo I a.C. Al igual que sucede con muchas otras ramas de la ciencia, la criptografía tuvo grandes avances durante las dos guerras mundiales, debido a la necesidad de transmitir información de forma secreta entre las distintas unidades militares de un mismo ejército. De hecho, en la Segunda Guerra Mundial, fue clave la intervención de Alan Turing, padre de la infórmatica y la inteligencia artificial, quien diseñó una máquina capaz de automatizar el proceso de criptoanálisis de la máquina Enigma, la cual utilizaba el bando alemán para cifrar y descifrar sus mensajes secretos. A la máquina creada por Turing se le considera el predecesor de los ordenadores actuales. Cabe destacar que todos los esquemas criptográficos mencionados se denominan sistemas de clave privada o

simétricos y requieren que, tanto el emisor, como el receptor dispongan de una clave en común para establecer la comunicación.

No fue hasta el siglo pasado, con la aparición de internet y el número de personas que necesitan proteger su información, que no fue necesario otro tipo de criptografía que no implicara un intercambio inicial de claves. La solución a este problema de la criptografía de clave secreta la dieron Diffie y Hellman en 1976 en su artículo [2]. Surge así la criptografía de clave pública, en la que cada participante genera dos claves, una que se hace pública y otra que se mantiene privada. Así, toda persona puede mandarnos un mensaje haciendo uso de la clave pública, y sólo nosotros podemos descifrarlo al disponer de la clave privada. Este tipo de criptografía se basa en la complejidad de resolver ciertos problemas matemáticos sin conocer la clave privada mientras que, si contamos con ella se resuelven de forma sencilla. Actualmente, la seguridad de la criptografía de clave pública que se utiliza en medios oficiales, se basa únicamente en dos problemas, siendo el primero de ellos el problema de factorización números enteros y el segundo el problema del logaritmo discreto. Un ejemplo de criptosistema de clave pública basado en la factorización de enteros es el famoso criptosistema RSA creado en 1978 por Rivest, Shamir y Adlman (ver [1]).

Estos problemas son problemas complejos con las herramientas actuales (ordenadores clásicos). Sin embargo, en 1994, Peter Shor introduce un algoritmo cuántico que permite resolver los dos problemas que hemos mencionado con un ordenador cuántico de forma rápida (ver [8]). Actualmente no es posible crear un ordenador cuántico que utilice todo su potencial y, por tanto, ponga en jaque a la comunidad. No obstante, en 2019 IBM presentó un primer ordenador cuántico comercial (IBM Q System One), el cual ya dispone de un procesador cuántico cuya potencia no puede ser simulada por ordenadores convencionales. Previo a su lanzamiento, el NIST (Instituto Nacional de Estándares y Tecnología) lanza en 2017 el programa 'Post-Quantum Cryptography Standardization', una competición con el objetivo de encontrar criptosistemas seguros frente a los ordenadores cuánticos. Se propusieron en noviembre de 2017, 69 sistemas, de los cuales 22 están basados en retículos. Tras un total de 4 rondas, en julio de 2022 se presentaron los 4 criptosistemas finalistas, todos ellos basados en retículos.

La finalidad de esta memoria no es indagar en estos criptosistemas tan complejos, pero si que el lector comprenda el funcionamiento de la criptografía basada en retículos y presentar un sistema de criptografía sencillo basado en estas interesantes estructuras. Siguiendo nuestro objetivo, se establecen los distintos capítulos de este trabajo, los cuales están distribuidos de la siguiente manera.

En el primer capítulo se formaliza la definición de retículo y se introducen conceptos básicos sobre el mismo. En particular, estudiamos como obtener bases de un retículo dado, también un invariante esencial de los mismos que es el *determinante* y vemos como los retículos establecen posibles particiones del espacio vectorial que generan mediante las *regiones fundamentales*.

En el segundo capítulo se presentan una serie de parámetros asociados a esta estructura que nos van a ayudar a entender mejor la geometría del retículo. Además, estos parámetros nos permiten introducir los dos problemas asociados a retículos en los que se centra esta memoria, el Problema del vector más corto o SVP y el Problema del vector más cercano o CVP. Estos son problemas para los que a día de hoy no se conocen algoritmos eficaces que lo resuelvan, lo que hace posible su aplicación en criptografía. De hecho, se han demostrado que son NP-completos.

Por último, en el tercer capítulo veremos una breve introducción a la criptografía y como construir un sistema criptográfico seguro basado en la complejidad del CVP. A su vez, abordaremos el Algoritmo LLL como elemento fundamental del criptoanálisis basado en retículos.

1

Introducción a la teoría de retículos

Un retículo es una disposición discreta y regular de puntos en un espacio real Euclídeo n-dimensional. Que sea discreto quiere decir que cada par de puntos se encuentran al menos distanciado por un valor $\epsilon > 0$ fijo, y por regular entendemos que tiene una estructura de grupo con la suma usual de $\mathbb{R}^n$. El ejemplo más sencillo es $\mathbb{Z}^n$. De forma más precisa, un retículo se define como sigue.

Definición 1.1. *Decimos que $\mathcal{L}$ es un retículo en $\mathbb{R}^n$ si existe una matriz $\mathbf{B} \in M_{n\times d}(\mathbb{R})$ de rango$(\mathbf{B}) = d \leq n$ tal que*

$$\mathcal{L} = \mathcal{L}(\mathbf{B}) := \mathbf{B}\mathbb{Z}^d = \{\mathbf{Bx} : \mathbf{x} \in \mathbb{Z}^d\} \subseteq \mathbb{R}^n,$$

donde el vector $\mathbf{x}$ se toma como vector columna.

Por ejemplo, en la siguiente figura los puntos negros representan el retículo $\mathcal{L}$ definido por la matriz $\mathbf{B} = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -0.5 \end{smallmatrix}\right)$.

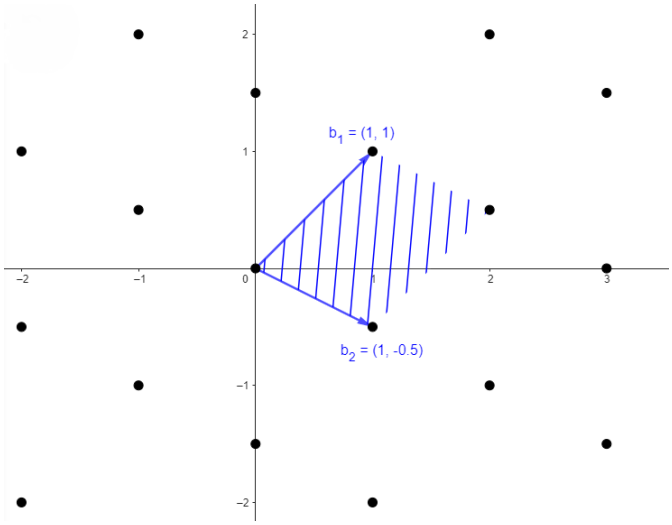


Figura 1.1. Retículo 2-dimensional $\mathbf{B}\mathbb{Z}^2$ generado por $\mathbf{B} = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -0.5 \end{smallmatrix}\right)$.

Escribiremos $\mathcal{L}$ en vez de $\mathcal{L}(\mathbf{B})$ cuando sea claro, o no queramos remarcar, cual es la matriz $\mathbf{B}$ que define el retículo. Notemos que si escribimos $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_d}]$ como la matriz dada por los vectores columna $\mathbf{b_1}, \ldots, \mathbf{b_d} \in \mathbb{R}^n$, se tiene que $\mathcal{L}(\mathbf{B})$ es el conjunto cuyos elementos son de la forma $\sum_{i=1}^{d} \mathbf{b_i} x_i$, con $x_i \in \mathbb{Z}$. Así pues, es fácil ver que todo retículo $\mathcal{L}$ es un grupo con la suma usual de $\mathbb{R}^n$, y es isomorfo a $\mathbb{Z}^d$ bajo el isomorfismo definido por $\mathbf{b_i} \mapsto \mathbf{e_i}$ donde $\mathbf{e_1}, \ldots, \mathbf{e_d}$ conforman la base canónica de $\mathbb{Z}^d$.

En este capítulo introducimos conceptos básicos de retículos y estudiamos su geometría. Por simplicidad y por ser el objeto de interés de esta memoria, nos centraremos en aquellos retículos que se dicen de *dimensión completa*, esto es, cuando $d = n$ y $\mathbf{B}$ es una matriz cuadrada de rango máximo, es decir, $\mathbf{B} \in M_{n\times n}(\mathbb{R})^*$. Sin embargo, muchos de los resultados y definiciones se pueden extender a cualquier rango $d \leq n$.

1.1. La base de un retículo

Sea $\mathcal{L} \subseteq \mathbb{R}^n$ un retículo de dimensión completa y sea $\{\mathbf{b_1}, \ldots, \mathbf{b_n}\}$ un conjunto de vectores linealmente independientes en $\mathbb{R}^n$. Decimos que $\{\mathbf{b_1}, \ldots, \mathbf{b_n}\}$ o $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$ es una base de $\mathcal{L}$ si $\mathcal{L} = \mathcal{L}(\mathbf{B})$.

Cabe destacar que, a diferencia de los espacios vectoriales, no todo conjunto linealmente independiente de vectores en $\mathcal{L}$ forma una base del retículo.

Ejemplo 1.2. Consideremos el retículo $\mathcal{L} = \mathbb{Z}^2$ y los vectores linealmente independientes $\mathbf{v_1}$ = (2,0), $\mathbf{v_2}$ =(0,2) $\in \mathcal{L}$. Es fácil comprobar que $\mathcal{L}([\mathbf{v_1}, \mathbf{v_2}]) = \{(2a, 2b) \mid a, b \in \mathbb{Z}^2\}$ es un subretículo de $\mathbb{Z}^2$ formado por aquellos vectores de coordenadas pares y, por tanto, $\{\mathbf{v_1}, \mathbf{v_2}\}$ no es base de $\mathbb{Z}^2$.

Sin embargo, al igual que en los espacios vectoriales, la base no tiene por qué ser única. Continuando con el mismo ejemplo, se puede observar que

$$\mathbb{Z}^2 = \mathcal{L}(\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)) = \mathcal{L}(\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)).$$

Por eso, dedicamos el resto esta sección a caracterizar cómo son las bases de un retículo.

Lema 1.3. *Sean* $\mathbf{B}$, $\mathbf{C} \in M_{n\times n}(\mathbb{R})^*$, *si existe* $\mathbf{U} \in M_{n\times n}(\mathbb{Z})$ *una matriz entera tal que* $\mathbf{B} = \mathbf{CU}$, *entonces* $\mathcal{L}(\mathbf{B}) \subseteq \mathcal{L}(\mathbf{C})$.

Demostración. La demostración se sigue del hecho de que para todo $\mathbf{x} \in \mathbb{Z}^n$, se tiene que $\mathbf{Bx} = \mathbf{CUx} = \mathbf{Cy}$, siendo $\mathbf{y} = \mathbf{Ux} \in \mathbb{Z}^n$, pues $\mathbf{U} \in M_{n\times n}(\mathbb{Z})$. Luego, $\mathbf{Bx} = \mathbf{Cy} \in \mathcal{L}(\mathbf{C})$ y se concluye que todo elemento de $\mathcal{L}(\mathbf{B})$ se encuentra también en $\mathcal{L}(\mathbf{C})$.

□

Corolario 1.4. *Sean* $\mathbf{B}$, $\mathbf{C} \in M_{n\times n}(\mathbb{R})^*$ *y* $\mathbf{U}$ *una matriz entera invertible en* $M_{n\times n}(\mathbb{Z})$ *tal que* $\mathbf{B} = \mathbf{CU}$, *entonces* $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$.

Demostración. Supongamos que $\mathbf{B} = \mathbf{CU}$ para cierta matriz invertible $\mathbf{U} \in M_{n\times n}(\mathbb{Z})$. Esto implica que existe una matriz entera $\mathbf{U}^{-1}$ tal que $\mathbf{C} = \mathbf{BU}^{-1}$, por lo que aplicando el lema anterior:

$$\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{CU}) \subseteq \mathcal{L}(\mathbf{C}) = \mathcal{L}(\mathbf{BU}^{-1}) \subseteq \mathcal{L}(\mathbf{B}).$$

Se concluye que $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$, es decir, ambas bases generan el mismo retículo.

□

Nótese que las matrices invertibles de $M_{n\times n}(\mathbb{Z})$ no se caracterizan únicamente por tener determinate no nulo, sino que también han de ser unimodulares, es decir, $|det(\mathbf{U})| = 1$. En efecto, recordemos que $det(\mathbf{AB}) = det(\mathbf{A})det(\mathbf{B})$ para toda matriz cuadrada $\mathbf{A}$, $\mathbf{B}$. Esto implica que $det(\mathbf{U})det(\mathbf{U^{-1}}) = det(\mathbf{UU^{-1}}) = det(\mathbf{I}) = 1$. Dado que $\mathbf{U}, \mathbf{U^{-1}} \in M_{n\times n}(\mathbb{Z})$ se tiene que $det(\mathbf{U}), det(\mathbf{U^{-1}}) \in \mathbb{Z}$. Sabemos que el producto de ambos determinantes es igual a 1, por lo que se concluye que $det(\mathbf{U}) = det(\mathbf{U^{-1}}) = 1$ o $det(\mathbf{U}) = det(\mathbf{U^{-1}}) = -1$.

Finalmente demostramos que el recíproco del colorario anterior es también cierto. Es decir, que dos bases cualquieras de un mismo retículo van a estar relacionadas mediante una matriz invertible $\mathbf{U} \in M_{n\times n}(\mathbb{Z})$.

Teorema 1.5. *Sean* $\mathbf{B}$, $\mathbf{C} \in M_{n\times n}(\mathbb{R})^*$, *entonces* $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$ *si, y sólo si, existe una matriz invertible* $\mathbf{U} \in M_{n\times n}(\mathbb{Z})$ *tal que* $\mathbf{B} = \mathbf{CU}$.

Demostración.

⇐ Basta con aplicar el Corolario 1.4.

⇒ Asumimos que $\mathbf{B}$, $\mathbf{C} \in M_{n\times n}(\mathbb{R})^*$ son tales que $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$. Esto implica que todo elemento de $\mathcal{L}(\mathbf{B})$ está en $\mathcal{L}(\mathbf{C})$. En particular, si tomamos el vector columna $\mathbf{e_i} \in \mathbb{Z}^n$ correspondiente al i-ésimo vector de la base canónica de $\mathbb{Z}^n$, tenemos que $\mathbf{Be_i} \in \mathcal{L}(\mathbf{C})$. Esto es, si $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$, entonces existe $\mathbf{u_i} \in \mathbb{Z}^n$ tal que $\mathbf{Be_i} = \mathbf{b_i} = \mathbf{Cu_i}$, para todo $i \in \{1, \ldots, n\}$. Por tanto, hemos encontrado una matriz entera $\mathbf{U} = [\mathbf{u_1}, \ldots, \mathbf{u_n}] \in M_{n\times n}(\mathbb{Z})$ tal que $\mathbf{B} = \mathbf{CU}$. Falta concluir que $\mathbf{U}$ es invertible en $M_{n\times n}(\mathbb{Z})$. Para ello repetimos el mismo argumento anterior intercambiando las matrices $\mathbf{B}$ y $\mathbf{C}$ y así, obtenemos una matriz $\mathbf{V} \in M_{n\times n}(\mathbb{Z})$ tal que $\mathbf{C} = \mathbf{BV}$. Veamos que $\mathbf{V}$ es la matriz inversa de $\mathbf{U}$. En efecto, tenemos que

$$\mathbf{B} = \mathbf{CU} = \mathbf{BVU} \Rightarrow \mathbf{B}(\mathbf{I} - \mathbf{VU}) = 0.$$

Al ser $\mathbf{B}$ una unidad en $M_{n\times n}(\mathbb{R})$, no puede ser un divisor de cero y, por tanto, $\mathbf{I} - \mathbf{VU} = 0$, es decir, $\mathbf{I} = \mathbf{VU}$. De forma análoga se comprueba que $\mathbf{I} = \mathbf{UV}$, demostrando así que $\mathbf{U} \in M_{n\times n}(\mathbb{Z})$ es una matriz invertible, con inversa $\mathbf{U}^{-1} = \mathbf{V}$.

□

1.2. Ortogonalización de Gram-Schmidt

Uno de los procedimientos más recurrentes en álgebra lineal es la ortogonalización de una base $\mathbf{B}$ dada de un espacio vectorial. Esto es un procedimiento en el que, a partir de la base dada, se busca una base $\mathbf{B}^*$ del mismo espacio vectorial, pero cuyos vectores sean ortogonales dos a dos. En el caso de retículos, veremos en el próximo capítulo que tener bases ortogonales del mismo facilita enormemente la resolución de problemas geométricos subyacentes. Sin embargo, como indicamos en esta sección, no todo retículo posee una base ortogonal que lo genere. Aún así, trabajar con ortogonalizaciones de sus bases facilita muchos aspectos computacionales. Veamos cómo obtener dicha base a través del proceso conocido como ortogonalización de Gram-Schmidt. Para ello es necesario introducir el concepto de *componente ortogonal*.

