CRIPTOGRAFÍA POSTCUÁNTICA

TRABAJO DE FIN DE GRADO

Curso 2019/20

FACULTAD DE CIENCIAS MATEMÁTICAS

GRADO EN MATEMÁTICAS

Nombre del estudiante: Alba Gurpegui Ramón

Nombre del tutor/es: David Pérez García

Madrid, 8 de Julio de 2020

Índice general

Summary

In 1979 the RSA public-key cryptosystem was developed, an algorithm whose security depends on the integer factorization problem and whose importance cannot be underestimated: all the cybersecurity nowadays relies mostly on it.

Nevertheless, due to the emergence of quantum computing, the recent advances made by Google in the process of building a quantum computer, and for this case in particular, Shor's algorithm, which is capable of decomposing efficiently a number into its primer factors, it can be said that current cybersecurity is under real threat.

As a consequence, NIST (National Institute of Standards and Technology) requested public collaboration to find an alternative cryptosystem that remains secure to quantum attacks. 82 proposals from researchers from all around the globe were initially presented. After the first filtering of these initial approaches, in 2019 the list of the 26 candidate systems that passed to the second round was published. Among the selected, 14 of them are based on the Learning With Errors problem (LWE), which was developed by Oded Regev in 2009.

The objective of this work will be the description and security analysis of the simplest cryptosystem based on the LWE problem. In particular, we will show that the LWE problem is as difficult as the problem of finding the shortest vector in an n dimensional lattice, known to be NP-hard.

Capítulo 1

Introducción

Los sistemas criptográficos de clave pública más utilizados en la actualidad están basados en el RSA, desarrollado en 1979, y cuya dificultad reside en el problema de factorización de los números enteros.

Sin embargo, utilizando el algoritmo cuántico de Shor (1994), que descompone en factores un número N en tiempo polinómico en el número de bits de N, un ordenador cuántico podría romper el RSA, comprometiendo así la privacidad de las comunicaciones.

Los recientes avances en el proceso de construcción de un ordenador cuántico por parte de Google [**8**] demuestran que las consecuencias que esto trae consigo no tienen por qué ser cosa de un futuro lejano, al contrario. Por ello, el NIST (la agencia nacional de estándares de EEUU) ha lanzado una competición, solicitando la colaboración pública para conseguir afrontar la amenaza a nuestra ciberseguridad actual por parte de los ordenadores cuánticos, buscando cambiar los estándares actuales de criptografía por un criptosistema que sea resistente a eventuales ataques cuánticos. En consecuencia, investigadores de todo el mundo buscan a día de hoy algoritmos que sean resistentes a la computación cuántica, dando paso a una rama de la criptografía que se denomina criptografía post-cuántica [**6**]. Dentro de esta rama de la criptografía se encuentra la criptografía basada en retículos, en la que nos centraremos.

En 2017 el NIST comenzó el proceso de estandarización, en el que inicialmente se presentaron 82 propuestas de las cuales se aceptaron 69. En enero de 2019, tras eliminar errores y proponer numerosos ataques, se hizo pública la lista de criptosistemas candidatos que pasaron a la segunda ronda [**7**], tras eliminar algunas de las propuestas, o juntarlas en base a su gran similitud. En esta segunda ronda se busca mejorar el rendimiento de los algoritmos y se espera que tenga una duración de 12-18 meses, hasta el comienzo de una tercera fase. Hasta el 2022-2024 no se esperan conseguir los primeros borradores del proceso.

A día de hoy, en la ronda 2, de los 26 seleccionados, 14 de ellos están basados en el problema de Learning With Errors (LWE), cimentado en el trabajo pionero desarrollado por Oded Regev en 2009 [**2**]. Además de lo anterior, el problema LWE es también la primitiva en la que se han basado algunos de los mayores hitos más recientes en criptografía, por ejemplo, la existencia de "fully homomorphic encryption" debida a Gentry en 2009 [**4**], que garantiza la posibilidad de realizar cualquier computación sobre datos cifrados (sin conocer la clave secreta) y que permite, por

tanto, tener computación privada en la nube.

El objetivo de este TFG será describir el criptosistema más sencillo basado en LWE y demostrar su seguridad.

Introduciremos primero el problema de LWE, describiremos su criptosistema asociado y demostraremos su corrección (Capítulo 2). La mayor parte del trabajo irá enfocada a mostrar la seguridad (Capítulo 3), demostrando de forma rigurosa que LWE es al menos tan difícil como encontrar el vector más corto en una retícula n-dimensional, que es un conocido problema NP-duro. Para ello necesitaremos numerosos conceptos y resultados sobre retículas y las distribuciones de probabilidad definidas sobre ellas, así como nociones y resultados de computación cuántica, que iremos presentando según vayan siendo necesarios.

Capítulo 2

LWE y su sistema de cifrado asociado

Supongamos un sistema de n ecuaciones con n incógnitas en $\mathbb{Z}_p$, por ejemplo:

$$\begin{aligned} 9s_1 + 5s_2 + 9s_3 + 6s_4 &= 9 \quad \text{mód } 17 \\ 3s_1 + 6s_2 + 4s_3 + 5s_4 &= 7 \quad \text{mód } 17 \\ s_1 + s_2 + s_3 + s_4 &= 2 \quad \text{mód } 17 \\ 2s_1 + s_2 + s_3 + 5s_4 &= 3 \quad \text{mód } 17 \end{aligned}$$

Si añadimos una cierta cantidad de ruido al vector de términos independientes, ¿se podrá seguir recuperando la solución $(s_1, \ldots, s_4)$?