Definición 1.6. *Sea $\mathbf{b} \in \mathbb{R}^n$ un vector cualquiera y S un subconjunto de $\mathbb{R}^n$, llamamos componente de $\mathbf{b}$ ortogonal a S al vector $\mathbf{b} \perp S$ que viene definido por las siguientes condiciones:*

- *$(\mathbf{b} \perp S)$ es un punto de $\mathbf{b} + span(S)$.*
- *$(\mathbf{b} \perp S)$ es un vector ortogonal a todo elemento de S.*

Geométricamente, $(\mathbf{b} \perp S)$ puede verse como el vector más corto perteneciente a $\mathbf{b} + span(S)$, donde $span(S)$ denota al espacio vectorial real generado por S.

Ejemplo 1.7. Sea $\mathbf{b} = (1, 0) \in \mathbb{R}^2$ y $S = \{(1, 2)\} \subseteq \mathbb{R}^2$. Veamos cuál es la componente de $\mathbf{b}$ ortogonal a S. Se tiene que $span(S) = span(\{(1, 2)\})$ corresponde al espacio vectorial generado por el punto $(1, 2)$. Geométricamente, esto no es más que la recta que pasa por el $(0, 0)$ y tiene a $(1, 2)$ como vector director. Por su parte, tenemos que $\mathbf{b} + span(\{(1, 2)\})$ es la recta trasladada en la dirección de $\mathbf{b} = (1, 0)$. Esto es, la recta que pasa por $(1, 0)$ y tiene a $(1, 2)$ como vector director. Por tanto,

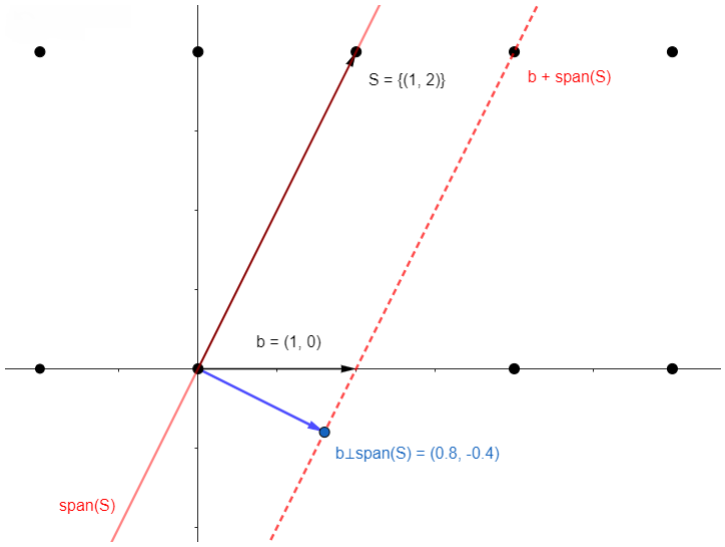
$$\mathbf{b} + span(\{(1, 2)\}) = \{(x, y) \in \mathbb{R}^n \mid y = 2(x - 1)\}.$$

Por definición, la componente de $\mathbf{b}$ ortogonal a S es un vector de este conjunto que es ortogonal a $(1, 2)$. Por tanto, esta componente es el punto (x, y) que satisface el sistema de ecuaciones:

$$\left.\begin{array}{c} y = 2(x - 1) \\ \langle (x, y), (1, 2) \rangle = x + 2y = 0. \end{array}\right\}$$

Resolviendo el sistema obtenemos que $\mathbf{b} \perp span(S) = (4/5, -2/5)$.

En la siguiente figura podemos ver que, geométricamente, $\mathbf{b} \perp span(S)$ es el vector más pequeño de $\mathbf{b} + span(\{(1, 2)\})$ con la distancia Euclídea usual.



De esta forma, se define la ortogonalización de Gram-Schmidt como sigue.

Definición 1.8. *Sea* $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}] \in M_{n\times n}(\mathbb{R})^*$. *Denominamos ortogonalización de Gram-Schmidt de* $\mathbf{B}$ *a la matriz* $\mathbf{B}^* = [\mathbf{b_1^*}, \ldots, \mathbf{b_n^*}]$ *que se obtiene tomando* $\mathbf{b_1^*} = \mathbf{b_1}$ *y* $\mathbf{b_i^*} = \mathbf{b_i} \perp \{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\}$ *para todo* $i \in \{2, \ldots, n\}$.

Observamos que, por construcción, tanto $\mathbf{B}$ como $\mathbf{B}^*$ generan el mismo espacio vectorial, que en este caso es $\mathbb{R}^n$. Sin embargo, $\mathbf{B}$ y $\mathbf{B}^*$, en general, no generan el mismo retículo.

Ejemplo 1.9. Sea $\mathbf{B} = \left(\begin{smallmatrix} 1 & 1 \\ 2 & 0 \end{smallmatrix}\right)$, por defincón se tiene que $\mathbf{B}^* = [\mathbf{b_1^*}, \mathbf{b_2^*}]$ con $\mathbf{b_1^*} = (1, 2)$ y $\mathbf{b_2^*} = (1, 0) \perp \{(1, 2)\} = (4/5, -2/5)$, como ya vimos en el Ejemplo 1.7. Notemos que, al ser $\mathbf{B}$ una matriz entera, se tiene que $\mathcal{L}(\mathbf{B}) \subseteq \mathbb{Z}^2$ es un subretículo de $\mathbb{Z}^2$. Sin embargo, $\mathcal{L}(\mathbf{B}^*) \nsubseteq \mathbb{Z}^2$ pues $\mathbf{b_2^*} = (4/5, -2/5) \in \mathcal{L}(\mathbf{B}^*)$. Esto muestra que $\mathbf{B}$ y $\mathbf{B}^*$ no definen el mismo retículo.

Siguiendo con este ejemplo, merece la pena destacar que el orden elegido en la base $\mathbf{B}$ afecta al proceso de ortogonalización. En efecto, supongamos ahora que tomamos $\mathbf{C} = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right) = [\mathbf{b_2}, \mathbf{b_1}]$. Claramente $\mathcal{L}(\mathbf{C}) = \mathcal{L}(\mathbf{B})$ pues no es más que una reordenación de los elementos de la base. Sin embargo, siguiendo el proceso de ortogonalización similar al que se muestra en la Figura 1.7, no es difícil ver que $\mathbf{C}^* = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)$ y, en este caso, $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C}) = \mathcal{L}(\mathbf{C}^*)$. Luego, no solamente $\mathbf{C}^*$ y $\mathbf{B}^*$ son matrices distintas, sino que generan retículos diferentes también.

A pesar de que en este ejemplo, hemos podido encontrar una base ortogonal $\mathbf{C}^*$ del retículo inicial $\mathcal{L}(\mathbf{B})$, en general, esto no es siempre posible. Por ejemplo, si consideramos el retículo dado en la Figura 1.1, éste no tiene ninguna base ortogonal. Esto es fácil demostrarlo una vez hayamos visto herramientas que se

introducen en el Capítulo 2. Luego, retomaremos este ejemplo al final de ese capítulo.

El método de ortogonalización de Gram-Schmidt establece explícitamente como calcular la matriz $\mathbf{B}^*$ a partir de una base dada $\mathbf{B}$, tal y como se demuestra en la siguiente proposición.

Proposición 1.10. *Sea* $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}] \in M_{n\times n}(\mathbb{R})^*$, *su ortogonalización de Gram-Schmidt es* $\mathbf{B}^* = [\mathbf{b_1^*}, \ldots, \mathbf{b_n^*}]$ *con*

$$\mathbf{b_i^*} = \mathbf{b_i} - \sum_{j<i} \mu_{i,j}\mathbf{b_j^*}, \text{ donde } \mu_{i,j} = \frac{\langle \mathbf{b_i}, \mathbf{b_j^*}\rangle}{\langle \mathbf{b_j^*}, \mathbf{b_j^*}\rangle}.$$

Demostración. Sea $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}] \in M_{n\times n}(\mathbb{R})^*$ y sea $\mathbf{B}^*$ la matriz dada en el enunciado. Para comprobar que $\mathbf{B}^*$ es la ortogonalización de Gram-Schmidt de $\mathbf{B}$ basta con ver que $\mathbf{b_1}^* = \mathbf{b_1}$ y $\mathbf{b_i}^* = \mathbf{b_i} \perp [\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}]$ para todo $i \in \{2, \ldots, n\}$, es decir que

(i) $\mathbf{b_i}^*$ es un punto de $\mathbf{b_i} + span(\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\})$, y
(ii) $\mathbf{b_i}^*$ es un vector ortogonal a todos los elementos de $\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\}$ donde

$$\mathbf{b_i}^* = \mathbf{b_i} - \sum_{\mathbf{j<i}} \mu_{\mathbf{i,j}}\mathbf{b_j^*}.$$

Observamos que $\mathbf{b_i}^* \in \mathbf{b_i} + span(\{\mathbf{b_1}^*, \ldots, \mathbf{b_{i-1}}^*\})$. Sin embargo, por definición de componente ortogonal, $span(\{\mathbf{b_1}^*, \ldots, \mathbf{b_{i-1}}^*\}) = span(\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\})$, luego $\mathbf{b_i}^* \in \mathbf{b_i} + span(\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\})$ y se tiene (i).

Para demostrar (ii) veamos que $\mathbf{b_i}^*$ es ortogonal al conjunto $\{\mathbf{b_1^*}, \ldots, \mathbf{b_{i-1}}^*\}$ y, por tanto, que es ortogonal al espacio vectorial que genera. Así, en particular, se tiene que $\mathbf{b_i}^*$ es ortogonal a $\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\} \subseteq span(\{\mathbf{b_1^*}, \ldots, \mathbf{b_{i-1}}^*\})$. Para ello, procederemos por inducción sobre i.

Sea $i = 2$, vemos que efectivamente $\mathbf{b_2^*}$ es ortogonal a $\mathbf{b_1}^*$, pues

$$\langle \mathbf{b_2}^*, \mathbf{b_1^*}\rangle = \langle \mathbf{b_2} - \mu_{2,1}\mathbf{b_1^*}, \mathbf{b_1^*}\rangle = \langle \mathbf{b_2}, \mathbf{b_1^*}\rangle - \mu_{2,1}\langle \mathbf{b_1^*}, \mathbf{b_1^*}\rangle =$$

$$\langle \mathbf{b_2}, \mathbf{b_1^*}\rangle - \frac{\langle \mathbf{b_2}, \mathbf{b_1^*}\rangle}{\langle \mathbf{b_1^*}, \mathbf{b_1^*}\rangle}\langle \mathbf{b_1^*}, \mathbf{b_1^*}\rangle = \langle \mathbf{b_2}, \mathbf{b_1^*}\rangle - \langle \mathbf{b_2}, \mathbf{b_1^*}\rangle = 0.$$

Supongamos ahora cierto hasta i, es decir, el conjunto $\{\mathbf{b_1^*}, \ldots, \mathbf{b_i}^*\}$ es ortogonal dos a dos. Veamos que esto se cumple también para $i + 1$. Para ello, basta con demostrar que $\langle \mathbf{b_{i+1}}^*, \mathbf{b_j^*}\rangle = 0$ para todo $j = 1, \ldots, i$. En efecto,

$$\langle \mathbf{b_{i+1}}^*, \mathbf{b_j^*}\rangle = \langle \mathbf{b_{i+1}} - \mu_{i+1,1}\mathbf{b_1^*} - \mu_{i+1,2}\mathbf{b_2^*} - \cdots - \mu_{i+1,i}\mathbf{b_i^*}, \mathbf{b_j^*}\rangle =$$

$$= \langle \mathbf{b_{i+1}}, \mathbf{b_j^*}\rangle - \mu_{i+1,1}\langle \mathbf{b_1^*}, \mathbf{b_j^*}\rangle - \mu_{i+1,2}\langle \mathbf{b_2^*}, \mathbf{b_j^*}\rangle - \cdots - \mu_{i+1,i}\langle \mathbf{b_i^*}, \mathbf{b_j^*}\rangle.$$

Aplicando la hipótesis de inducción, se tiene que $\langle \mathbf{b_k^*}, \mathbf{b_j^*} \rangle = 0$ para todo $k \in \{2, \dots, i\} \setminus \{j\}$ y, por tanto,

$$\langle \mathbf{b_{i+1}}^*, \mathbf{b_j^*} \rangle = \langle \mathbf{b_{i+1}}, \mathbf{b_j^*} \rangle - \mu_{i+1,j} \langle \mathbf{b_j^*}, \mathbf{b_j^*} \rangle =$$

$$= \langle \mathbf{b_{i+1}}, \mathbf{b_j^*} \rangle - \frac{\langle \mathbf{b_{i+1}}, \mathbf{b_j^*} \rangle}{\langle \mathbf{b_j^*}, \mathbf{b_j^*} \rangle} \langle \mathbf{b_j^*}, \mathbf{b_j^*} \rangle = 0.$$

Se concluye que $\mathbf{B}^*$ es la matriz que resulta de aplicar el método de ortogonalización de Gram-Schmidt.

□

Observación 1.11. La proposición anterior nos dice que la ortogonalización de Gram-Schmidt $\mathbf{B}^* = [\mathbf{b_1^*}, \dots, \mathbf{b_n^*}]$ satisface la relación $\mathbf{B} = \mathbf{B}^*\mathbf{T}$, siendo $\mathbf{T}$ la matriz triangular

$$\mathbf{T} = \begin{bmatrix} 1 & \mu_{2,1} & \dots & \mu_{n,1} \\ & \ddots & & \vdots \\ & & 1 & \mu_{n,n-1} \\ & & & 1 \end{bmatrix}.$$

1.3. El determinante

Hemos visto previamente que para cualquier retículo $\mathcal{L}$, podemos encontrar una base $\mathbf{B}$ que lo genere, por lo que es natural plantearnos cuál es el determinante del mismo. Lo visto en las dos secciones anteriores nos permite introducir la siguiente definición.

Definición 1.12. *Sea $\mathcal{L}$ un retículo y $\mathbf{B} \in M_{n\times n}(\mathbb{R})^*$ una base de $\mathcal{L}$ con $\mathbf{B}^*$ su correspondiente ortagonalización de Gram-Schmidt. Definimos el* determinante *de $\mathcal{L}$ como*

$$det(\mathcal{L}) := |det(\mathbf{B})| = |det(\mathbf{B}^*)|.$$

Notamos que está bien definido dado que es independiente de la base escogida. Esto se sigue del hecho de que dadas $\mathbf{B}$ y $\mathbf{C} \in M_{n\times n}(\mathbb{R}^*)$ bases del retículo $\mathcal{L}$, el Teorema 1.5 nos dice que existe una matriz invertible $\mathbf{U} \in M_{n\times n}(\mathbb{Z})$ que las relaciona, resultando así que

$$|det(\mathbf{B})| = |det(\mathbf{CU})| = |det(\mathbf{C})||det(\mathbf{U})| = |det(\mathbf{C})|,$$

pues $\mathbf{U}$ es unimodular. Además, aplicando la Observación 1.11 se comprueba que

$$|det(\mathbf{B})| = |det(\mathbf{B}^*\mathbf{T})| = |det(\mathbf{B}^*)||det(\mathbf{T})| = |det(\mathbf{B}^*)|.$$

Recordemos que, geométricamente, el determinante de una matriz $\mathbf{B}$ coincide con el volumen del paralelepípedo construido a partir de los vectores que constituyen dicha matriz (ver [9]). En el caso de los retículos, esto nos dice que el determinante del retículo generado por una base $\mathbf{B}$, va a coincidir con el volumen del paralelepípedo conocido como *paralelepípedo fundamental* $\mathcal{P}(\mathbf{B})$, el cual se define como

$$\mathcal{P}(\mathbf{B}) := \mathbf{B}\mathbb{T}^n = \left\{ \sum_{i=1}^{n} \mathbf{b_i} x_i \mid 0 \leq x_i < 1, \forall i \right\} \tag{1.1}$$

siendo $\mathbb{T} = [0, 1)$, es decir, el intervalo unitario semiabierto. Esto nos permite extender la definición del determinante de un retículo para todo retículo de dimensión no necesariamente completa tomando $det(\mathcal{L}) := vol(\mathcal{P}(\mathbf{B}))$, donde $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_d}] \in M_{n\times d}(\mathbb{R})$ es una matriz de rango $d \leq n$. En este caso, siguiendo [9, Theorem 7], se tiene que

$$det(\mathcal{L}) := vol(\mathcal{P}(\mathbf{B})) = \sqrt{det(\mathbf{B^T B})}. \tag{1.2}$$

No es difícil observar que es posible rellenar el espacio vectorial generado por un retículo $\mathcal{L}$ con copias trasladadas del paralelepípedo fundamental $\mathcal{P}(\mathbf{B})$, siendo $\mathbf{B}$ una base que genere a $\mathcal{L}$. Véase como ejemplo la Figura 1.1, donde el área sombreada coincide con el paralelepípedo fundamental. Formalmente hablando, $\mathcal{P}(\mathbf{B})$ es una *región fundamental*.

Definición 1.13. *Sea $\mathcal{L}$ un retículo de dimensión completa. Dado $S \subset \mathbb{R}^n$, decimos que S es una región fundamental del retículo si el conjunto $\{\mathbf{v} + S \mid \mathbf{v} \in \mathcal{L}\}$ forma una partición de $\mathbb{R}^n$.*

Equivalentemente, S es una región fundamental de un retículo $\mathcal{L}$ si para todo punto $\mathbf{t} \in \mathbb{R}^n = span(\mathcal{L})$, existe un único $\mathbf{v} \in \mathcal{L}$ tal que $\mathbf{t} \in \mathbf{v} + S$.

Notación. Sea $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$. En lo que sigue, usamos la notación $\lfloor \mathbf{x} \rfloor := (\lfloor x_1 \rfloor, \ldots, \lfloor x_n \rfloor)$, donde $\lfloor x_i \rfloor$ indica el mayor número entero igual o menor que x_i. Análogamente, usamos $\lfloor \mathbf{x} \rceil := (\lfloor x_1 \rceil, \ldots, \lfloor x_n \rceil)$ para denotar el redondeo usual, aplicado a cada una de las coordenadas de $\mathbf{x}$.

Proposición 1.14. *Sea $\mathcal{L}$ un retículo y $\mathbf{B} \in M_{n\times n}(\mathbb{R})^*$ una base de $\mathcal{L}$. El paralelepípedo fundamental $\mathcal{P}(\mathbf{B})$ asociado a la base $\mathbf{B}$ constituye una región fundamental del retículo.*

Demostración. Para comprobar que $\mathcal{P}(\mathbf{B})$ es una región fundamental de $\mathcal{L}$, debemos ver que $\{\mathbf{v} + \mathcal{P}(\mathbf{B}) \mid \mathbf{v} \in \mathcal{L}\}$ forma una partición de $span(\mathbf{B}) = \mathbb{R}^n$. En efecto, sea $t \in \mathbb{R}^n$, dado que $span(\mathbf{B}) = \mathbb{R}^n$, existe $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ tal que

$$\mathbf{t} = \mathbf{b_1}x_1 + \cdots + \mathbf{b_n}x_n = \mathbf{Bx}.$$

Si reescribimos $\mathbf{x}$ como

$$\mathbf{x} = \lfloor (x_1, \ldots, x_n) \rfloor + (x_1 - \lfloor x_1 \rfloor, \ldots, x_n - \lfloor x_n \rfloor),$$

se tiene que $\lfloor x \rfloor \in \mathbb{Z}^n$ y $(x_1 - \lfloor x_1 \rfloor, \ldots, x_n - \lfloor x_n \rfloor) \in [0,1)^n$. De esto se sigue que $t \in v + \mathcal{P}(\mathbf{B})$, donde $\mathbf{v} = \mathbf{B}\lfloor \mathbf{x} \rfloor$. Para concluir que $\mathcal{P}(\mathbf{B})$ es una partición, falta ver que este $\mathbf{v}$ es único. Supongamos por reducción al absurdo que existen $\mathbf{v}_1$, $\mathbf{v}_2 \in \mathcal{L}$ distintos e $\mathbf{y}_1, \mathbf{y}_2 \in [0,1)^n$, tales que

$$\mathbf{t} = \mathbf{v}_1 + \mathbf{B}\mathbf{y}_1 \in \mathbf{v}_1 + \mathcal{P}(\mathbf{B}),$$

$$\mathbf{t} = \mathbf{v}_2 + \mathbf{B}\mathbf{y}_2 \in \mathbf{v}_2 + \mathcal{P}(\mathbf{B}).$$

Restando ambas expresiones llegamos a que,

$$\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{B}(\mathbf{y}_2 - \mathbf{y}_1) \tag{1.3}$$

y dado que $\mathbf{v}_1 - \mathbf{v}_2 \in \mathcal{L}$, el vector $\mathbf{B}(\mathbf{y}_2 - \mathbf{y}_1)$ también debe pertener al retículo, lo que implica que $\mathbf{y}_2 - \mathbf{y}_1 \in \mathbb{Z}^n$. No obstante, sabemos que $\mathbf{y}_1$, $\mathbf{y}_2 \in [0,1)^n$, así que la única opción es que $\mathbf{y}_2 - \mathbf{y}_1$ sea igual al vector $\mathbf{0}$. De esta manera llegamos a una contradicción, pues de (1.3) se sigue que $\mathbf{v}_1 = \mathbf{v}_2$ y esto es absurdo por hipótesis.

□

Otro ejemplo de región fundamental de un retículo es el *paralelepípedo centrado semiabierto*, el cual se define como

$$\mathcal{C}(\mathbf{B}) := \left\{ \mathbf{Bx} \mid -\frac{1}{2} \leq \mathbf{x} < \frac{1}{2} \right\}.$$

Para demostrar que efectivamente $\mathcal{C}(\mathbf{B})$ también es una región fundamental, nos basta con repetir el procedimiento seguido en la demostración de la Proposición 1.14, pero considerando esta vez

$$\mathbf{t} = \mathbf{B}\lfloor (x_1, \ldots, x_n) \rceil + \mathbf{B}(x_1 - \lfloor x_1 \rceil, \ldots, x_n - \lfloor x_n \rceil).$$

Asimismo, si cambiamos la base $\mathbf{B}$ por su ortogonalización de Gram-Schmidt, obtenemos el *paralelepípedo centrado ortogonal* $\mathcal{C}(\mathbf{B}^*)$, el cual es clave en la construcción de algoritmos asociados a retículos y se define como sigue:

$$\mathcal{C}(\mathbf{B}^*) := \left\{ \mathbf{B}^*\mathbf{x} \mid -\frac{1}{2} \leq \mathbf{x} < \frac{1}{2} \right\}.$$

Proposición 1.15. *Sea $\mathcal{L}$ un retículo y $\mathbf{B} \in M_{n\times n}(\mathbb{R})^*$ una base de $\mathcal{L}$. El paralelepípedo $\mathcal{C}(\mathbf{B}^*)$, donde $\mathbf{B}^*$ es la ortogonalización de Gram-Schmidt de $\mathbf{B}$, es una región fundamental del retículo.*

Demostración. Sea $\mathbf{B}$ una base de $\mathcal{L}$ y $\mathbf{B}^* = [\mathbf{b_1}^*, \ldots, \mathbf{b_n}^*]$ su corrrespondiente ortogonalización de Gram-Schmidt. Por definción, se tiene que $\mathbf{b_n}^*$ es ortogonal a $span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}])$ y, por tanto, podemos descomponer $\mathbb{R}^n$ en capas ortogonales a $\mathbf{b_n}^*$. Es decir, $\mathbb{R}^n$ admite una partición en hiperplanos de la forma

$$\{c\mathbf{b_n}^* + span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}]) \mid c \in \mathbb{R}\}. \tag{1.4}$$

Además, teniendo en cuenta que, por definición, $\mathbf{b_n}^* \in \mathbf{b_n} + span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}])$, observamos que para todo $c \in \mathbb{R}$, las siguientes capas coinciden

$$c\mathbf{b_n}^* + span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}]) = c\mathbf{b_n} + span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}]). \tag{1.5}$$

Ahora, sea $\mathbf{t} \in \mathbb{R}^n$ un punto cualquiera. Al ser (1.4) una partición, tenemos que existe un único $c_n \in \mathbb{R}$ tal que $\mathbf{t} \in c_n\mathbf{b_n}^* + span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}])$. De (1.5) se sigue que, descomponiendo $c_n = \lfloor c_n \rceil + c_n'$ con $c_n' = c_n - \lfloor c_n \rceil \in \left[-\frac{1}{2}, \frac{1}{2}\right)$, el hiperplano que contiene a $\mathbf{t}$ puede escribirse como $\lfloor c_n \rceil \mathbf{b_n} + c_n'\mathbf{b_n}^* + span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}])$. En otras palabras, el vector $\mathbf{t}_{n-1} := \mathbf{t} - \lfloor c_n \rceil \mathbf{b_n} - c_n'\mathbf{b_n}^* \in span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}])$.

Repetimos el proceso particionando el hiperplano $span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}])$ en capas $\{c\mathbf{b_{n-1}}^* + span([\mathbf{b_1}, \ldots, \mathbf{b_{n-2}}]) \mid c \in \mathbb{R}\}$, y obtenemos $\mathbf{t_{n-2}} = \mathbf{t_{n-1}} - \lfloor c_{n-1} \rceil \mathbf{b_{n-1}} - c_{n-1}'\mathbf{b_{n-1}}^* \in span([\mathbf{b_1}, \ldots, \mathbf{b_{n-2}}])$ con $c_{n-1}' \in \left[-\frac{1}{2}, \frac{1}{2}\right)$. Continuando con el procedimiento de forma recursiva, se obtiene en última instancia

$$\mathbf{t}_0 := \mathbf{t} - \sum_{i=1}^{n} \lfloor c_i \rceil \mathbf{b_i} - \sum_{i=1}^{n} c_i'\mathbf{b_i} \in span(\{0\}) = \{0\},$$

lo que implica que

$$\mathbf{t} = \sum_{i=1}^{n} \lfloor c_i \rceil \mathbf{b_i} + \sum_{i=1}^{n} c_i'\mathbf{b_i} \in \mathcal{L}(\mathbf{B}) + \mathcal{C}(\mathbf{B}^*),$$

pues $\lfloor c_i \rceil \in \mathbb{Z}$ y $c_i' \in [-\frac{1}{2}, \frac{1}{2})$ para todo $i = 1, \ldots, n$.

Se concluye que para todo $\mathbf{t} \in \mathbb{R}^n$, podemos encontrar $\mathbf{v} \in \mathcal{L}$, que además es único, y $\mathbf{x} \in \mathcal{C}(\mathbf{B}^*)$ tales que $\mathbf{t} = \mathbf{v} + \mathbf{x}$, resultando así que $\mathcal{C}(\mathbf{B}^*)$ es una región fundamental del retículo.

□

2

Problemas basados en retículos

En este capítulo profundizamos aún más en la geometría de los retículos. En concreto, estudiaremos parámetros que nos permiten hacernos una idea sobre las distancias que se establecen entre los distintos puntos del retículo. Además introducimos dos problemas fundamentales de retículos que se consideran problemas NP-completos y, por tanto, tienen aplicaciones en criptografía.

2.1. Parámetros fundamentales

Dado un retículo $\mathcal{L}$, dos parámetros que miden cómo de lejos o de cerca se encuentran los puntos del retículo son el *radio recubridor* y el *radio de empaquetamiento*.

Definición 2.1. *El radio recubridor de un retículo $\mathcal{L}$ se denota como $\rho(\mathcal{L})$ y es el menor número real r tal que las esferas de radio r centradas en cada uno de los puntos de $\mathcal{L}$ cubren $\mathbb{R}^n$. Es decir, es el menor $r \in \mathbb{R}^+$ tal que $\bigcup_{\mathbf{v}\in\mathcal{L}} \overline{\mathcal{B}(\mathbf{v}, r)} = \mathbb{R}^n$.*

Este radio nos aporta información geométrica sobre qué tan lejos se encuentran los puntos del retículo entre sí. Por otra parte, si consideramos las mismas esferas pero disminuimos su radio hasta un valor r lo suficientemente pequeño tal que estas son disjuntas, aparece el concepto radio de empaquetamiento, el cual nos ayuda a comprender qué tan cerca pueden encontrarse dos puntos del retículo.

Definición 2.2. *El radio de empaquetamiento de un retículo $\mathcal{L}$ se denota como $\mu(\mathcal{L})$ y es el mayor número real r tal que las esferas de radio r centradas en cada uno de los puntos de $\mathcal{L}$ no se intersectan. Es decir, es el mayor $r \in \mathbb{R}^+$, tal que $\mathcal{B}(\mathbf{x}, r) \cap \mathcal{B}(\mathbf{y}, r) = \emptyset$ para todo $\mathbf{x}, \mathbf{y} \in \mathcal{L}$.*

Notamos que, en general, si tenemos $\mathbf{x}$ e $\mathbf{y} \in \mathcal{L}$ y tomamos las bolas $\mathcal{B}(\mathbf{x}, r)$ y $\mathcal{B}(\mathbf{y}, r)$, éstas no se van a intersectar siempre y cuando r sea a lo sumo la mitad de $dist(\mathbf{x}, \mathbf{y})$. Esto es cierto para cualquier par de puntos del retículo. En

particular, si $\mathbf{x}$ e $\mathbf{y}$ se encuentran lo más cerccano posible, $\mu(\mathcal{L})$ debe ser como mucho la mitad de esta distancia, a la que se denomina *distancia mínima*. De hecho, no es difícil ver que $\mu(\mathcal{L})$ es exactamente la mitad de dicha distancia.

Definición 2.3. *Sea $\mathcal{L}$ un retículo, la distancia mínima de $\mathcal{L}$ es la menor distancia entre dos puntos cualesquiera del mismo, es decir,*

$$\lambda(\mathcal{L}) = inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{x}, \mathbf{y} \in \mathcal{L}, \mathbf{x} \neq \mathbf{y}\},$$

donde $\|\cdot\|$ denota la norma Euclídea usual.

Considerando que $\mathcal{L}$ tiene estructura de grupo con la suma usual, si $\mathbf{x}, \mathbf{y} \in \mathcal{L}$ son tales que $\lambda(\mathcal{L}) = \|\mathbf{x} - \mathbf{y}\|$, podemos tomar el vector $\mathbf{v} = \mathbf{x} - \mathbf{y} \in \mathcal{L}$ y así definir, de forma equivalente, la distancia mínima como la longitud del menor vector no nulo del retículo, es decir,

$$\lambda(\mathcal{L}) = inf\{\|\mathbf{v}\| : \mathbf{v} \in \mathcal{L} \setminus \{\mathbf{0}\}\}.$$

Nótese, que el valor $\lambda(\mathcal{L})$ siempre se va a alcanzar por ser $\mathcal{L}$ un conjunto discreto, por lo que la distancia mínima del retículo está bien definida. Determinar esta distancia no es un problema sencillo, y en gran medida depende de la base $\mathbf{B}$ dada. Veremos esto con mayor detalle al final del capítulo. De momento, podemos dar la siguiente cota inferior.

Teorema 2.4. *Sea $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$ base de un retículo y $\mathbf{B}^*$ su ortogonalización de Gram-Schmidt, se tiene que $\lambda(\mathcal{L}(\mathbf{B})) \geq \min_i \|\mathbf{b_i^*}\|$.*

Demostración. Sea $\mathbf{Bx}$ un punto no nulo cualquiera del retículo, donde $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$ y sea k el mayor índice para el cual $x_k \neq 0$. Probaremos que

$$\|\mathbf{Bx}\| \geq \|\mathbf{b_k}^*\| \geq min_i \|\mathbf{b_i^*}\|.$$

Así, como caso particular, tenemos que $\lambda(\mathcal{L}(\mathbf{B})) \geq min_i \|\mathbf{b_i^*}\|$.

Para comprobar esto, consideramos el producto escalar de $\mathbf{Bx} = \sum \mathbf{b_i} x_i$ y $\mathbf{b_k^*}$. Como $x_i = 0$ para todo $i > k$, se tiene que

$$\langle \mathbf{Bx}, \mathbf{b_k^*} \rangle = \sum_{i \leq k} \langle \mathbf{b_i} x_i, \mathbf{b_k^*} \rangle = x_k \langle \mathbf{b_k}, \mathbf{b_k^*} \rangle,$$

donde la última igualdad se obtiene al ser $\mathbf{b_k^*}$ ortogonal a $\{\mathbf{b}_1, \ldots, \mathbf{b}_{k-1}\}$ por definición de $\mathbf{b_k^*}$. Teniendo en cuenta la Proposición 1.10, tenemos que

$$\mathbf{b_k} = \mathbf{b_k}^* + \sum_{j<k} \mu_{k,j} \mathbf{b_j^*},$$

y se sigue que

$$x_k\langle \mathbf{b_k}, \mathbf{b_k^*}\rangle = x_k\langle \mathbf{b_k}^*, \mathbf{b_k^*}\rangle + \sum_{j<k} x_k\mu_{k,j}\langle \mathbf{b_j}^*, \mathbf{b_k^*}\rangle = x_k\langle \mathbf{b_k}^*, \mathbf{b_k^*}\rangle.$$

De nuevo, la última igualdad se obtiene por ortogonalidad de $\mathbf{b_k^*}$ y, en este caso, $\{\mathbf{b_1}^*, \ldots, \mathbf{b_{k-1}}^*\}$. Es decir, $\langle \mathbf{Bx}, \mathbf{b_k^*}\rangle = x_k\|\mathbf{b_k^*}\|^2$.