Este es exactamente el problema LWE (Learning With Errors) introducido por Oded Regev en [**2**]. Formalmente se puede definir de la siguiente manera.

Definición 1. Dado un parámetro $n \in \mathbb{N}$, dos enteros $p, m \geq 2$ de tamaño polinomial en n y una distribución de probabilidad $\chi : \mathbb{Z}_p \to \mathbb{R}^+$, definimos la variable aleatoria

$$A_{s,\chi} = (\mathbf{A}, \mathbf{A} \cdot \mathbf{s} + \mathbf{e}) = (\mathbf{a}_1, \langle \mathbf{a}_1, \mathbf{s} \rangle + \mathbf{e}_1), ..., (\mathbf{a}_m, \langle \mathbf{a}_m, \mathbf{s} \rangle + \mathbf{e}_m),$$

donde para cada $i \in \{1, ..., m\}$, $\mathbf{a}_i \leftarrow \mathbb{Z}_p^n, e_i \leftarrow \chi$, y A es la matriz de filas $\mathbf{a}_1, ..., \mathbf{a}_m$.

En la definición hemos utilizado como es usual la notación $e_i \in \mathbb{Z}_p \leftarrow \chi$ para referirnos a que e_i se muestrea según la distribución χ. Si no se especifica χ, como en el caso de los $\mathbf{a}_i$ se sobreentiende que se muestrea con la distribución uniforme.

Se puede definir análogamente $A_{s,\phi}$ donde ϕ es una distribución de probabilidad continua en el toro $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. En este caso, los m vectores son variables aleatorias independientes de la forma

$$\left(\mathbf{a}, \frac{\langle \mathbf{a}, \mathbf{s} \rangle}{p} + \mathbf{e} \quad \text{mód } 1 \right)$$

donde a se muestrea uniformemente y e se muestrea según ϕ.

Definición 2. Diremos entonces que, un algoritmo resuelve el problema $\text{LWE}_{p,\chi}$ si $\forall \mathbf{s} \in \mathbb{Z}_p^n$ dado un número arbitrario de muestras tomadas independientemente de $A_{s,\chi}$, el algoritmo devuelve $\mathbf{s}$. Se define análogamente $\text{LWE}_{p,\phi}$

Observación 3. : Resulta sencillo comprobar que, de ser el error inexistente en nuestra instancia del problema, su resolución sería trivial sin más que usar la eliminación Gaussiana.

Como se ha comentado en la introducción, la importancia del problema LWE se debe a su aplicación a la criptografía, ya que es una de las propuestas más prometedoras para pasar a ser el nuevo estándar de criptografía de clave pública

en el concurso que tiene abierto el NIST para sustituir el RSA por protocolos resistentes a ataques con ordenadores cuánticos.

Vamos mostrar a continuación en qué consiste esta propuesta:

Parámetros: Sean $n, m, p \in \mathbb{Z}$, tomaremos

$$p \geq 2, n^2 < p < 2n^2, m = (1+\varepsilon)(n+1)\log p \tag{1}$$

para $\varepsilon > 0$ y $\chi = \overline{\Psi}_{\alpha(n)}$ con

$$\alpha(n) = \frac{1}{\sqrt{n}\log^2 n} \tag{2}$$

donde $\overline{\Psi}_\alpha$ es la discretización de la distribución normal periódica de parámetro α (ver Sección 2.3 más abajo para una definición precisa).

A partir de ahora supondremos n, m, p, α como en (1), (2).

Clave privada: Una muestra aleatoria uniforme en $\mathbb{Z}_p^n$. Es decir, $\mathbf{s} \leftarrow \mathbb{Z}_p^n$.

Clave pública: Una muestra aleatoria de $A_{s,\chi}$. Es decir, $(\mathbf{a}_i, b_i)_{i=1}^m$, donde $\mathbf{a}_i$ se elige uniformemente al azar en $\mathbb{Z}_p^n$, e_i se muestrea de χ y $b_i = \langle \mathbf{a}_i, \mathbf{s}\rangle + e_i$.

Encriptado: para encriptar un bit escogeremos uniformemente al azar un conjunto auxiliar S de entre los 2^m subconjuntos posibles de $\{1, \dots m\}$, para el cual si el bit es 0, el encriptado es $\left(\sum_{i\in S} \mathbf{a}_i, \sum_{i\in S} b_i\right)$ y, si es 1, el encriptado es $\left(\sum_{i\in S} \mathbf{a}_i, \left\lfloor \frac{p}{2} \right\rfloor + \sum_{i\in S} b_i\right)$.

Desencriptado del par $(\mathbf{a}, b)$ será 0 si $b - \langle \mathbf{a}, s\rangle$ se encuentra más cercano a 0 que a $\left\lfloor \frac{p}{2} \right\rfloor$ módulo p. En caso contrario será 1.

Observación 4. Bajo la propuesta de parámetros, el tamaño de la clave pública es $O(mn\log p)$ y el proceso de cifrado incrementa el tamaño de un mensaje por un factor $O(n\log n)$.

Una vez introducido el criptosistema, necesitamos ver que es *correcto* (en el sentido de que si se encripta y luego se desencripta, el mensaje inicial no se altera) y que es *seguro* (no se puede obtener información del bit encriptado si no se conoce la clave secreta).

1. Corrección

Para $k \in \mathbb{Z}_p$ definimos $\chi^{\star k}$ como la distribución resultante de sumar (módulo p) k muestras independientes de χ con el convenio de que $\chi^{\star 0}$ es la distribución que tiene toda la masa en el 0.

Para ver la corrección necesitamos el siguiente:

Lema 5. *Sea $\delta > 0$, supongamos que $\forall k \in \{0, 1, ..., m\}, \chi^{\star k}$ cumple que*

$$\Pr_{e\sim\chi^{\star k}}\left\{|e| < \left\lfloor \frac{p}{2} \right\rfloor /2\right\} > 1 - \delta$$

entonces la probabilidad de error en el desencriptado es, como mucho, δ.

DEMOSTRACIÓN. Supongamos que estamos en el caso de tener que encriptar el 0, (para el 1 la demostración es análoga), con

$$(\mathbf{a}, b) = \left(\mathbf{a} = \sum_{i \in S} \mathbf{a}_i, b = \sum_{i \in S} b_i = \sum_{i \in S} \langle \mathbf{a}_i, \mathbf{s} \rangle + \sum_{i \in S} e_i \right)$$

Entonces, como despejando $b - \langle \mathbf{a}, \mathbf{s} \rangle = \sum_{i \in S} e_i$, la distribución del error es $\chi^{\star |S|}$, y, por hipótesis, se tiene que $\left| \sum_{i \in S} e_i \right| < \left\lfloor \frac{p}{2} \right\rfloor /2$ con probabilidad, como poco, $1 - \delta$. Por encontrarnos en el caso en que es más cercano a 0 que a $\left\lfloor \frac{p}{2} \right\rfloor /2$ el desencriptado es correcto con probabilidad al menos $1 - \delta$. □

Utilizaremos también la siguiente definición estándar:

DEFINICIÓN 6 (Función despreciable). Una función $f(n)$ es despreciable en n si para todo $c > 0$,

$$\lim_{n \to \infty} n^c f(n) = 0.$$

Es decir, si converge a 0 cuando $n \to \infty$ más rápido que cualquier polinomio.

LEMA 7. *Dada la selección de parámetros establecida, se cumple para todo $k \in \{0, 1, ..., m\}$ y alguna función despreciable $\delta(n)$, que*

$$\Pr_{e \sim \bar{\Psi}_\alpha^{\star k}} \left\{ |e| < \left\lfloor \frac{p}{2} \right\rfloor /2 \right\} > 1 - \delta(n)$$

Postpondremos la demostración de este lema al momento en que definamos $\bar{\Psi}_\alpha$ (Definición 19).

Como consecuencia de ambos lemas, para cada bit $c \in \{0, 1\}$, si empleamos el protocolo anterior para elegir la clave pública y la clave privada para encriptar y desencriptar c, el resultado será c con probabilidad $1 - \delta(n)$ para una función despreciable δn.

2. Seguridad

Probar la seguridad del criptosistema anterior será el principal objetivo de este TFG. El objetivo es demostrar que romper el criptosistema implicaría la existencia de un algoritmo cuántico capaz de encontrar en tiempo polinomial el vector más corto en una retícula, algo que se cree imposible [1].

Para formalizar esta afirmación necesitamos las siguientes definiciones básicas sobre retículas

2.1. Retículas I.

DEFINICIÓN 8 (Retícula). Dados n vectores linealmente independientes $b_1, \cdots, b_n \in \mathbb{R}^m$, la retícula generada por ellos se define como:

$$\Lambda(b_1, b_2, ..., b_n) = \left\{ \sum x_i b_i \mid x_i \in \mathbb{Z} \right\}.$$

Identificaremos muchas veces la retícula con la matriz que tiene los vectores b_i por columnas. Así hablaremos de $\det(\Lambda)$, Λ^{-1}, ...

DEFINICIÓN 9 (Retícula dual). Definimos la retícula dual de Λ como la generada por el conjunto de vectores b_i^* tales que $\langle b_i, b_j^* \rangle = \delta_{i,j}$. La denotamos por Λ^*.

[1] Nótese que ese problema es NP-duro [**1**]

Es trivial comprobar la siguiente identidad (que también caracteriza a Λ^*).

$$\Lambda^* = \{y \in \mathbb{R}^n \mid \forall x \in \Lambda, \langle x, y\rangle \in \mathbb{Z}\}.$$

Nótese que, por la defínción de Λ^*, se tiene que $\Lambda^{-1} = (\Lambda^*)^T$ y, por tanto, $\det(\Lambda^*) = 1/\det(\Lambda)$.

DEFINICIÓN 10 (Longitud del vector no nulo más corto de una retícula λ_1). Se trata del más pequeño r tal que los puntos de una retícula contenidos en una bola de radio r generan un espacio de dimension 1. Podemos generalizar esto como sigue

$$\lambda_i(\Lambda) = \text{ínf}\left\{r \mid \dim(gen(\Lambda \cap \overline{\mathbf{B}}(0,r))) \geq i\right\}$$

con

$$\overline{\mathbf{B}}(0,r) = \{x \in \mathbb{R}^m \mid \|x\| \leq r\}$$

la bola cerrada de centro 0 y radio r.

Podemos ya definir el problema GapSVP_γ, que no es más que el problema de encontrar el vector más corto (SVP son las siglas de *Shortest Vector Problem*) con un error relativo γ.

DEFINICIÓN 11 (GapSVP_γ). Dada un retícula n-dimensional y un número $d > 0$ se devolverá SI cuando $\lambda_1(\Lambda) \leq d$ y NO cuando $\lambda_1(\Lambda) > d \cdot \gamma$

Esto nos permite enunciar ya el Teorema principal del trabajo.

TEOREMA 12. *Si existe un algoritmo W que, en tiempo polinómico en n, distingue entre los encriptados de 0 y 1 con probabilidad no despreciable y para una fracción no despreciable de* $\mathbf{s}$*, entonces existe un algoritmo cuántico que, en tiempo polinómico en n, resuelve el problema* $\text{GapSVP}_{O(n/\alpha(n))}$.