Por último, recordemos que Cauchy-Schwarz nos dice que $\|\mathbf{Bx}\| \cdot \|\mathbf{b_k^*}\| \geq |\langle \mathbf{Bx}, \mathbf{b_k^*}\rangle|$. Esto sumado a lo anterior nos lleva a que $\|\mathbf{Bx}\| \geq |x_k| \cdot \|\mathbf{b_k^*}\|$. Sabiendo que $x_k \in \mathbb{Z} \setminus \{0\}$ tenemos que $|x_k| \geq 1$, luego $\|\mathbf{Bx}\| \geq |x_k| \cdot \|\mathbf{b_k^*}\| \geq \|\mathbf{b_k^*}\|$, como queríamos demostrar.

□

La longitud del menor vector no nulo puede verse también como el radio de la menor bola centrada en el origen que contiene a un vector no nulo del retículo. Con esta idea se extiende de forma natural la definición de distancia mínima como sigue:

Definición 2.5. *Denominamos* mínimos sucesivos *del retículo $\mathcal{L}$ a la secuencia de parámetros $\lambda_1, \ldots, \lambda_n$, donde para todo $i = 1, \ldots, n$, $\lambda_i(\mathcal{L}) = \lambda_i$ es el menor real positivo tal que la bola $\overline{\mathcal{B}(\mathbf{0}, \lambda_i)} = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq \lambda_i\}$ de radio λ_i centrada en el origen contiene al menos i vectores linealmente independientes.*

En efecto, esto generaliza el concepto de distancia mínima pues $\lambda_1 = \lambda(\mathcal{L})$.

La naturaleza discreta de los retículos nos permite afirmar que siempre podemos encontrar vectores linealmente independientes con una longitud dada por los mínimos sucesivos. Es decir, para todo retículo $\mathcal{L}$, existen $\mathbf{v_1}, \ldots, \mathbf{v_n} \in \mathcal{L}$ linealmente independientes tales que $\|\mathbf{v_i}\| = \lambda_i(\mathcal{L})$ para todo $i = 1, \ldots, n$. En este caso, dada la definición de mínimos sucesivos, la intuición nos puede llevar a pensar en un primer momento que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ constituye una base de $\mathcal{L}$ ya que son linealmente independientes y son del menor tamaño posible, sin embargo, esto no es cierto en general.

Ejemplo 2.6. Sea $n \geq 5$ y sea $\mathcal{L}$ el subretículo de $\mathbb{Z}^n$ formado por los vectores cuyas coordenadas son o bien todas pares o bien todas impares. $\mathcal{L}$ es un retículo pues es un subgrupo de $\mathbb{Z}^n$ y una posible base es la formada por los vectores $\mathbf{b_i} = \mathbf{2e_i}$ para todo $i = 1, \ldots, n-1$ y $\mathbf{b_n} = (1, 1, \ldots, 1)$. No es difícil ver que, si $\mathbf{v} \in \mathcal{L}$ es un vector de coordenadas todas pares, entonces $\|\mathbf{v}\| \geq \|\mathbf{2e_i}\| = 2$, mientras que, si $\mathbf{v}$ tiene coordenadas todas impares $\|\mathbf{v}\| \geq \|(1, 1, \ldots, 1)\| = \sqrt{n} > 2$, pues $n \geq 5$. De esto se deduce que $\lambda_i = 2$ para todo $i = 1, \ldots, n$, pues $S = \{\mathbf{2e_1}, \ldots, \mathbf{2e_n}\}$ es un conjunto de vectores linealmente independientes en $\overline{\mathcal{B}(\mathbf{0}, 2)}$ y no existen vectores no nulos en $\mathcal{L}$ de menor longitud. Sin embargo, S no es una base de $\mathcal{L}$ pues no hay forma de cubrir los vectores de coordenadas impares con ella. Por este motivo, deducimos que toda base de $\mathcal{L}$ debe contener al menos un vector $\mathbf{b}$ de coordenadas todas impares, y que además $\|\mathbf{b}\| \geq \sqrt{n} > 2$. Por tanto, no existe ninguna base $\mathbf{B}$ de $\mathcal{L}$ que cumpla que $\mathbf{B} \subseteq \overline{\mathcal{B}(\mathbf{0}, \lambda_n = 2)}$.

2.2. Acotando los parámetros fundamentales

Observamos, de las propias definiciones, que estos parámetros fundamentales dependen exclusivamente del retículo $\mathcal{L}$ y no de la base elegida, como es natural. Determinar estos parámetros es uno de los grandes y más fundamentales problemas sobre retículos, pues su descripción da información sobre la geometría del retículo. Dedicamos esta sección a determinar algunas de ellas.

En primer lugar es fácil observar, que para todo retículo n-dimensional $\mathcal{L}$, se cumple que

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n, \tag{2.1}$$

pues si $\overline{\mathcal{B}(\mathbf{0}, \lambda_i)}$ contiene i vectores linealmente independientes, también contendrá $i-1$, es decir, $\lambda_{i-1} \leq \lambda_i$, para todo $i = 2, 3, \ldots, n$.

Por otra parte, el radio recubridor ρ y λ_n cumplen

$$\lambda_n \leq 2\rho \leq \sqrt{n}\lambda_n.$$

Empezamos demostrando la primera desigualdad.

Proposición 2.7. *Para cualquier retículo $\mathcal{L}$, el radio recubridor verifica $\rho(\mathcal{L}) \geq \lambda_n/2$.*

Demostración. Supongamos por reducción al absurdo que $\rho(\mathcal{L}) < \lambda_n/2$, es decir, $\epsilon = \frac{1}{2}\lambda_n - \rho > 0$ es un valor real fijo y positivo. Veremos que de ser así, podemos encontrar n vectores linealmente independientes a distancia $\lambda_n - \epsilon$ del origen, que es absurdo por definición de λ_n.

Sea $\mathbf{t}_1 \in \mathbb{R}^n$ tal que $\|\mathbf{t}_1\| = \rho + \epsilon$. Por definición de radio recubridor, existe $\mathbf{v}_1 \in \mathcal{L}$ tal que $\|\mathbf{t_1} - \mathbf{v_1}\| \leq \rho$. De esto tenemos que $\mathbf{v}_1$ no es el vector cero, por cómo hemos elegido $\mathbf{t}_1$. Además, por la desigualdad triangular, $\|\mathbf{v}_1\| = \|\mathbf{t}_1 + (\mathbf{v}_1 - \mathbf{t}_1)\| \leq \|\mathbf{t}_1\| + \|\mathbf{v}_1 - \mathbf{t}_1\| \leq (\rho + \epsilon) + \rho = 2\rho + \epsilon = \lambda_n - \epsilon$. Supongamos ahora inductivamente que hemos construido un conjunto de vectores linealmente independientes $\{\mathbf{v_1}, \ldots, \mathbf{v_{i-1}}\} \subseteq \overline{\mathcal{B}(\mathbf{0}, \lambda_\mathbf{n} - \epsilon)}$, con $i \in \{2, \ldots, n\}$. Denotemos $V_{i-1} = span(\{\mathbf{v_1}, \ldots, \mathbf{v_{i-1}}\})$ y sea $\mathbf{t}_i \in \mathbb{R}^n$ un vector ortogonal a V_{i-1} tal que $\|\mathbf{t}_i\| = \rho + \epsilon$. De la misma forma, por definición de ρ, existe $\mathbf{v}_i \in \mathcal{L}$ tal que $\|\mathbf{t_i} - \mathbf{v_i}\| \leq \rho$. Al ser $\mathbf{t}_i$ ortogonal a V_{i-1} y de longitud $\rho + \epsilon$, se deduce que $\mathbf{v}_i \notin V_{i-1}$, luego $\mathbf{v_1}, \ldots, \mathbf{v_i}$ son linealmente independientes y $\|\mathbf{v}_i\| \leq \|\mathbf{t}_i\| + \|\mathbf{v}_i - \mathbf{t}_i\| \leq 2\rho + \epsilon = \lambda_n - \epsilon$. Esto demuestra la existencia de $\mathbf{v_1}, \ldots, \mathbf{v_n} \in \mathcal{L}$ vectores linealmente independientes y de longitud a lo sumo $\lambda_n - \epsilon$, lo que es absurdo por definición de λ_n.

□

Proposición 2.8. *Para cualquier retículo $\mathcal{L}$, se cumple que $\rho(\mathcal{L}) \leq \frac{\sqrt{n}}{2}\lambda_n(\mathcal{L})$.*

Demostración. Sea $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ un conjunto de vectores linealmente independientes tales que $\|\mathbf{v_i}\| \leq \lambda_\mathbf{i}$ para todo $i = 1, \ldots, n$. Sea $\mathcal{L}' = \mathcal{L}(\mathbf{V})$ el subretículo de $\mathcal{L}$ definido por la matriz $\mathbf{V} = [\mathbf{v_1}, \ldots, \mathbf{v_n}]$. De la definición de radio recubridor es fácil ver que $\rho \leq \rho' := \rho(\mathcal{L}')$. Veamos que $\rho' \leq \frac{\sqrt{n}}{2}\lambda_n$.

Para ello recordemos que la Proposición 1.15 nos dice que $\mathcal{C}(\mathbf{V}^*)$ es una región fundamental de $\mathcal{L}'$. Es decir, para todo $\mathbf{t} \in span(\mathcal{L}')$, existen $\mathbf{v} \in \mathcal{L}'$ y $\mathbf{x} \in \mathcal{C}(\mathbf{V}^*)$ tales que $\mathbf{t} = \mathbf{v} + \mathbf{x}$. Luego, $\mathbf{t} \in \overline{\mathcal{B}(\mathbf{v}, \|\mathbf{x}\|)}$. Esto es, todo punto de $\mathbb{R}^n$ está a distancia $\|\mathbf{x}\|$ del retículo para algún $\mathbf{x} \in \mathcal{C}(\mathbf{V}^*)$. Por definición de ρ', esto implica que $\rho' \leq max\{\|\mathbf{x}\| \mid \mathbf{x} \in \mathcal{C}(\mathbf{V}^*)\}$. Para concluir la demostración basta con ver que $\|\mathbf{x}\| \leq \frac{\sqrt{n}}{2}\lambda_n$ para todo $\mathbf{x} \in \mathcal{C}(\mathbf{V}^*)$. En efecto, sea $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{v}_i^* \in \mathcal{C}(\mathbf{V}^*)$ con $-\frac{1}{2} \leq x_i < \frac{1}{2}$. Como $\{\mathbf{v_1}^*, \ldots, \mathbf{v_n}^*\}$ es un conjunto de vectores ortogonales dos a dos, el teorema de Pitágoras nos dice que:

$$\|\mathbf{x}\|^2 = \|x_1\mathbf{v_1}^*\|^2 + \cdots + \|x_n\mathbf{v_n}^*\|^2 = |x_1|^2\|\mathbf{v_1}^*\|^2 + \cdots + |x_n|^2\|\mathbf{v_n}^*\|^2 \leq$$

$$\leq \left(\frac{1}{2}\right)^2 \|\mathbf{v_1}^*\| + \cdots + \left(\frac{1}{2}\right)^2 \|\mathbf{v_n}^*\|,$$

Asimismo, teniendo en cuenta que $\|\mathbf{v_i}^*\| \leq \|\mathbf{v_i}\|$ para todo $i = 1, \ldots, n$, se tiene que

$$\left(\frac{1}{2}\right)^2 \|\mathbf{v_1}^*\| + \cdots + \left(\frac{1}{2}\right)^2 \|\mathbf{v_n}^*\| \leq \frac{1}{4}(\|\mathbf{v_1}\|^2 + \cdots + \|\mathbf{v_n}\|^2) \leq$$

$$\leq \frac{1}{4}(\lambda_n^2 + \cdots + \lambda_n^2) = \frac{n}{4}\lambda_n^2.$$

Tomando raíces cuadradas, se concluye que $\|\mathbf{x}\| \leq \frac{\sqrt{n}}{2}\lambda_n$, como queríamos ver.

□

Esto nos permite acotar los mínimos sucesivos utilizando el radio recubridor y viceversa. A continuación, vemos como acotar el menor vector no nulo del retículo mediante el determinante, resultado que demostró el matemático alemán Hermann Minkowski en [7].

Teorema 2.9. *(Teorema de Minkowski) Sea $\mathcal{L}$ un retículo n-dimensional con $n \geq 2$, se tiene que*

$$\lambda(\mathcal{L}) < \left(\prod_{i=1}^{n} \lambda_i(\mathcal{L})\right)^{1/n} < \sqrt{n} \cdot det(\mathcal{L})^{1/n}$$

La primera desigualdad se deduce fácilmente de (2.1). La demostración de la segunda desigualdad no es tan evidente, para ello introducimos primero el *Teorema de Blichfeldt* y un corolario del mismo al que se conoce como *Teorema del cuerpo convexo.*

Teorema 2.10. *(Teorema de Blichfeldt) Dados un retículo $\mathcal{L}$ y un conjunto $S \subseteq \mathbb{R}^n$, si $vol(S) > det(\mathcal{L})$ entonces S contiene dos puntos distintos $\mathbf{z_1}, \mathbf{z_2} \in S$ tales que $\mathbf{z_1} - \mathbf{z_2} \in \mathcal{L}$.*

Demostración. Sea $\mathbf{B}$ una base cualquiera del retículo, podemos definir la familia de conjuntos $S_\mathbf{x} := S \cap (\mathbf{x} + \mathcal{P}(\mathbf{B}))$, donde $\mathbf{x} \in \mathcal{L}(\mathbf{B})$. Recordamos de la Proposición 1.14 que $\mathcal{P}(\mathbf{B})$ es una región fundamental, es decir, que $\{\mathbf{x} + \mathcal{P}(\mathbf{B}) \mid \mathbf{x} \in \mathcal{L}(\mathbf{B})\}$ es una partición de $\mathbb{R}^n$. Por tanto, los conjuntos $S_\mathbf{x}$ forman una partición de S. Se sigue así que

$$vol(S) = \sum_{\mathbf{x} \in \mathcal{L}(\mathbf{B})} vol(S_\mathbf{x}).$$

Consideramos ahora los conjuntos trasladados $S_\mathbf{x} - \mathbf{x} = (S - \mathbf{x}) \cap \mathcal{P}(\mathbf{B})$, los cuales están contenidos en $\mathcal{P}(\mathbf{B})$. Puesto que $vol(S_\mathbf{x}) = vol(S_\mathbf{x} - \mathbf{x})$ y por hipótesis $vol(S) > det(\mathcal{L})$, se tiene que

$$vol(\mathcal{P}(\mathbf{B})) = det(\mathcal{L}(\mathbf{B})) < vol(S) = \sum_{\mathbf{x} \in \mathcal{L}(\mathbf{B})} vol(S_\mathbf{x}) = \sum_{\mathbf{x} \in \mathcal{L}(\mathbf{B})} vol(S_\mathbf{x} - \mathbf{x}).$$

De esta manera, al estar $S_\mathbf{x} - \mathbf{x} \subseteq \mathcal{P}(\mathbf{B})$ y ser $\sum_{\mathbf{x} \in \mathcal{L}(\mathbf{B})} vol(S_\mathbf{x} - \mathbf{x}) > vol(\mathcal{P}(\mathbf{B}))$, se deduce que no todos los conjuntos pueden ser disjuntos, es decir, existen dos vectores distintos $\mathbf{x}, \mathbf{y} \in \mathcal{L}(\mathbf{B})$ tales que $(S_\mathbf{x} - \mathbf{x}) \cap (S_\mathbf{y} - \mathbf{y}) \neq \emptyset$. Por ello, podemos tomar $\mathbf{z} \in (S_\mathbf{x} - \mathbf{x}) \cap (S_\mathbf{y} - \mathbf{y})$ y definir

$$\mathbf{z_1} := \mathbf{z} + \mathbf{x} \in S_\mathbf{x} \subseteq S,$$
$$\mathbf{z_2} := \mathbf{z} + \mathbf{y} \in S_\mathbf{y} \subseteq S,$$

siendo así $\mathbf{z_1}$, $\mathbf{z_2}$ dos vectores distintos de S tales que

$$\mathbf{z_1} - \mathbf{z_2} = \mathbf{x} - \mathbf{y} \in \mathcal{L}(\mathbf{B}),$$

como queríamos demostrar.