2.2. Esquema de la demostración del Teorema 12. No es difícil ver [**2**], que romper el criptosistema implica resolver el problema asociado de LWE. Por tanto, nuestro objetivo será ver cómo disponer de un algoritmo que resuelva el problema de LWE implica la existencia de un algoritmo cuántico para resolver $GapSVP_{O(n/\alpha(n))}$.

El principal resultado para probar el Teorema 12 será ver que resolver LWE (con ciertos parámetros) implica la existencia de un algoritmo cuántico que permite muestrear de la distribución gaussiana discreta (con ciertos parámetros) en una retícula.

Veamos pues las definiciones y resultados que necesitamos sobre gaussianas discretas

2.3. Gaussianas discretas.

DEFINICIÓN 13. Sea un vector $\mathbf{x}$ y un $s > 0$ definimos la función gaussiana en $\mathbb{R}^n$ escalada por el factor s como

$$\rho_s(\mathbf{x}) := \exp\left(-\pi \|\mathbf{x}/s\|^2\right).$$

Denotaremos por

$$N_s := \rho_s / s^n$$

a la distribución gaussiana asociada.

Definición 14 (Distribución gaussiana discreta). Dado un subconjunto contable $A \subset \mathbb{R}^n$ y un parámetro $s > 0$, la distribución gaussiana discreta sobre A de parámetro s $D_{A,s}$ asigna el valor 0 $\forall x \notin A$ y los valores

$$D_{A,s}(x) = \frac{\rho_s(x)}{\rho_s(A)} \quad \forall x \in A,$$

donde $\rho_s(A) := \sum_{x \in A} \rho_s(x)$.

De hecho, para una función $f : \mathbb{R}^n \to \mathbb{C}$, y un conjunto discreto $L \subset \mathbb{R}^n$, denotaremos

$$f(L) := \sum_{x \in L} f(x). \tag{3}$$

Las gaussianas discretas se comportan esencialmente como sus análogas continuas. El parámetro de suavizado se encarga precisamente de cuantificarlo. Para definirlo, utilizaremos la retícula dual.

Definición 15 (Parámetro de suavizado). Sea una retícula n-dimensional Λ y $\varepsilon > 0$, se define el parámetro de suavizado $\eta_\varepsilon(\Lambda)$ como el más pequeño s tal que $\rho_{1/s}(\Lambda^* \setminus \{0\}) \leq \varepsilon$.

Observación 16. El parámetro de suavizado $\eta_\varepsilon(\Lambda)$ nos da la escala límite r para la cual $D_{\Lambda,r}$ se comporta como una distribución gaussiana continua. Si $r \geq \eta_\varepsilon(\Lambda)$, $D_{\Lambda,r}$ se comportará como una gaussiana continua y, por tanto, los vectores muestreados de $D_{\Lambda,r}$, tendrán con una alta probabilidad norma $r\sqrt{n}$. En el Capítulo 3 se formalizarán estas afirmaciones. Si conseguimos hacer r suficientemente pequeño, esto nos proporcionará vectores de norma pequeña, resolviendo por tanto el problema GapSVP. De ahí que para probar el Teorema 12 baste lograr muestrear de una gaussiana discreta con r suficientemente pequeña.

Por último necesitamos definir la distribución normal periódica en $\mathbb{R}$ y su discretización en $\mathbb{Z}_p$.

Definición 17 (distribución normal periódica). Sea $\beta \in \mathbb{R}^+$. Definimos la distribución normal periódica de parámetro β como la distribución obtenida de tomar muestras de una variable normal de media 0 y desviación típica $\frac{\beta}{\sqrt{2\pi}}$ reducida modulo 1. Es decir,

$$\forall r \in [0,1)\,, \Psi_\beta(r) := \sum_{k=-\infty}^{\infty} \frac{1}{\beta} \cdot exp\left(-\pi \left(\frac{r-k}{\beta}\right)^2\right) \tag{4}$$

Observación 18. A la distancia L_1 entre dos funciones de densidad, la llamaremos usualmente *distancia estadística* entre las correspondientes distribuciones de probabilidad.

Definición 19 (Discretización de una función de densidad). Dada una función de densidad en el toro $\phi : \mathbb{T} \to \mathbb{R}^+$, y un entero $p \geq 1$, definimos su discretización

$$\overline{\phi} : \mathbb{Z}_p \to \mathbb{R}^+$$

como la distribución de probabilidad discreta que se obtiene al tomar muestras de ϕ multiplicada por p y redondeada al entero más cercano módulo p. Es decir,

$$\overline{\phi}(i) = \int_{(i-1/2)/p}^{(i+1/2)/p} \phi(x)dx$$

Tenemos ya por tanto definida la distribución $\overline{\Psi}_\alpha$ utilizada en el criptosistema asociado al problema de LWE.

Podemos además hacer la demostración pendiente del Lema 7, que reescribimos nuevamente por comodidad del lector.