□

Corolario 2.11. *(Teorema del cuerpo convexo) Sea $\mathcal{L}$ un retículo y sea $S \subseteq \mathbb{R}^n$ un conjunto convexo y simétrico respecto al origen con volumen $vol(S) > 2^n det(\mathcal{L})$. Entonces S contiene un vector del retículo distinto del vector nulo.*

Demostración. Consideremos el conjunto $S' = S/2 = \{\mathbf{x} \mid 2\mathbf{x} \in S\}$. El volumen de S' satisface que

$$vol(S/2) = 2^{-n} vol(S) > det(\mathcal{L})$$

Aplicando el Teorema de Blichfeld, existen $\mathbf{z_1}, \mathbf{z_2} \in S'$ tales que $\mathbf{z_1} - \mathbf{z_2} \in \mathcal{L}\backslash\{\mathbf{0}\}$. Por definición de S' se tiene que $2\mathbf{z_1}$, $2\mathbf{z_2} \in S$ y, por simetría de S, $-2\mathbf{z_2} \in S$. Finalmente, por convexidad de S, el punto medio del segmento con extremos $2\mathbf{z_1}$ y -$2\mathbf{z_2}$ debe pertenecer a S, es decir,

$$\frac{2\mathbf{z_1} + (-2\mathbf{z_2})}{2} = \mathbf{z_1} - \mathbf{z_2} \in S$$

es un vector no nulo del retículo contenido en el conjunto S.

□

Una vez demostrado el Teorema del cuerpo convexo, podemos demostrar la segunda desigualdad del Teorema de Minkowski. Para ello, necesitamos también el siguiente lema técnico referente a la función gamma Γ. Recordamos que dicha función cumple la recursividad $\Gamma(x+1) = x\Gamma(x)$ y que $\Gamma(1) = 1$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Lema 2.12. *Sea Γ la función gamma. Se tiene que para todo $n \in \mathbb{N}$ con $n \geq 2$,*

$$\Gamma\left(\frac{n}{2}+1\right) \leq \left(\frac{n}{2}\right)^{\frac{n}{2}}.$$

Demostración. La recursividad de Γ nos permite demostrar el resultado para todo n sabiendo que es cierto para $n-2$. Procedemos así por inducción sobre n y demostramos en primer lugar los dos casos base.

- Si $n = 2$, se cumple que

$$\Gamma\left(\frac{2}{2}+1\right) = \Gamma(2) = 1 \cdot \Gamma(1) = 1 = \left(\frac{2}{2}\right)^{\frac{2}{2}}.$$

- Si $n = 3$,

$$\Gamma\left(\frac{3}{2}+1\right) = \frac{3}{2} \cdot \Gamma\left(\frac{3}{2}\right) = \frac{3}{2} \cdot \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{3}{4}\sqrt{\pi} \leq \frac{3}{4}\sqrt{6} = \left(\frac{3}{2}\right)^{\frac{3}{2}}.$$

  Supongamos ahora cierto para $n-2$. Entonces,

$$\Gamma\left(\frac{n}{2}+1\right) = \frac{n}{2} \cdot \Gamma\left(\frac{n}{2}\right) = \frac{n}{2} \cdot \Gamma\left(\frac{n-2}{2}+1\right),$$

  y aplicando la hipótesis de inducción se sique que

$$\Gamma\left(\frac{n}{2}+1\right) \leq \frac{n}{2} \cdot \left(\frac{n-2}{2}\right)^{\frac{n-2}{2}} \leq \frac{n}{2} \cdot \left(\frac{n}{2}\right)^{\frac{n-2}{2}} = \left(\frac{n}{2}\right)^{\frac{n}{2}}.$$

□

Con esto ya estamos en disposición de demostrar el Teorema de Minkowki.

Demostración (Teorema 2.9). La desigualdad a demostrar es

$$\left(\prod_{i=1}^{n} \lambda_i(L)\right)^{1/n} < \sqrt{n} \cdot det(L)^{1/n}.$$

Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ vectores linealmente independientes del retículo cumpliendo que $\|\mathbf{v}_i\| = \lambda_i$. Asumamos por reducción al absurdo que $\prod_{i=1}^{n} \lambda_i \geq (\sqrt{n})^n \cdot det(\mathcal{L})$. En lo que sigue, consideramos la base $B = \{\mathbf{v}_1^*, \ldots, \mathbf{v}_n^*\}$ del espacio vectorial $\mathbb{R}^n$, donde los $\mathbf{v}_i^*$ son los obtenidos mediante la ortogonalización de Gram-Schmidt de $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ la transformación lineal definida como sigue

$$T\left(\sum_{i=1}^{n} c_i \mathbf{v}_i^*\right) = \sum_{i=1}^{n} \lambda_i c_i \mathbf{v}_i^*,$$

que expande cada coordenada $\mathbf{v}_i^*$ por un factor λ_i, y sea $S = \mathcal{B}(\mathbf{0}, 1)$ la bola abierta n-dimensional de radio 1 y centrada en el origen. Si aplicamos la transformación T al conjunto S, se obtiene un conjunto simétrico convexo $T(S)$ de volumen

$$vol(T(S)) = \left(\prod_{i=1}^{n} \lambda_i\right) vol(S) \geq (\sqrt{n})^n \cdot det(\mathcal{L}) vol(S).$$

Recordamos que el volumen de una esfera n-dimensional de radio 1 viene dado por $V_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)}$, por lo que aplicando el Lema 2.12 se tiene que

$$vol(S) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} \geq \frac{\pi^{\frac{n}{2}}}{\left(\frac{n}{2}\right)^{\frac{n}{2}}} = \frac{2^{\frac{n}{2}} \pi^{\frac{n}{2}}}{n^{\frac{n}{2}}} > \frac{2^n}{n^{\frac{n}{2}}}.$$

Luego, $vol(T(S)) > 2^n det(\mathcal{L})$, por lo que nos encontramos en las hipótesis del Teorema del cuerpo convexo y podemos afirmar que $T(S)$ contiene un punto $\mathbf{y}$ del retículo diferente del origen. Al pertenecer $\mathbf{y}$ al conjunto $T(S)$, se debe dar que $\mathbf{y} = T(\mathbf{x})$ para algún $\mathbf{x} \in \mathbb{R}^n$ con $\|\mathbf{x}\| < 1$, por definición de S. Expresando $\mathbf{x}$ e $\mathbf{y}$ en función de la base ortogonal se tiene

$$\mathbf{x} = \sum_{i=1}^{n} c_i \mathbf{v}_i^* \qquad e \qquad \mathbf{y} = \sum_{i=1}^{n} \lambda_i c_i \mathbf{v}_i^*.$$

Dado que $\mathbf{y} \neq \mathbf{0}$, existe $i \in \{1, \ldots, n\}$ tal que $c_i \neq 0$. Sea k el mayor índice para el que $c_k \neq 0$, y sea $k' \leq k$ el índice más pequeño cumpliendo que $\lambda_{k'} = \lambda_k$. Teniendo en cuenta que $c_k \neq 0$ y que B es una base, se tiene que $\mathbf{y}$ es linealmente independiente de $\mathbf{v}_1^*, \ldots, \mathbf{v}_{k'-1}^*$. Por tanto, $\mathbf{y}$ es linealmente independiente a $\mathbf{v}_1, \ldots, \mathbf{v}_{k'-1}$, pues generan el mismo espacio vectorial. Además, teniendo en cuenta que B es ortogonal, que tenemos (2.1) y que $\mathbf{x}$ es tal que $\|\mathbf{x}\| < 1$, se tiene que

$$\|\mathbf{y}\|^2 = \sum_{i \leq k} \lambda_i^2 \|c_i \mathbf{v}_i^*\|^2 \leq \lambda_k^2 \sum_{i \leq k} \|c_i \mathbf{v}_i^*\|^2 = \lambda_k^2 \|\mathbf{x}\|^2 < \lambda_k^2.$$

De esta forma hemos encontrado k' vectores linealmente independientes $\mathbf{y}$, $\mathbf{v}_1$, $\ldots$, $\mathbf{v}_{k'-1}$ tales que $\|\mathbf{y}\| < \lambda_k = \lambda_{k'}$, y $\|\mathbf{v}_i\| = \lambda_i < \lambda_{k'}$ para todo $i = 1, \ldots, k'-1$, por cómo hemos elegido k'. Esto es absurdo por definición de $\lambda_{k'}$, donde el absurdo procede de asumir que no se cumple la segunda desigualdad del Teorema de Minkowski.

□

El Teorema de Minkowski, además de tener aplicaciones en criptografía de retículos, nos permite demostrar ciertos resultados de Teoría de Números, siendo un ejemplo de ello el siguiente teorema.

Teorema 2.13. *Para todo número primo $p \equiv 1 \ (mod\ 4)$ existen a, $b \in \mathbb{Z}$ tales que $p = a^2 + b^2$.*

Demostración. Para demostar este resultado, buscaremos un retículo entero $\mathcal{L} \subseteq \mathbb{Z}^2$ tal que $\lambda(\mathcal{L})^2 = p$. De esta manera, si (a, b) es el vector más pequeño del retículo, entonces se tiene que $p = a^2 + b^2$.

Sea $p \in \mathbb{Z}$ tal que $p \equiv 1 \ (mod\ 4)$, entonces $\mathbb{Z}_p^*$ es un grupo cíclico con el producto usual y se cumple que $4 \mid o(\mathbb{Z}_p^*) = p - 1$. Por tanto, va a existir un elemento $n \in \mathbb{Z}_p^*$ de orden 4, o lo que es lo mismo, podemos encontrar un entero n tal que $n^2 = -1 \ (mod\ p)$. De esto se sigue que

$$p \mid n^2 + 1. \tag{2.2}$$

Consideremos ahora el retículo $\mathcal{L}$ generado por la base

$$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ n & p \end{bmatrix},$$

el cual tiene determinante p. Luego, aplicando el Teorema de Minkowski, sabemos que existe un vector no nulo $\mathbf{v} \in \mathcal{L}(\mathbf{B})$ tal que $\|\mathbf{v}\|^2 < 2p$. Dado que la norma al cuadrado de todo vector del retículo es un múltiplo de p, sea $\mathbf{x} = (x_1, x_2) \in \mathbb{Z}^2$, se tiene que

$$\|\mathbf{Bx}\|^2 = x_1^2 + (nx_1 + px_2)^2 = (1 + n^2)x_1^2 + p(2nx_1x_2 + px_2^2)$$

es un múltiplo de p teniendo en cuenta (2.2). Es decir, hemos llegado a que todo vector $\mathbf{v} \in \mathcal{L}(\mathbf{B})$ es tal que $p \mid \|\mathbf{v}\|^2$ y que existe $\mathbf{v} \in \mathcal{L}(\mathbf{B}) \setminus \{\mathbf{0}\}$ con $\|\mathbf{v}\|^2 \leq 2p$. Esto implica la existencia de $\mathbf{v} = (a, b) \in \mathbb{Z}^2$ con $\|\mathbf{v}\|^2 = a^2 + b^2 = p$.

□

2.3. Problemas principales basados en retículo

Los problemas que se presentan en esta sección destacan por ser los más relevantes dentro del área de estudio de la critografía basada en retículos, dado que no se han encontrado actualmente algoritmos eficientes que sean capaces de encontrar una solución. Se conocen como el *El problema del vector más corto* o SVP, por sus siglas en inglés, y el *El problema del vector más cercano* o CVP. Ambos problemas se ha demostrado que son NP-completos (ver [3]).

Definición 2.14. *El 'Problema del vector más corto' consiste en, dado un retículo $\mathcal{L} \subset \mathbb{R}^n$, encontrar un vector no nulo de menor longitud. Es decir, el SVP consiste en hallar $\mathbf{v} \in \mathcal{L}$ tal que $\|\mathbf{v}\| = \lambda(\mathcal{L})$.*

Si bien el SVP, en general, es díficil de resolver, si $\mathcal{L}$ es un retículo para el cual se conoce una base ortogonal, entonces el SVP tiene solución trivial. En efecto, sea $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$ una base ortogonal de $\mathcal{L}$, y $\mathbf{v} = \sum_{i=1}^{n} \mathbf{b_i} x_i \in \mathcal{L}(\mathbf{B})$ con $x_i \in \mathbb{Z}$, aplicando el Teorema de Pitágoras se tiene que

$$\|\mathbf{v}\|^2 = \sum_{i=1}^{n} |x_i|^2 \|\mathbf{b_i}\|^2.$$

Luego, $min\{\|\mathbf{v}\| \mid \mathbf{v} \in \mathcal{L}(\mathbf{B})\} = min_i \|\mathbf{b_i}\|$ y de esta manera queda resuelto el SVP.

El otro problema, en el que centraremos nuestro interés, se enuncia como sigue.

Definición 2.15. *El 'Problema del vector de más cercano' consiste en, dados un retículo $\mathcal{L} \subset \mathbb{R}^n$, un punto seleccionado $\mathbf{t} \in \mathbb{R}^n$, encontrar el punto del retículo más cercano a $\mathbf{t}$. Es decir, el CVP consiste en, dado $\mathbf{t} \in \mathbb{R}^n$, encontrar $v \in \mathcal{L}$ tal que $\|\mathbf{t} - \mathbf{v}\| \leq \|\mathbf{t} - \mathbf{v}'\|$ para todo $v' \in \mathcal{L}$.*

Al igual que sucedía con el SVP, el CVP tiene solución trivial si trabajamos con una base ortogonal del retículo. En efecto, sea $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$ una base ortogonal de $\mathcal{L}$, y $\mathbf{t} = \sum_{i=1}^{n} \mathbf{b_i} a_i \in span(\mathbf{B}) = \mathbb{R}^n$ con $a_i \in \mathbb{R}$. Sea $f : \mathbb{Z}^n \to \mathbb{R}^+$ tal que $f(\mathbf{x}) = \|\mathbf{t} - \mathbf{B}\mathbf{x}\|^2$. Se tiene que el vector más cercano viene dado por el $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n$ que minimiza f. Así, aplicando el Teorema de Pitágoras se tiene que

$$f(\mathbf{x}) = \left\| \sum_{i=1}^{n} \mathbf{b_i}(a_i - x_i) \right\|^2 = \sum_{i=1}^{n} \|\mathbf{b_i}\|^2 |(a_i - x_i)|^2.$$

De esto se sigue que el valor mínimo se alcanza en los $x_1, \ldots, x_n \in \mathbb{Z}$ que minimizan $|(a_i - x_i)|$, es decir, cuando $x_i = \lfloor a_i \rceil$ para todo $i = 1, \ldots, n$. Por

tanto, para resolver el CVP basta con tomar como vector más cercano a $\mathbf{Bx}$, donde $\mathbf{x} = \lfloor \mathbf{a} \rceil$ y $\mathbf{a}$ es el vector tal que $\mathbf{t} = \mathbf{Ba}$, es decir, $\mathbf{a} = \mathbf{B}^{-1}\mathbf{t}$.

Se ha comprobado que la resolución del SVP y del CVP es automática si trabajamos con retículos que admiten una base ortogonal y conocemos dicha base. No obstante, este procedimiento no resuelve estos problemas si no se conoce una base ortogonal del retículo. Por eso, en criptografía se consideran versiones aproximadas del SVP y el CVP. Sea $\gamma \geq 1$ un *factor de aproximación*, definimos γ-SVP como la versión γ-aproximada del problema SVP, la cual plantea hallar $\mathbf{v} \in \mathcal{L}$ tal que $\|\mathbf{v}\| \leq \gamma \cdot \lambda(\mathcal{L})$. Por otra parte, definimos la versión γ-aproximada del problema CVP como el problema que consiste en encontrar $\mathbf{v} \in \mathcal{L}$ tal que $\|\mathbf{t} - \mathbf{v}\| \leq \gamma \cdot \|\mathbf{t} - \mathbf{v}'\|$ para todo $\mathbf{v}' \in \mathcal{L}$. Está claro que cuando $\gamma = 1$, recuperamos la versión exacta de estos problemas. En el siguiente capítulo introduciremos algunas bases para las cuales se conocen algoritmos que resuelven versiones aproximadas de ambos problemas con un factor de aproximación razonable.

Se hace necesario introducir estas versiones aproximadas porque, además, como ya adelantamos en el capítulo anterior, no todo retículo admite una base ortogonal. Veámoslo en el siguiente ejemplo.

Ejemplo 2.16. Sea $\mathcal{L} = \mathcal{L}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -0.5 \end{smallmatrix}\right)$ el retículo de la Figura 1.1, veamos que no tiene ninguna base ortogonal.

Supongamos por reducción al absurdo que $\mathbf{B} = [\mathbf{b_1}, \mathbf{b_2}]$ es una base ortogonal de $\mathcal{L}$. Asumiendo sin pérdida de generalidad que $\|\mathbf{b_1}\| \leq \|\mathbf{b_2}\|$, se tiene que $\mathbf{b_1}$ es el vector más corto del retículo. Esto es así por lo visto a la hora de demostrar que el SVP tiene solución trivial si se emplea una base ortogonal. En el caso de este ejemplo, el vector más corto de $\mathcal{L}$ es $\mathbf{b_1} = (1, -0.5)$ o su opuesto. Sea ahora $\mathbf{b_2}$ un vector ortogonal a $\mathbf{b_1}$, es decir, $\mathbf{b_2} = (a, b)$ es tal que

$$\langle \mathbf{b_1}, \mathbf{b_2} \rangle = a - 0.5b = 0.$$

Despejando b obtenemos que $\mathbf{b_2}$ es de la forma $\mathbf{b_2} = (a, 2a)$ para algún $a \in \mathbb{R}$, que podemos asumir positivo, sin pérdida de generalidad. Luego, $\mathbf{B} = \left(\begin{smallmatrix} 1 & 2 \\ -0.5 & 2a \end{smallmatrix}\right)$.

Por otra parte, recordemos que el determinante es un invariante en los retículos. Por tanto, si $\mathbf{B}$ es una base de $\mathcal{L}$, se debe cumplir que

$$|det(\mathbf{B})| = \left| det\left(\begin{bmatrix} 1 & 1 \\ 1 & -0.5 \end{bmatrix} \right) \right| = \left| -\frac{3}{2} \right| = \frac{3}{2}.$$

Como $|det(\mathbf{B})| = 2a + 0.5a = \frac{5}{2}a$, obtenemos así que para que $\mathbf{B}$ sea base, se debe dar que $a = \frac{3}{5}$. Sin embargo, $\mathbf{b_2} = \left(\frac{3}{5}, \frac{6}{5}\right)$ no es un vector del retículo $\mathcal{L}$, pues no es difícil comprobar que no existen $x, y \in \mathbb{Z}$ tales que $\left(\frac{3}{5}, \frac{6}{5}\right) = x(1, 1) + y(1, -0.5)$. Se concluye así que $\mathcal{L}$ no tiene ninguna base que sea ortogonal.

3

Criptografía basada en retículos

Llegados a este punto en el que han sido introducidos los elementos indispensables de la teoría de retículos, nos encontramos en disposición de ver como se aplican estos en la criptografía. Para ello, haremos primero una breve introducción con los conceptos básicos sobre criptografía que nos permitan comprender la construcción del *criptosistema* GGH y además presentaremos un posible ataque al mismo.

3.1. Conceptos básicos sobre criptografía

La *criptografía* es la ciencia que se encarga del cifrado y el descifrado de los datos con la finalidad de hacerlos ininteligibles a terceros para los que no está destinado el mensaje.

Formalmente, al mensaje original que queremos transmitir se le conoce como *texto plano*, que se transforma de manera eficiente para obtener el *texto cifrado*, con el cual una tercera persona no tiene forma de conocer la información que se quiere transmitir. Esta transformación del mensaje es denominada *cifrado*, mientras que el proceso por el cual se recupera el mensaje original recibe el nombre de *descifrado*. Para que sean posible el cifrado y el descifrado, se requiere de una *clave* que permita llevar a cabo dichos procesos. En el caso de la *criptografía simétrica* o *criptografía de clave secreta* se utiliza la misma clave para ambos procesos. Sin embargo, en emphcriptografía asimétrica o *criptografía de clave pública*, la clave de función cifrado es diferente de la empleada para descifrar. En conjunto, estos conceptos nos permiten definir los *criptosistemas*. Estos se definen como una terna $(\mathcal{P}, \mathcal{C}, \mathcal{S})$ de conjuntos donde

$$\mathcal{P} \equiv \{\text{textos planos}\},$$

$$\mathcal{C} \equiv \{\text{textos cifrados}\},$$

$$\mathcal{S} \equiv \{\text{claves}\}.$$

Y las siguientes relaciones entre los conjuntos anteriores: *funciones de cifrado*

$$e_{\mathbf{s}} : \mathcal{P} \longrightarrow \mathcal{C} \text{ para todo } \mathbf{s} \in \mathcal{S}$$

y *funciones de descifrado*

$$d_{\mathbf{s}} : \mathcal{C} \longrightarrow \mathcal{P} \text{ para todo } \mathbf{s} \in \mathcal{S}.$$

Además, debe verificarse que

$$d_{\mathbf{s}}(e_{\mathbf{s}}(\mathbf{x})) = \mathbf{x} \text{ para todo } \mathbf{x} \in \mathcal{P} \text{ y para todo } \mathbf{s} \in \mathcal{S}.$$

El critosistema que tratamos en esta memoria pertenece a la familia de criptosistemas de clave pública o asimétricos. La principal diferencia con la criptografía simétrica, es que no se requiere que el emisor y el receptor tengan un intercambio inicial de la clave a utilizar.

La criptografía de clave pública nace con el objetivo de evitar el intercambio de claves privadas, tarea que, en la actualidad, debido a la gran cantidad de usuarios de las redes de comunicación se torna complicada.

En este tipo de criptosistemas, cada participante $\mathcal{A}$ dispone de dos claves, una clave pública $\mathbf{p}_{\mathcal{A}}$ que cualquier persona puede utilizar en la función de cifrado para enviar un mensaje a $\mathcal{A}$ y una clave privada $\mathbf{s}_{\mathcal{A}}$, que no comparte, y que utiliza $\mathcal{A}$ para leer sus mensajes. Es decir, descifrar los mensajes que le han enviado. Se tiene así, que la composición de ambas funciones devuelve el mensaje original $d_{\mathbf{s}_{\mathcal{A}}}(e_{\mathbf{p}_{\mathcal{A}}}(\mathbf{x})) = \mathbf{x}$.

3.2. Criptosistema de clave pública GGH

El criptosistema GGH presentado en 1997 y denominado así por ser propuesto por Goldreich, Goldwasser y Halevi en [4], se basa en la dificultad de resolver el problema del vector más cercano y está inspirado en el criptosistema de McEliece propuesto en 1978, el cual en vez de implicar retículos, hace uso de códigos lineales en cuerpos finitos. El criptosistema GGH parte de $\mathcal{P} \equiv \mathbb{Z}^n$, donde la dimensión n se fija previamente y es el parámetro de seguridad de nuestro sistema. Además, los textos cifrados son vectores de $\mathcal{C} \equiv \mathbb{R}^n$ y las claves vienen dadas por $\mathcal{S} \equiv \{\mathbf{B} \in M_{n\times n}(\mathbb{Z}) \mid \mathcal{L}(\mathbf{B}) \textit{ es un subretículo de } \mathbb{Z}^n\}$. En este contexto el sistema criptográfico GGH se establece como sigue:

- La clave privada es una 'buena' base $\mathbf{B} \in \mathcal{S}$ definiendo un retículo $\mathcal{L} = \mathcal{L}(\mathbf{B})$. Como veremos a continuación, una 'buena' base es una cuyos vectores son *casi ortogonales*.
- La clave pública es $\mathbf{H} \in \mathcal{S}$ es una 'mala' base para el retículo $\mathcal{L}$, es decir, $\mathcal{L}(\mathbf{H}) = \mathcal{L}(\mathbf{B})$. Esta base se puede obtener tomando una matriz entera $\mathbf{U} \in M_{n\times n}(\mathbb{Z})^*$ tal que $\mathbf{H} = \mathbf{UB}$ no es *muy ortogonal*.

- De esta forma, dado un mensaje a trasmitir $\mathbf{x} \in \mathcal{P}$, la función de cifrado $e_{\mathbf{H}} : \mathbb{Z}^n \to \mathbb{R}^n$ se define como la transformación lineal $e_{\mathbf{H}}(\mathbf{x}) = \mathbf{Hx} + \mathbf{r}$. El vector $\mathbf{r} \in \mathbb{R}^n$ es un *vector error* no demasiado grande y es escogido de manera aleatoria cada vez que se quiere mandar un mensaje. Así, el mensaje $\mathbf{x}$ se cifra en un punto $\mathbf{c} = \mathbf{Hx} + \mathbf{r}$ cercano al vector del retículo $\mathbf{v} = \mathbf{Hx} \in \mathcal{L}(\mathbf{B})$.
- El proceso de descifrado consiste en hallar el vector del retículo más cercano al mensaje cifrado recibido $\mathbf{c}$. Por tanto, descifrar corresponde a resolver una instancia del CVP, lo cual es posible usando la buena base $\mathbf{B}$ tal y como explicaremos en breve. Una vez hallado $\mathbf{v}$, se calcula $\mathbf{H}^{-1}\mathbf{v}$, recuperando así el mensaje original $\mathbf{x}$.

Se observa que efectivamente este criptosistema se basa en el CVP, ya que, para descifrar $\mathbf{c}$, basta con encontrar el vector del retículo más cercano a $\mathbf{c}$, que como hemos visto esto es un problema NP-duro en general. Sin embargo, al acotar el tamaño del vector error, realmente la seguridad se basa en una versión aproximada del CVP. De hecho, la confianza en la seguridad de este criptosistema se basa en la inexistencia de un ataque que lo rompa por completo para cualquier dimensión n del retículo.

3.2.1. Elección de una buena base

Al final del Capítulo 2, vimos como de tener una base ortogonal, se puede resolver el CVP. Sin embargo, también establecimos que, en general, los retículos no siempre admiten una base ortogonal. Por tanto, se hace necesario establecer un criterio para determinar cuando una base se considera *buena*. Lo natural, es decir que una base es buena cuando permite resolver el γ-CVP para valores de γ lo más próximo posibles a 1. De esta forma, el algoritmo que sigue el GGH para descifrar los mensajes recibidos, es capaz de encontrar el vector más cercano con un mayor nivel de error $\mathbf{r}$ permitido, cuanto más cerca de ser ortogonal sea la clave privada.

Recordamos de la Definición 1.12 que dada una base $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$ de $\mathcal{L}$, entonces $det(\mathcal{L}) = det(\mathbf{B}) = det(\mathbf{B}^*)$. Además, al ser $\mathbf{B}^*$ ortogonal, se tiene que

$$det(\mathbf{B}^*) = \prod_i \|\mathbf{b_i}^*\|.$$

A su vez, de la definición de componente ortogonal, habíamos visto que $\mathbf{b_i}^*$ es el menor vector de $\mathbf{b_i} + span(\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\})$, luego $\|\mathbf{b_i}^*\| \leq \|\mathbf{b_i}\|$ y se tiene

$$det(\mathcal{L}) = \prod_i \|\mathbf{b_i}^*\| \leq \prod_i \|\mathbf{b_i}\|.$$

Así, Hadamard propuso usar lo que se conoce como *radio de Hadamard* para establecer el nivel de ortogonalidad de una base $\mathbf{B}$ dada, que se define como

$$\mathcal{H} = \mathcal{H}(\mathbf{B}) = \left(\frac{\prod_i \|\mathbf{b_i}^*\|}{\prod_i \|\mathbf{b_i}\|} \right)^{\frac{1}{n}}.$$

Se tiene que $0 < \mathcal{H} \leq 1$, y $\mathbf{B}$ es ortogonal si, y solo si, $\mathcal{H} = 1$. Geométricamente se observa que entre más proximo a 1 sea el valor de este radio, más ortogonal es la base. Luego, diremos que $\mathbf{B}$ es casi ortogonal si $\mathcal{H}(\mathbf{B})$ es próximo a 1. El nivel de proximidad, está estrechamente vinculado a la capacidad de corrección de error de la función de descifrado.

Además, como $\prod_{i=1}^{n} \|\mathbf{b_i}\| = det(\mathcal{L})$, que es un invariante, encontrar una base casi ortogonal es equivalente a encontrar una base cuyos vectores sean lo más cortos posibles.

3.2.2. Elección de una mala base

En contrapartida, una mala base es aquella en la que el radio de Hadamard está próximo a 0, o una en la que la longitud de los vectores es grande. Para determinar esta base, como mencionamos anteriormente, es suficiente con establecer una matriz $\mathbf{U} \in M_{n\times n}(\mathbb{Z})^*$ tal que $\mathbf{H} = \mathbf{BU}$ cumple esta condición, donde $\mathbf{B}$ es la clave privada.

En [5] Micciancо propone usar como clave pública la *Forma Normal de Hermite* de la clave privada $\mathbf{B}$. Recordamos que esta matriz se define como sigue.

Definición 3.1. *Sea* $\mathbf{H} = (h_{ij}) \in \mathbb{Z}^{n\times n}$ *y rango máximo. Diremos que* $\mathbf{H}$ *se encuentra en la Forma Normal de Hermite si* $\mathbf{H}$ *es una matriz triangular cumpliendo que* $h_{ii} > 0$ *y* $|h_{ji}| < h_{ii}$ *para* $1 \leq i \leq n$ *y* $1 \leq j < i \leq n$.

Recordamos también que, para toda matriz $\mathbf{B}$, existe una única matriz entera $\mathbf{U} \in M_{n\times n}(\mathbb{Z})$ tal que $det(\mathbf{U}) = \pm 1$ y la matriz $\mathbf{H} = \mathbf{BU}$ está en la Forma Normal de Hermite. Por tanto, dada una base $\mathbf{B}$ existe una única matriz en la Forma Normal de Hermite $\mathbf{H}$ tal que $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{H})$. Además, se conocen algoritmos eficientes que llevan cualquier matriz a su Forma Normal de Hermite. Por ello, Micciano razona que la Forma Normal de Hermite es la peor base que se puede dar.

De esto también se deduce que usando la Forma Normal de Hermite se tiene un algoritmo eficiente para establecer si, dos matrices $\mathbf{B}$ y $\mathbf{C}$, definen el mismo retículo, es decir, si $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$. Para ello, basta con poner cada una de las matrices en su Forma Normal de Hermite y ver si estas coinciden.

3.2.3. Cifrado y descifrado

Sean $\mathbf{B}$ una base casi ortogonal, $\mathbf{H}$ una mala base del retículo $\mathcal{L}(\mathbf{B})$, y $\mathbf{x} \in \mathbb{Z}^n$ el mensaje original a transmitir. La función de cifrado $e_{\mathbf{H}} : \mathbb{Z}^n \longrightarrow \mathbb{R}^n$ se define como

$$e_{\mathbf{H}}(\mathbf{x}) = \mathbf{Hx} + \mathbf{r},$$

donde $\mathbf{r}$ es un vector error.

Por otra parte, la función de descifrado $d_{\mathbf{B}} : \mathcal{C} \longrightarrow \mathcal{P}$, hace uso de la clave privada $\mathbf{B}$ para retomar el mensaje original a partir de $\mathbf{c} = e_{\mathbf{H}}(\mathbf{x})$, siempre y cuando el error $\mathbf{r}$ no sea demasiado grande. Esta función de descifrado sigue el Algoritmo 1 conocido como el algoritmo de Babai o el algoritmo del *Nearest Plane*. Recibe este último nombre pues se basa en un proceso de búsqueda de planos cercanos al vector de $\mathbb{R}^n$ dado, siguiendo la misma línea de la demostración de la Proposición 1.15, tal y como explicamos a continuación.

Sea $\mathbf{B}$ una base, y sea $\mathbf{t} \in \mathbb{R}^{\mathbf{n}}$. La Proposición 1.15 nos dice que $\mathcal{C}(\mathbf{B}^*)$ es una región fundamental de $\mathcal{L} = \mathcal{L}(\mathbf{B})$. Es decir, existen $\mathbf{v} \in \mathcal{L}$ y $\mathbf{x} \in \mathcal{C}(\mathbf{B}^*)$ tales que $\mathbf{t} = \mathbf{v} + \mathbf{x}$. Además, la demostración de dicha proposición establece explícitamente quién es $\mathbf{v}$, el cual vimos que se obtiene particionando en primer lugar $\mathbb{R}^n$ en hiperplanos de la forma

$$\{c\mathbf{b_n} + span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}]) \mid c \in \mathbb{R}\},$$

y siguiendo luego el proceso de forma recursiva para $span([\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}])$ con $i \in \{2, \ldots, n-1\}$. El algoritmo del *Nearest Plane* sigue este mismo proceso recursivo lo que, en cada paso, se busca el hiperplano $l_k\mathbf{b_k} + span([\mathbf{b_1}, \ldots, \mathbf{b_{k-1}}])$ más cercano a $\mathbf{t} - \sum_{i=k+1}^{n} l_i\mathbf{b_i}$, donde $l_k \in \mathbb{Z}$. Veamos a continuación quiénes son explícitamente estos coeficientes $l_1, \ldots, l_n$.