LEMA. *Dada la selección de parámetros establecida, se cumple para todo* $k \in \{0, 1, ..., m\}$ *y alguna función despreciable* $\delta(n)$*, que*

$$\Pr_{e \sim \bar{\Psi}_\alpha^{\star k}} \left\{ |e| < \left\lfloor \frac{p}{2} \right\rfloor /2 \right\} > 1 - \delta(n)$$

DEMOSTRACIÓN. Obtenemos una muestra de $\bar{\Psi}_\alpha^{\star k}$ tomando $x_1, ..., x_k$ de Ψ_α y calculando $\sum_{i=1}^k \lfloor p x_i \rceil \mod p$ (ver Definición 19 más abajo). Por estar este último valor alejado como mucho [2], $k \leq m < p/32$ de $\sum_{i=1}^k p x_i \mod p$, basta demostrar que, con probabilidad alta, $\left| \sum_{i=1}^k p x_i \mod p \right| < p/16$, o equivalentemente, $\left| \sum_{i=1}^k x_i \mod 1 \right| < 1/16$. Por estar $\sum_{i=1}^k x_i \mod 1$ distribuido de acuerdo a $\Psi_{\sqrt{k} \cdot \alpha}$ y $\sqrt{k} \cdot \alpha = o(1/\log n)$ la probabilidad de que $\left| \sum_{i=1}^k x_i \mod 1 \right| < 1/16$ es $1 - \delta(n)$ para alguna función despreciable $\delta(n)$, lo que es trivial a partir de la definición de Ψ_β en la ecuación (4).

□

[2]El número 32 no es especial. Se puede conseguir sin más que elegir n suficientemente grande.

Capítulo 3

Lema principal

Gracias a la Observación 16 hemos reducido el problema a probar que tener un oráculo que resuelva LWE nos permite muestrear una gaussiana discreta en la retícula Λ con parámetro r suficientemente pequeño. Esto será consecuencia del siguiente lema, que muestra cómo se puede reducir iterativamente el parámetro r de forma exponencial (r se divide por la mitad en cada iteración).

LEMA 20 (El paso iterativo). *Sea una función despreciable $\varepsilon = \varepsilon(n)$, n, p, α como en (1), (2). Supongamos que tenemos acceso a un oráculo W que resuelve el problema LWE_{p,Ψ_α} a partir de un número polinomial en n de muestras. Entonces dada una retícula n-dimensional Λ, un $r > \sqrt{2}p\eta_\varepsilon(\Lambda)$ y n^c muestras de $D_{\Lambda,r}$, (para una cierta constante $c > 0$), existe un algoritmo cuántico eficiente que produce una muestra de la distribución $D_{\Lambda,r\sqrt{n}/\alpha p}$.*

Nótese que $\frac{r\sqrt{n}}{\alpha p} \leq \frac{r}{2}$ por las condiciones (1) y (2).

La demostración pasará por un problema intermedio de retículas llamado CVP (Closest Vector Problem).

DEFINICIÓN 21 ($\text{CVP}_{\Lambda,d}$). En el problema $\text{CVP}_{\Lambda,d}$ dado un punto $\mathbf{x} \in \mathbb{R}^n$ cuya distancia a la retícula Λ es como mucho $d > 0$, se quiere encontrar el punto más cercano de Λ a $\mathbf{x}$. A dicho punto lo denotamos por $\kappa_\Lambda(x)$

Veremos en la siguiente sección la existencia de un algoritmo que, a partir de W y las muestras de $D_{\Lambda,r}$ podrá resolver el problema $CVP_{\Lambda^*,\alpha p/(\sqrt{2}r)}$. Por último, veremos en la Sección 2 la existencia de un algoritmo cuántico que, dado un oráculo que resuelve el problema $CVP_{\Lambda^*,\alpha p/(\sqrt{2}r)}$, devuelve una muestra de $D_{\Lambda,r/\sqrt{n}(\alpha p)}$.

1. De las muestras al CVP

El objetivo de esta sección es probar el siguiente lema.

LEMA 22. *Sea $\varepsilon = \varepsilon(n)$ una función despreciable, $\alpha = \alpha(n) \in (0,1)$ un número real, y el entero $p = p(n) \geq 2$. Supongamos tenemos acceso a un oráculo W que resuelve el problema LWE_{p,Ψ_α} a partir de un número polinomial en n de muestras; entonces dada una retícula n-dimensional Λ, n^c muestras de $D_{\Lambda,r}$ y un $r > \sqrt{2}p\eta_\varepsilon(\Lambda)$; existe una constante $c > 0$ y un algoritmo eficiente que resuelve el problema $CVP_{\Lambda^*,\alpha p(\sqrt{2}r)}$.*

Para realizar la prueba, deberemos primero construir una reducción del problema $CVP_{\Lambda,d}$ al $CVP^{(p)}_{\Lambda,d}$, donde diremos que un algorito resueve el problema $CVP^{(p)}_{\Lambda,d}$ si dados Λ, $p \geq 2$, $0 < d < \lambda_1(\Lambda)/2$ y un punto $\mathbf{x} \in \mathbb{R}^n$ a distancia como mucho d de Λ nos devuelve $\Lambda^{-1}\kappa_\Lambda(\mathbf{x}) \bmod p \in \mathbb{Z}_p^n$. Después, tras comprobar cómo nos es suficiente con dar un algoritmo que resuelva el problema $CVP^{(p)}_{\Lambda,d}$ para demostrar el Lema 22, construiremos el algoritmo para este último problema.

LEMA 23. *Sea una retícula* Λ, $p \geq 2$, $d < \lambda_1(\Lambda)/2$, *y supongamos que tenemos acceso a un oráculo que resuelve el* $CVP^{(p)}_{\Lambda,d}$. *Entonces existe un algoritmo para resolver el* $CVP_{\Lambda,d}$ *de manera eficiente.*

La clave de la demostración será la existencia de un algoritmo eficiente, llamado Nearest Plane Algorithm (NPA) y desarrollado por Babai en 1986, que toma como entrada un punto x que está a distancia $\leq d$ de una retícula Λ y devuelve un punto de la retícula que está a distancia como mucho $2^n d$, donde n es el rango de Λ.