Empezamos explicando cómo obtener l_n. Sea $L_n = dist(t, span(\mathbf{b_1}, \ldots, \mathbf{b_{n-1}})$. Geométricamente, se observa que L_n coincide con la longitud de la componente ortogonal de $\mathbf{t}$ respecto de $\{\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}\}$, es decir,

$$L_n = \|\mathbf{t} \perp \{\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}\}\| = \|\mathbf{t}\| \cdot cos(\alpha),$$

siendo α el ángulo que forma $\mathbf{t}$ con $\mathbf{t} \perp \{\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}\}$. Por otra parte, como $\mathbf{b_n}^*$, por definición, tiene la misma dirección que $\mathbf{t} \perp \{\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}\}$, el ángulo que forma, bien $\mathbf{b_n}^*$ o bien $-\mathbf{b_n}^*$, con $\mathbf{t}$ es también α. Luego, el producto escalar de $\mathbf{t}$ y $\mathbf{b_n}^*$ es

$$\langle \mathbf{t}, \mathbf{b_n}^* \rangle = \pm\|\mathbf{t}\|\|\mathbf{b_n^*}\| \cdot cos(\alpha),$$

tomando el signo positivo si $\mathbf{b_n}^*$ tiene el mismo sentido que $\mathbf{t} \perp \{\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}\}$, y menos si tiene sentido contrario. Con esto se tiene que, sustituyendo en la expresión anterior,

$$L_n = \pm\frac{\langle \mathbf{t}, \mathbf{b_n^*} \rangle}{\|\mathbf{b_n^*}\|}.$$

Por lo tanto, t es un vector del plano

$$\pm L_n \frac{\mathbf{b_n^*}}{\|\mathbf{b_n^*}\|} + span(\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}) = \frac{\langle \mathbf{t}, \mathbf{b_n^*} \rangle}{\|\mathbf{b_n^{*2}}\|}\mathbf{b_n^*} + span(\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}).$$

Recordando que, por definición de componente ortogonal, las capas

$$c\mathbf{b_n}^* + span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}]) = c\mathbf{b_n} + span([\mathbf{b_1}, \ldots, \mathbf{b_{n-1}}]),$$

coinciden para todo $c \in \mathbb{R}^n$, llegamos a que $\mathbf{t}$ vive en la capa $c_n = \frac{\langle \mathbf{t}, \mathbf{b_n^*} \rangle}{\|\mathbf{b_n^*}^2\|}$, donde c_n sigue la notación introducida en la demostración de la Proposición 1.15. Finalemnte, obtenemos nuestro $l_n = \lfloor c_n \rceil = \left\lfloor \frac{\langle \mathbf{t}, \mathbf{b_n^*} \rangle}{\|\mathbf{b_n^*}^2\|} \right\rceil$. El mismo proceso demuestra que esto se extiende para el resto de coeficientes, siendo $l_k = \left\lfloor \frac{\langle \mathbf{t}, \mathbf{b_k^*} \rangle}{\|\mathbf{b_k^*}^2\|} \right\rceil$. Además, tal y como se demostró en la Proposición 1.15, se tiene que

$$\mathbf{t} - \sum_{i=1}^{n} l_i \mathbf{b_i} \in \mathcal{C}(\mathbf{B}^*).$$

Este proceso recursivo puede implementarse en un algoritmo tal y como se demuestra a continuación, y es el conocido algoritmo del *Nearest Plane*.

Algorithm 1: Nearest Plane.

Input: $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$ y un punto seleccionado $\mathbf{t} \in \mathbb{R}^n$.
Output: Un punto del retículo $\mathbf{v} \in \mathcal{L}(\mathbf{B})$ tal que $\mathbf{t} - \mathbf{v} \in \mathcal{C}(\mathbf{B}^*)$.

NearestPlane($\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_k}]$, $\mathbf{t}$):
if $k = 0$ **then**
  return 0
else
  $\mathbf{b}^* \leftarrow$ **GramSchmidt**($\mathbf{B}$)
  $l_k \leftarrow \left\lfloor \frac{\langle \mathbf{t}, \mathbf{b_k^*} \rangle}{\|\mathbf{b_k^*}\|^2} \right\rceil$
  return $l_k \cdot \mathbf{b_k}$ + **NearestPlane**($\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_k}]$, $\mathbf{t} - l_k \cdot \mathbf{b_k}$)

En nuestro caso, si recibimos el texto cifrado $\mathbf{c} = \mathbf{Hx} + \mathbf{r} \in \mathbb{R}^n$, y le aplicamos el algoritmo *Nearest plane*, el resultado es un vector $\mathbf{v} \in \mathcal{L} = \mathcal{L}(\mathbf{H})$ tal que $\mathbf{c} - \mathbf{v} \in \mathcal{C}(\mathbf{B}^*)$. Al ser $\mathcal{C}(\mathbf{B}^*)$ una región fundamental, notamos que si el vector error $\mathbf{r}$ es un punto de $\mathcal{C}(\mathbf{B}^*)$, necesariamente debe ser $\mathbf{v} = \mathbf{Hx}$, y podemos obtener el mensaje original calculando $\mathbf{H}^{-1}\mathbf{v}$. Sin embargo, la región fundamental $\mathcal{C}(\mathbf{B}^*)$ no puede ser pública, ya que con ella cualquier persona puede descifrar el mensaje. Por ello, basta con indicar que el vector error debe ser tal que $\|\mathbf{r}\| \leq \frac{1}{2} min_i \|\mathbf{b_i}^*\|$. Esto justifica que una buena base es una lo más ortogonal posible, pues en caso contrario, al ser el determinante un invariante del retículo, si la base es poco ortogonal, su ortogonalización de Gram-Schmidt tendrá vectores muy pequeños, siendo menor su capacidad correctora.

3.3. Posible ataque: Algoritmo LLL

Por lo visto en la sección anterior, sabemos que conociendo una buena base del retículo es posible resolver el CVP con cierto margen de error usando el

Algoritmo 1. Además vimos que una buena base es aquella cuyos vectores son lo más pequeños posible. En esta sección estudiaremos el algoritmo LLL propuesto por Lenstra, Lenstra y Lovasz en [6] como posible ataque al criptosistema GGH y, por tanto, elemento fundamental en el criptoanálisis basado en retículo, pues es capaz de encontrar de manera eficiente una base cuyos vectores son de tamaño reducido. Estas bases se conocen como *bases reducidas*.

3.3.1. Bases reducidas

Para la obtención de una base reducida es fundamental considerar las proyecciones ortogonales de un vector $\mathbf{x}$ de $\mathbb{R}^n$ dado, respecto a una base $\mathbf{B} = \{\mathbf{b_1}, \ldots, \mathbf{b_n}\} \in M_{n\times n}(\mathbb{R})^*$. Esto es, definimos $\pi_i(\mathbf{x}) := \mathbf{x} \perp span(\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\})$, para todo $i = 1, \ldots, n$. Geométricamente, siguiendo un razonamiento análogo a la sección anterior, se puede comprobar que

$$\pi_i(\mathbf{x}) = \sum_{j=i}^{n} \frac{\langle \mathbf{x}, \mathbf{b_j^*} \rangle}{\langle \mathbf{b_j^*}, \mathbf{b_j^*} \rangle} \mathbf{b_j^*}. \tag{3.1}$$

En particular, se tiene que $\mathbf{b_i^*} = \pi_i(\mathbf{b_i})$.

Una vez aclarada la notación, nos encontramos en disposición de definir el concepto de *base LLL reducida*. Estas bases vienen dadas en función de un parámetro real $\frac{1}{4} < \delta < 1$, por lo que se habla de bases δ-LLL reducidas.

Definición 3.2. *Una base* $\mathbf{B} = \{\mathbf{b_1}, \ldots, \mathbf{b_n}\} \in \mathbb{R}^{d\times n}$ *es* δ*-LLL reducida si*

(i) $|\mu_{i,j}| \leq \frac{1}{2}$ *para todo* $i > j$,
(ii) para cualquier par de vectores consecutivos $\mathbf{b_i}, \mathbf{b_{i+1}}$, *se tiene*

$$\delta \|\pi_i(\mathbf{b_i})\|^2 \leq \|\pi_i(\mathbf{b_{i+1}})\|^2;$$

donde $\mu_{i,j}$ *son los dados por el método de ortogonalización de Gram-Schmidt visto en Proposición 1.10.*

Teniendo en cuenta (3.1), la segunda condición se puede reescribir al tener

$$\|\pi_i(\mathbf{b_{i+1}})\|^2 = \left\| \frac{\langle \mathbf{b_{i+1}}, \mathbf{b_i^*} \rangle}{\langle \mathbf{b_i^*}, \mathbf{b_i^*} \rangle} \mathbf{b_i^*} + \sum_{j=i+1}^{n} \frac{\langle \mathbf{b_{i+1}}, \mathbf{b_j^*} \rangle}{\langle \mathbf{b_j^*}, \mathbf{b_j^*} \rangle} \mathbf{b_j^*} \right\|^2 = \|\mu_{i+1,i}\mathbf{b_i}^* + \mathbf{b_{i+1}}^*\|^2 =$$

$$= \|\mu_{i+1,i}\mathbf{b_i}^*\|^2 + \|\mathbf{b_{i+1}}^*\|^2 = (\mu_{i+1,i})^2 \|\mathbf{b_i}^*\|^2 + \|\mathbf{b_{i+1}}^*\|^2.$$

Así, la segunda condición en la definición de base δ-LLL reducida es equvialente a

$$(\delta - \mu_{i+1,i}^2)\|\mathbf{b_i}^*\|^2 \leq \|\mathbf{b_{i+1}}^*\|^2. \tag{3.2}$$

De esto, se deduce el siguiente teorema que, en particular demuestra que obtener una base reducida permite resolver una aproximación del SVP.

Teorema 3.3. *Para cualquier* $\frac{1}{4} < \delta < 1$*, si* $\mathbf{B}$ *es una base* δ*-LLL reducida de un retículo* $\mathcal{L}$*, se tiene que*

$(i)\ \|\mathbf{b_1}\| \leq \alpha^{(n-1)/2}\lambda(\mathcal{L})$
$(ii)\|\mathbf{b_1}\| \leq \alpha^{(n-1)/4}det(\mathbf{B})^{1/n}$

donde $\alpha = \frac{1}{\delta - 1/4} \geq 4/3$.

Demostración. Para la demostración de (i), nótese que sea $\mathbf{B}$ una base δ-LLL reducida, se tiene que $\mu_{i,j}^2 \leq \frac{1}{4}$, para todo $i > j$ por la condición (i). Si definimos $\alpha := \frac{1}{\delta - \frac{1}{4}}$, de (3.2) se sigue que

$$\|\mathbf{b_i}^*\|^2 \leq \alpha\|\mathbf{b_{i+1}}^*\|^2.$$

Aplicando esta desigualdad repetidamente, obtenemos que

$$\|\mathbf{b_1}^*\|^2 \leq \alpha^{i-1}\|\mathbf{b_i}^*\|^2 \leq \alpha^{n-1}\|\mathbf{b_i}^*\|^2. \tag{3.3}$$

donde la última desigualdad se sigue pues $\alpha \geq 4/3 > 1$. Además, como se cumple para todo $i \in \{1, \ldots, n\}$, se tiene que el primer vector en una base LLL reducida satisface

$$\|\mathbf{b_1}\| \leq \alpha^{(n-1)/2} min_i \|\mathbf{b_i}^*\| \leq \alpha^{(n-1)/2}\lambda_1,$$

ya que $\lambda(\mathcal{L}) \geq min_i\|\mathbf{b_i}^*\|$ por el Teorema 2.4. Esto demuestra la primera propiedad del teorema.

Una vez demostrado (i), veamos que (ii) también es cierto. Para ello relacionamos la longitud del primer vector de la base LLL reducida con el determinante del retículo. Esto se deduce de aplicar (3.3) en el siguiente producto

$$\|\mathbf{b_1}\|^n \leq \prod_i \alpha^{(i-1)/2}\|\mathbf{b_i}^*\| = \prod_i \alpha^{(i-1)/2} \prod_i \|\mathbf{b_i}^*\| = \alpha^{n(n-1)/4}det(\mathbf{B}),$$

lo que implica que

$$\|\mathbf{b_1}\| \leq \alpha^{(n-1)/4}det(\mathbf{B})^{1/n}.$$

□

Asumiendo que podemos obtener una base δ-LLL reducida del retículo $\mathcal{L}$ con $\frac{1}{4} < \delta < 1$, del teorema anterior se sigue el siguiente corolario.

Corolario 3.4. *Sea* $\mathbf{B}$ *una base* δ*-LLL reducida de* $\mathcal{L}$*. El primer vector de* $\mathbf{B}$ *es una solución del problema* γ*-SVP, donde* $\gamma = \alpha^{(n-1)/2}$ *con* $\alpha = \frac{1}{\delta - 1/4} \geq 4/3$.

3.3.2. El Algoritmo LLL

El algoritmo LLL se divide en dos pasos, siendo el objetivo de cada uno de ellos lograr que se cumplan las dos propiedades de una base δ-LLL reducida.

El primer paso consiste en reducir el tamaño de la base hasta conseguir la condición (i). Esto se consigue haciendo un uso iterado del Algoritmo 1. En efecto, observamos que si $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$ es tal que $\mathbf{b_i} - \mathbf{b_i}^* \in \mathcal{C}(\mathbf{B}^*)$, para todo $i = 1, \ldots, n$, se tiene que $|\mu_{i,j}| \leq \frac{1}{2}$ para todo $j < i$, por definición de $\mathcal{C}(\mathbf{B}^*)$. De esta manera, el Algoritmo 2 se basa en comprobar si $\mathbf{b_i} - \mathbf{b_i}^* \in \mathcal{C}(\mathbf{B}^*)$, para en caso contrario redefinir $\mathbf{b_i}$, tomando el nuevo $\mathbf{b_i}$ como $\mathbf{b_i} - \mathbf{v_i}$, siendo $\mathbf{v_i} = \mathbf{NearestPlane}(\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}], \mathbf{b_i} - \mathbf{b_i}^*)$. Siguiendo este proceso obtenemos una base reducida, pues recordemos que en la sección anterior vimos que el Algoritmo 1 es tal que $\mathbf{b_i} - \mathbf{v_i} \in \mathcal{C}(\mathbf{B}^*)$.

Algorithm 2: Size Reduce

Input: $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$ una base del retículo.
Output: $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$ una base cuyos vectores cumplen la condición de longitud reducida.

SizeReduce($\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$):
for $i = 2$ *to* n **do**
  $\mathbf{v_i} \leftarrow \mathbf{NearestPlane}(\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}], \mathbf{b_i} - \mathbf{b_i}^*)$
  $\mathbf{b_i} \leftarrow \mathbf{b_i} - \mathbf{v_i}$
return $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$

Una vez reducida la base de entrada $\mathbf{B}$, la única forma de que no sea una base LLL redudica es no dándose la segunda condición, es decir, $\delta\|\pi_i(\mathbf{b_i})\|^2 > \|\pi_i(\mathbf{b_{i+1}})\|^2$ para algún índice $i \in \{1, \ldots, n\}$. Esto puede sudecer para más de un par $(\mathbf{b_i}, \mathbf{b_{i+1}})$. El Algoritmo 3, lo que hace es cambiar el orden de $\mathbf{b_i}$ y $\mathbf{b_{i+1}}$, y no va a importar el par por el que se comience a realizar estos cambios. El algoritmo LLL original, empieza por intercambiar el par para el cual $\mathbf{b_i}$ es el menor vector no ordenado, aunque cualquier otra selección sería adecuada. De hecho, se podría optar por intercambiar varios pares a la vez, obteniendo así una versión paralela del algoritmo LLL. Después del cambio de orden en los pares que incumplen la segunda condición, la base resultante no necesariamente sigue siendo de longitud reducida, por lo que se debe repetir el Algoritmo 2 y el proceso de intercambio desde el principio. Así se define el Algoritmo 3 que, de terminar, devuelve una base LLL reducida de $\mathcal{L}$.

Resta ver que efectivamente el proceso finaliza. Para ello, estudiaremos el número máximo de cambios de orden que se pueden dar, lo que nos lleva a asociar a la base de entrada $\mathbf{B}$ un entero positivo y comprobar que cada vez que se produce un cambio de orden entre dos vectores, este entero decrece según un factor constante.

Algorithm 3: Algoritmo LLL.

Input: $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$ una base del retículo y un parámetro real δ.
Output: $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$ una base LLL-reducida del retículo.

LLL($\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$, δ):
SizeReduce($\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$)
if $\delta\|\pi_i(\mathbf{b_i})\|^2 > \|\pi_i(\mathbf{b_{i+1}})\|^2$ *para algún* i **then**
  swap($\mathbf{b_i}, \mathbf{b_{i+1}}$);
  return LLL($\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$, δ)
else
  return $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}]$

Sea $\mathbf{B} = [\mathbf{b_1}, \ldots, \mathbf{b_n}] \in M_{n\times n}(\mathbb{Z})$, definimos dicho entero como

$$\mathcal{D} = \prod_{k=1}^{n} det(\mathcal{L}([\mathbf{b_1}, \ldots, \mathbf{b_k}]))^2 \in \mathbb{Z}.$$

Queremos demostrar que $\mathcal{D}$ decrece según un factor δ en cada iteración. Para ello recordamos que

$$det(\mathcal{L}([\mathbf{b_1}, \ldots, \mathbf{b_k}]))^2 = \prod_{i=1}^{k} \|\mathbf{b_i}^*\|^2. \tag{3.4}$$

Veamos primero que, dada una base $\mathbf{B}$, el Algoritmo 2 no varía el valor de este entero. Esto se sigue del hecho de que el Algoritmo 2 no afecta a la matriz ortogonal $\mathbf{B}^*$ y de la ecuación (3.4). En efecto, $\mathbf{B}^*$ es también la correspondiente ortogonalización de Gram-Schmidt de $\mathbf{B}'$ =**SizeReduce**($\mathbf{B}$). Esto se sigue de la definición de $\mathbf{B}^*$, ya que $\mathbf{b_i}^* \perp span(\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\})$ para todo $i \in \{1, \ldots, n\}$, y, por tanto, $\mathbf{t} = \mathbf{b_i} - \mathbf{b_i}^* \in span(\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\})$. Luego, al hallar $\mathbf{v_i} = \mathbf{NearestPlane}(\mathbf{B}, \mathbf{b_i} - \mathbf{b_i}^*) \in span(\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\})$, se tiene que en el paso $k \geq i$, el hiperplano más cercano a $\mathbf{t}$ lo determina $\mathbf{c}_k = 0$, pues $\mathbf{t}$ pertenece a un subretículo de dimensión menor a k. Es así que, para poder determinar un hiperplano $c_k\mathbf{b_k} + span(\{\mathbf{b_1}, \ldots, \mathbf{b_{k-1}}\}) \neq span(\{\mathbf{b_1}, \ldots, \mathbf{b_{k-1}}\})$ que se encuentre cercano a $\mathbf{t}$, debemos encontrarnos en un paso $k < i$. Teniendo esto en cuenta, podemos afirmar que $\mathbf{b_i}' = \mathbf{b_i} - \mathbf{v_i} \in span(\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}\})$, lo que implica que $(\mathbf{b_i}')^* = \mathbf{b_i}^*$, demostrando así que $\mathbf{B}$ y $\mathbf{B}'$ tienen la misma ortogonalización de Gram-Schmidt.

Llegados a este punto, nos queda determinar qué efecto tiene en $\mathcal{D}$ cada uno de los cambios en el segundo paso del Algoritmo 3. Centrándonos en el efecto que tiene el cambio de $\mathbf{b_i}$ por $\mathbf{b_{i+1}}$ para algún $i \in \{1, \ldots, n\}$, denotamos por $\mathcal{D}'$ al entero asociado a la base $\{\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}, \mathbf{b_{i+1}}, \mathbf{b_i}, \ldots, \mathbf{b_n}\}$. Nótese que para $j \neq i$ el retículo $\mathcal{L}([\mathbf{b_1}, \ldots, \mathbf{b_j}])$ no es afectado por los cambios. Esto se puede comprobar considerando los casos $j < i$ y $j > i$. Cuando $j < i$, no hay cambios en la base $\{\mathbf{b_1}, \ldots, \mathbf{b_j}\}$ por lo que el valor de $det(\mathcal{L}([\mathbf{b_1}, \ldots, \mathbf{b_j}]))$ se mantiene.

Por otro lado, si $j > i$, el único cambio que se produce es que dos vectores en $\{\mathbf{b_1}, \ldots, \mathbf{b_j}\}$ han cambiado de orden, por lo que el retículo $\mathcal{L}([\mathbf{b_1}, \ldots, \mathbf{b_j}])$ no varía y el valor de $det(\mathcal{L}([\mathbf{b_1}, \ldots, \mathbf{b_j}]))$ sigue siendo el mismo. Por tanto, el único factor en $\mathcal{D}$ que se ve afectado es el correspondiente al retículo $\mathcal{L}([\mathbf{b_1}, \ldots, \mathbf{b_i}])$, pues $\mathbf{b_i}$ era previamente $\mathbf{b_{i+1}}$. Se tiene que

$$\frac{\mathcal{D}}{\mathcal{D}'} = \frac{det(\mathcal{L}([\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}, \mathbf{b_i}]))^2}{det(\mathcal{L}([\mathbf{b_1}, \ldots, \mathbf{b_{i-1}}, \mathbf{b_{i+1}}]))^2}.$$

De esta forma, aplicando (3.4) y la definición de proyección ortogonal, se sigue que

$$\frac{\mathcal{D}}{\mathcal{D}'} = \frac{\prod_{j=1}^{i} \|\pi_j(\mathbf{b_j}^*)\|^2}{(\prod_{j=1}^{i-1} \|\pi_j(\mathbf{b_j}^*)\|^2) \cdot \|\pi_i(\mathbf{b_{i+1}}^*)\|^2} = \frac{\|\pi_i(\mathbf{b_i}^*)\|^2}{\|\pi_i(\mathbf{b_{i+1}}^*)\|^2} \geq \frac{1}{\delta},$$

pues los cambios de orden sólo se efectúan cuando $\delta\|\pi_i(\mathbf{b_i}^*)\|^2 > \|\pi_i(\mathbf{b_{i+1}}^*)\|^2$.

Tenemos así que

$$\mathcal{D}' \leq \delta\mathcal{D}$$

y tras m iteraciones,

$$\mathcal{D}^{(m)} \leq \delta^m\mathcal{D},$$

donde $\mathcal{D}$ es el valor asignado a la base inicial y $\mathcal{D}^{(m)}$ el valor tras m iteraciones.

Además, como $\mathcal{D}$ es un entero positivo, $\mathcal{D} \geq 1$ y $(\frac{1}{\delta})^m \leq \mathcal{D}$, llegando así a que

$$m \leq log_{\frac{1}{\delta}}\mathcal{D}.$$

Esto prueba la existencia de una cota superior del número de iteraciones en función del valor inicial $\mathcal{D}$. Es más, en [6] se demuestra que el número de iteraciones es polinomial en n.

3.3.3. Aplicaciones del Algoritmo LLL

Como primera aplicación del Algoritmo LLL, veamos que es capaz de resolver el problema de Teoría de Números dado por el Teorema 2.13, es decir, el algoritmo demuestra que es posible escribir todo número primo $p \equiv 1 \ (mod\ 4)$ como $p = a^2 + b^2$ con $a, b \in \mathbb{Z}$. Recordamos de la demostración de dicho teorema que existe un retículo entero $\mathcal{L}$ generado por la base

$$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ n & p \end{bmatrix},$$

tal que todo vector $\mathbf{v} \in \mathcal{L}(\mathbf{B})$ es tal que $p \mid \|\mathbf{v}\|^2$. Además, por el Teorema de Minkowski 2.9 sabemos que existe $\mathbf{v} \in \mathcal{L}(\mathbf{B}) \setminus \{\mathbf{0}\}$ con $\|\mathbf{v}\|^2 \leq 2p$ y de esto se sigue que $\|\mathbf{v}\|^2 = a^2 + b^2 = p$. La cuestión ahora es determinar este vector, pero esto es posible aplicando el Algoritmo 3 a la base $\mathbf{B}$ para obtener una base

δ-reducida $\{\mathbf{b_1}, \mathbf{b_2}\}$ del mismo retículo, donde $\delta = \frac{3}{4}$. Del Teorema 3.3 se tiene que $\|\mathbf{b_1}\| \leq \alpha^{1/4} det(\mathbf{B})^{1/2}$. Luego, elevando al cuadrado y teniendo en cuenta que $det(\mathbf{B}) = p$, se obtiene que $\|\mathbf{b_1}\|^2 \leq \sqrt{2}p < 2p$, por lo que $\mathbf{b_1}$ es el vector que buscamos.

Por otro lado, el Algoritmo LLL es capaz de resolver de forma aproximada el CVP y es por ello que se considera como un posible ataque al Criptosistema GGH. Para la resolución aproximada del CVP, consideremos como punto seleccionado al texto cifrado $\mathbf{c}$ y la base $\mathbf{B}$ del retículo. Luego, es posible hallar una base LLL-reducida del retículo mediante el Algoritmo 3, la cual va a ser una buena base del retículo, por lo que podemos aplicar $\mathbf{NearestPlane}(\mathbf{LLL}(\mathbf{B}), \mathbf{c})$ y obtener una solución aproximada del CVP. Esto implica que es posible obtener con cierto error el mensaje original que se quería transmitir en el Criptosistema GGH. Para aumentar los niveles seguridad, basta con tomar n suficientemente grande, pues el error que corrige es de carácter exponencial.

Bibliografía

[1] L. Adleman, R. Rivest, A. Shamir, *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems*, Communications of the ACM, vol. 21 (2), (1978), pp. 120–126.

[2] W. Diffie and M. Hellman, *New directions in cryptography*, IEEE Transactions on Information Theory, vol. 22, no. 6, (1976), pp. 644–654.

[3] P. van Emde Boas, *Another NP-complete problem and the complexity of com- puting short vectors in a lattice*, University of Amsterdam, Department of Mathematics, Netherlands, (1981).

[4] O. Goldreich, S. Goldwasser and S. Halevi, *Public-key cryptosystems from lattice reduction problems*, Advances in cryptology, volume 1294 of Lecture Notes in Comput. Sci., Springer, (1997), pp. 112–131.

[5] D. Micciancio, *Improving lattice based cryptosystems using the hermite normal form*, in Cryptography and Lattices Conference, volume 2146 of Lecture Notes in Computer Science, (2001), pp. 126–145. Springer-Verlag, Providence, Rhode Island.

[6] A. K. Lenstra, H. W. Lenstra and L. Lovasz, *Factoring Polynomials with Rational Coefficients*, Math. Ann. 261, (1982), pp. 515–534.

[7] H. Minkowski, *Geometrie der Zahlen*, Teubner, Leipzig, (1896). [Reprinted Chelsea, New York, (1953)].

[8] P.W. Shor, *Algorithms for quantum computation: discrete logarithms and factoring*, Proceedings 35th Annual Symposium on Foundations of Computer Science, IEEE Comput. Soc. Press, (1994), pp. 124–134.

[9] B. Peng, *The Determinant: a Means to Calculate Volume*, (2007). Disponible en: https://www.math.uchicago.edu/~may/VIGRE/VIGRE2007/REUPapers/FINALAPP/Peng.pdf

Lattices structure and their aplication to cryptography

Sección de Matemáticas
Universidad de La Laguna

Manuel Betancort Pérez
Facultad de Ciencias • Sección de Matemáticas
Universidad de La Laguna
alu0101158501@ull.edu.es

Abstract

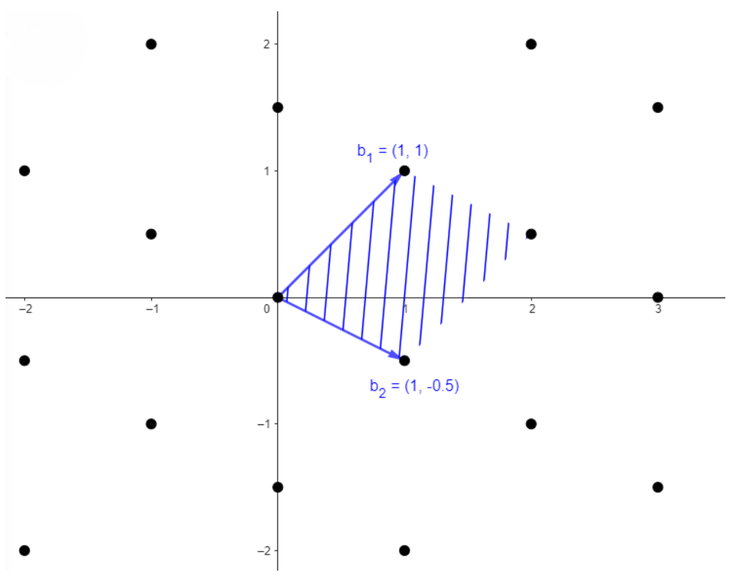
In this memory, we make an introduction to lattice theory. We present important results about them and study problems which are computationally hard to solve. We highlight the so called SVP and CVP, which are fundamental in lattice cryptography. Finally, we propose a simple cryptography system based in the aforementioned problems.

1. Introduction

We say that $\mathcal{L}$ is a lattice in $\mathbb{R}^n$ if there exists a matrix $\mathbf{B} \in M_{n\times d}(\mathbb{R})$ of rank $d \leq n$ such that

$$\mathcal{L} = \mathcal{L}(\mathbf{B}) := \mathbf{B}\mathbb{Z}^d = \{\mathbf{Bx} : \mathbf{x} \in \mathbb{Z}^d\} \subseteq \mathbb{R}^n.$$

For example, the black points in the next figure represent the lattice $\mathcal{L}$ defined by the matrix $\mathbf{B} = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -0.5 \end{smallmatrix}\right)$.



In this context we say that $\mathbf{B}$ is a basis of the lattice $\mathcal{L}$. Also, in this work we only consider lattices of full dimension, that is, when $d = n$.

A lattice $\mathcal{L}$ may have different basis and they all relate in the following way.

Theorem. Let **B**, **C** $\in M_{n\times n}(\mathbb{R})^*$, then $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$ if and only if there exists an invertible matrix $\mathbf{U} \in M_{n\times n}(\mathbb{Z})$ such that **B** $=$ **CU**.

2. Fundamental problems

We have two fundamental problems based on lattices known as the *Shortest vector problem* or SVP, and the *Closest vector problem* or CVP. Both problems are NP-complete as shown in [1], and they can be stated as follows. Given a lattice $\mathcal{L} \subseteq \mathbb{R}^n$:

- The 'Shortest vector problem' consists of finding the shortest non zero lattice point.
- The 'Closest vector problem' consists of, given a target point $\mathbf{t} \in \mathbb{R}^n$, finding the lattice point closest to $\mathbf{t}$.

3. Lattice cryptosistem

Criptography is the art of encrypting and decrypting data so that they are unintelligible to a third party. Since the advent of the Internet, public key cryptosystems have become essencial in Cryptography Theory. Moreover, with the introduction of quantum computers, there has been an increasing interest in *post-quantum cryptography*. In this regard, lattices seem to be a sensible object to consider when building cryptosystems in the post-quantum era, as approved by the famous NIST project.

The GGH cryptosistem, proposed by Goldreich, Goldwasser and Halevi in [2], is based in the complexity of solving the Closest Vector Problem. At a high level, the GGH cryptosistem works as follows:

- The private key is a 'good' basis $\mathbf{B} \in M_{n\times n}(\mathbb{Z})$ such that $\mathcal{L} = (\mathbf{B})$. Algorithmically, good bases allow to efficiently solve certain instances of the Closest Vector Problem.
- The public key $\mathbf{H} \in M_{n\times n}(\mathbb{Z})$ is a 'bad' basis of $\mathcal{L}$, i.e., $\mathcal{L}(\mathbf{H}) = \mathcal{L}(\mathbf{B})$. This basis can be computed by taking an integer matrix $\mathbf{U} \in M_{n\times n}(\mathbb{Z})^*$ such that $\mathbf{H} = \mathbf{UB}$. In [4], Micciancio proposed to use, as the public basis, the Hermite Normal Form of $\mathbf{B}$.
- In this way, let $\mathbf{x} \in \mathbb{Z}^n$ be the message to be sent, we define $\mathbf{c} = \mathbf{Hx} + \mathbf{r}$ as the ciphertext. The *error vector* $\mathbf{r} \in \mathbb{R}^n$ must be not too long and is taken randomly.
- The decryption problem corresponds to finding the lattice point $\mathbf{v}$ closest to the target ciphertext $\mathbf{c}$. We can solve this problem using the good basis $\mathbf{B}$. Then, if we have $\mathbf{v}$, we calculate $\mathbf{H}^{-1}\mathbf{v}$ and the original message is recovered.

A sense of 'good' basis is one whose vectors are as small as possible. In this direction, a possible attack to this cryptosistem is the famously known LLL Algorithm proposed by Lenstra, Lenstra and Lovasz, which was introduced in [3].

References

[1] P. van Emde Boas, *Another NP-complete problem and the complexity of com- puting short vectors in a lattice*, University of Amsterdam, Department of Mathematics, Netherlands, (1981).

[2] O. Goldreich, S. Goldwasser and S. Halevi, *Public-key cryptosystems from lattice reduction problems*, Advances in cryptology, volume 1294 of Lecture Notes in Comput. Sci., Springer, (1997), pp. 112–131.

[3] A. K. Lenstra, H. W. Lenstra and L. Lovasz, *Factoring Polynomials with Rational Coefficients*, Math. Ann. 261, (1982), pp. 515–534.

[4] D. Micciancio, *Improving lattice based cryptosystems using the hermite normal form*, in Cryptography and Lattices Conference, volume 2146 of Lecture Notes in Computer Science, (2001), pp. 126–145. Springer-Verlag, Providence, Rhode Island.