DEMOSTRACIÓN. Partimos de un punto $\mathbf{x}$ a distancia $\leq d$ de nuestra retícula. Vamos a construir el algoritmo de la siguiente manera:

Paso 1 Construimos iterativamente $\mathbf{x_1}, \mathbf{x_2}$,$\mathbf{x_3}$....
El primer punto $\mathbf{x_1} = \mathbf{x}$.
Dado $\mathbf{x_i}$, construimos $\mathbf{x_{i+1}}$ como sigue:
Sea $\mathbf{a}_i = \Lambda^{-1}\kappa_\Lambda(\mathbf{x}_i) \in \mathbb{Z}^n$, entonces

$$\mathbf{x_{i+1}} = (\mathbf{x_i} - \Lambda(\mathbf{a_i} \mod p))/p$$

Nótese que, aunque $\mathbf{a_i}$ no se puede obtener a priori, el oráculo nos permite obtener $\mathbf{a_i} \mod p$, que es lo único que hace falta en la definición iterativa de los $\mathbf{x_i}$

Por ser $\Lambda(\mathbf{a}_i - (\mathbf{a}_i \mod p))/p \in \Lambda$ el punto más cercano a $\mathbf{x}_{i+1}$, se deduce que $\mathbf{a}_{i+1} = (\mathbf{a}_i - (\mathbf{a}_i \mod p))/p$. Además, la distancia de $\mathbf{x_i}$ a Λ es a lo sumo d/p^i.
Así, tras n iteraciones, obtendremos un punto $\mathbf{x}_{n+1}$ a distancia d/p^n.

Paso 2: Aplicamos el algoritmo NPA para obtener un punto de la retícula, y por tanto de la forma $\Lambda\mathbf{a}$ con $\mathbf{a} \in \mathbb{Z}^n$, a distancia como mucho $2^n \cdot d/p^n \leq d < \lambda_1(\Lambda)/2$ de $\mathbf{x}_{n+1}$. Por tanto $\Lambda\mathbf{a}$ es el punto de Λ más cercano a $\mathbf{x}_{n+1}$, con lo que $\mathbf{a}$ es precisamente el $\mathbf{a_{n+1}}$ asociado a $\mathbf{x_{n+1}}$.

Paso 3: Para cada k desde n hasta 1 obtenemos $\mathbf{a_k}$ a partir de $\mathbf{a_{k+1}}$ mediante la ecuación $\mathbf{a}_n = p\mathbf{a}_{n+1} + (\mathbf{a}_n \mod p)$, donde $\mathbf{a}_n \mod p$ se obtiene nuevamente a través del oráculo.

La salida del algoritmo es $\Lambda\mathbf{a}_1$, que es precisamente el punto de la retícula más cercano a $\mathbf{x}_1 = \mathbf{x}$.

□

Necesitamos ahora del siguiente lema.

LEMA 24. *Sea una retícula n-dimensional* Λ *y* $\varepsilon > 0$. *Se tiene*

$$\eta_\varepsilon(\Lambda) \geq \sqrt{\frac{\log(1/\varepsilon)}{\pi}} \cdot \frac{1}{\lambda_1(\Lambda^*)}$$

DEMOSTRACIÓN. Sean $\mathbf{v} \in \Lambda^*$ un vector de longitud $\lambda_1(\Lambda^*)$ y $s = \eta_\varepsilon(\Lambda)$, entonces

$$\varepsilon = \rho_{1/s}\left(\Lambda^* \setminus \{0\}\right) \geq \rho_{1/s}(\mathbf{v}) = \exp\left(-\pi\left(s\lambda_1(\Lambda^*)\right)^2\right)$$

□

Como de este lema y las hipótesis del Lema 22 se obtiene que

$$\alpha p/\sqrt{2}r \leq 1/\eta_\varepsilon(\Lambda) \leq \lambda_1(\Lambda^*)/2,$$

estamos en las condiciones del lema 23. Por tanto, para probar el lema 22 nos es suficiente con dar con un algoritmo que que resuelva el problema $CVP^{(p)}_{\Lambda^*,\alpha p(\sqrt{2}r)}$ a partir de un oráculo que resuelva LWE.

Ese ese el contenido del siguiente lema.

LEMA 25. *Sea una función despreciable $\varepsilon = \varepsilon(n)$, un entero $p = p(n) \geq 2$ y un real $\alpha = \alpha(n) \in (0,1)$. Supongamos que tenemos acceso a un oráculo que $\forall \beta \leq \alpha$, con β desconocido encuentra $\boldsymbol{s}$ a partir de un número polinómico de muestras de la distribución A_{s,Ψ_β}. Entonces dada una retícula n-dimensional Λ, un $r > \sqrt{2}p\eta_\varepsilon(\Lambda)$ y un número polinómico de muestras de la distribución $D_{\Lambda,r}$, existe un algoritmo eficiente que resuelve el problema $CVP^{(p)}_{\Lambda^*,\alpha p/\sqrt{2}r}$.*

Para verlo necesitaremos de algunos lemas y conceptos previos.

1.1. Lemas y conceptos previos.

LEMA 26 (fórmula de sumación de Poisson). *Para toda retícula Λ y una función $f : \mathbb{R}^n \to \mathbb{C}$ se tiene que*

$$f(\Lambda) = \det(\Lambda^*)\widehat{f}(\Lambda^*)$$

siendo $\widehat{f}$ la transformada de Fourier de f:

$$\widehat{f}(\mathbf{w}) = \int_{\mathbb{R}^n} f(\mathbf{x}) e^{-2\pi i \langle \mathbf{x}, \mathbf{w} \rangle} d\mathbf{x}.$$

Los siguientes lemas formalizan las propiedades del parámetro de suavizado comentadas en la Observación 16 y que se necesitan en la demostración.

LEMA 27. *Sea una retícula Λ, $c \in \mathbb{R}^n$, $\varepsilon > 0$ y $r \geq \eta_\varepsilon$, se tiene que $\rho_r(\Lambda + \mathbf{c})$ está entre los valores $r^n \det(\Lambda^*)(1-\varepsilon)$ y $r^n \det(\Lambda^*)(1+\varepsilon)$, lo que denotamos por*

$$\rho_r(\Lambda + \boldsymbol{c}) \in r^n \det(\Lambda^*)(1 \pm \varepsilon)$$

DEMOSTRACIÓN. Definimos $\rho_{r,-\mathbf{c}}(\mathbf{x}) := \rho_r(\mathbf{x} + \mathbf{c})$. Aplicando la fórmula de sumación de Poisson tenemos que

$$\rho_r(\Lambda + \mathbf{c}) = \sum_{\mathbf{x} \in \Lambda} \rho_r(\mathbf{x} + \mathbf{c}) = \sum_{\mathbf{x} \in \Lambda} \rho_{r,-\mathbf{c}}(\mathbf{x}) = \det(\Lambda^*) \sum_{\mathbf{y} \in \Lambda^*} \widehat{\rho}_{r,-c}(\mathbf{y})$$

Usamos ahora el buen comportamiento de la transformada de Fourier con las traslaciones, en concreto, que

$$\widehat{\rho}_{r,-c}(\mathbf{y}) = \exp(2\pi i \langle \mathbf{c}, \mathbf{y} \rangle)\, \widehat{\rho}_r(\mathbf{y}),$$

y que la transformada de Fourier de la gaussiana es ella misma:

$$\widehat{\rho}_r(\mathbf{y}) = r^n\, \rho_{1/r}(\mathbf{y}).$$

Esto nos permite obtener que

$$\rho_r(\Lambda + \mathbf{c}) = r^n \det(\Lambda^*) \sum_{\mathbf{y} \in \Lambda^*} \exp(2\pi i \langle \mathbf{c}, \mathbf{y} \rangle) \rho_{1/r}(\mathbf{y}) \in r^n \det(\Lambda^*)(1 \pm \varepsilon)\,,$$

donde en el último paso hemos usado que $\rho_{1/r}(\Lambda^* \setminus \{0\}) \leq \varepsilon$, por ser $r \geq \eta_\varepsilon$. □

LEMA 28. *Sea una retícula Λ, un vector $\mathbf{u} \in \mathbb{R}^n$, dos reales $r, s > 0$ y $t = \sqrt{r^2 + s^2}$. Supongamos $rs/t = 1/\sqrt{1/r^2 + 1/s^2} \geq \eta_\varepsilon(\Lambda)$ para algun $\varepsilon < \frac{1}{2}$. Consideremos entonces la distribución continua Y en $\mathbb{R}^n$ obtenida al tomar muestras de $D_{\Lambda+\mathbf{u},r}$ y añadirles un vector de ruido tomado de N_s. Entonces la distancia estadística entre N_t e Y será, como mucho, 4ε.*

La idea del lema es probar que si sumamos una amplitud s gaussiana continua, a una amplitud r gaussiana discreta , donde ambas amplitudes son más grandes que el parámetro de suavizado, la distribución resultante está muy próxima a la distribución gaussiana de amplitud $\sqrt{r^2 + s^2}$.

Consecuentemente, para $r > \sqrt{2}\eta_\varepsilon(\Lambda)$, si muestreamos de la distribución $D_{\Lambda,r}$ y le añadimos el ruido gaussiano N_r obtenemos una distribución con distancia estadística como mucho 4ε con respecto de la distribución gaussiana $N_{\sqrt{2}r}$. Como $N_{\sqrt{2}r}$ es la distribución que se obtiene al sumar dos muestras independientes de N_r, se tiene que el ruido N_r es suficiente para esconder la estructura discreta de $D_{\Lambda,r}$.

DEMOSTRACIÓN. Usando que la distribución de probabildad de la suma de variables aleatorias independientes viene dada por la convolución, podemos escribir la función de densidad de Y como

$$\begin{aligned} Y(\mathbf{x}) &= \frac{1}{s^n \rho_r(\Lambda + \mathbf{u})} \sum_{\mathbf{y} \in \Lambda + \mathbf{u}} \rho_r(\mathbf{y}) \rho_s(\mathbf{x} - \mathbf{y}) \\ &= \frac{1}{s^n \rho_r(\Lambda + \mathbf{u})} \sum_{\mathbf{y} \in \Lambda + \mathbf{u}} e^{-\pi\left(\|\mathbf{y}/r\|^2 + \|(\mathbf{x}-\mathbf{y})/s\|^2\right)} \\ &= \frac{1}{s^n \rho_r(\Lambda + \mathbf{u})} \sum_{\mathbf{y} \in \Lambda + \mathbf{u}} e^{-\pi\left(\frac{r^2+s^2}{r^2 \cdot s^2} \cdot \left\|\mathbf{y} - \frac{r^2}{r^2+s^2}\mathbf{x}\right\|^2 + \frac{1}{r^2+s^2}\|\mathbf{x}\|^2\right)} \end{aligned}$$

donde la última igualdad se comprueba expandiendo $\|\cdot\|^2 = \langle \cdot, \cdot \rangle$. Operando y utilizando la notación para la gaussiana desplazada introducida en el la demostración del lema anterior, tenemos que

$$\begin{aligned} Y(\mathbf{x}) &= e^{-\frac{\pi}{r^2+s^2}\|\mathbf{x}\|^2} \frac{1}{s^n \rho_r(\Lambda + \mathbf{u})} \sum_{\mathbf{y} \in \Lambda + \mathbf{u}} e^{-\pi \frac{r^2+s^2}{r^2 \cdot s^2} \cdot \left\|\mathbf{y} - \frac{r^2}{r^2+s^2}\mathbf{x}\right\|^2} \\ &= \frac{1}{s^n} \rho_t(\mathbf{x}) \cdot \frac{\rho_{rs/t,(r/t)^2\mathbf{x}-\mathbf{u}}(\Lambda)}{\rho_{r,-\mathbf{u}}(\Lambda)} \end{aligned}$$

que, utilizando la fórmula de sumación de Poisson en el numerador y el denominador, es igual a

$$\begin{aligned} &= \frac{1}{s^n} \rho_t(\mathbf{x}) \cdot \frac{\widehat{\rho}_{rs/t,(r/t)^2\mathbf{x}-\mathbf{u}}(\Lambda^*)}{\widehat{\rho}_{r,-\mathbf{u}}(\Lambda^*)} \\ &= \frac{\rho_t(\mathbf{x})}{t^n} \cdot \frac{(t/rs)^n \widehat{\rho}_{rs/t,(r/t)^2\mathbf{x}-\mathbf{u}}(\Lambda^*)}{(1/r)^n \widehat{\rho}_{r,-\mathbf{u}}(\Lambda^*)} \quad (*) \end{aligned}$$

Ahora, utilizando como hicimos en el lema anterior el buen comportamiento de la transformada de Fourier con las traslaciones y las gaussianas, se tiene que

$$\widehat{\rho}_{rs/t,(r/t)^2\mathbf{x}-\mathbf{u}}(\mathbf{w}) = e^{-2\pi i \langle (r/t)^2\mathbf{x}-\mathbf{u}, \mathbf{w} \rangle} \cdot (rs/t)^n \rho_{t/rs}(\mathbf{w})$$

y que

$$\widehat{\rho}_{r,-u}(\mathbf{w}) = e^{2\pi i\langle \mathbf{u},\mathbf{w}\rangle} \cdot r^n \rho_{1/r}(\mathbf{w}),$$

de donde se desprende

$$\left|1 - (t/rs)^n \widehat{\rho}_{rs/t,(r/t^2)\mathbf{x}-\mathbf{u}}(\Lambda^*)\right| \leq \rho_{t/rs}(\Lambda^* \setminus \{0\}) \leq \varepsilon$$

$$|1 - (1/r)^n \widehat{\rho}_{r,-\mathbf{u}}(\Lambda^*)| \leq \rho_{1/r}(\Lambda^* \setminus \{0\}) \leq \varepsilon$$

por ser $r \geq rs/t \geq \eta_\epsilon(\Lambda)$.
Con lo que la fracción de la derecha en (*) queda entre $(1-\varepsilon)/(1+\varepsilon) \geq 1-2\varepsilon$ y $(1+\varepsilon)/(1-\varepsilon) \leq 1+4\varepsilon$ lo que implica

$$|Y(\mathbf{x}) - \rho_t(\mathbf{x})/t^n| \leq \rho_t(\mathbf{x})/t^n \cdot 4\varepsilon.$$

La demostración se concluye integrando el anterior resultado en $\mathbb{R}^n$. Recordamos que la distancia estadística se definía como la distancia en L_1 entre las respectivas funciones densidad. □

LEMA 29. *Sea la retícula Λ, los vectores $\mathbf{u}, \mathbf{z} \in \mathbb{R}^n$ y los reales $r, \alpha > 0$; supongamos que $1/\sqrt{1/r^2 + (\|\mathbf{z}\|/\alpha)^2} \geq \eta_\varepsilon(\Lambda)$ para $\varepsilon > \frac{1}{2}$. Entonces la distribución $\langle \mathbf{z}, \mathbf{v}\rangle + e$ donde v se distribuye de acuerdo con $D_{\Lambda+\mathbf{u},r}$ y e es una variable normal de media 0 y desviación típica $\alpha/\sqrt{2\pi}$ está a distancia estadística, como mucho, 4ε de una variable normal de media 0 y desviación típica $\sqrt{(r\|\mathbf{z}\|)^2 + \alpha^2}/\sqrt{2\pi}$. En particular, como la distancia estadística no puede aumentarse con la aplicación de una función, la distribución $\langle \mathbf{z}, \mathbf{v}\rangle + e$ mod 1 estará como mucho a distancia 4ε de $\Psi_{\sqrt{(r\|\mathbf{z}\|^2)+\alpha^2}}$.*

DEMOSTRACIÓN. Es trivial que $\langle \mathbf{z}, \mathbf{v}\rangle + e$ es exactamente $\langle \mathbf{z}, \mathbf{v}+\mathbf{h}\rangle$ donde $\mathbf{h}$ se distribuye de acuerdo a $N_{\alpha/\|\mathbf{z}\|}$. Además, por el lema 28 sabemos que la distribución de $\mathbf{v}+\mathbf{h}$ está como mucho a distancia 4ε de la distribución $N_{\sqrt{r^2+(\alpha/\|\mathbf{z}\|)^2}}$. Al tomar el producto escalar con $\mathbf{z}$ obtenemos la distribución normal de media 0 y desviación típica $\sqrt{(r\|\mathbf{z}\|)^2 + \alpha^2}/\sqrt{2\pi}$. □

Podemos ya demostrar el Lema 25.

1.2. Prueba del Lema 25.

Por comodidad del lector, reproducimos nuevamente el Lema 25

LEMA. *Sea una función despreciable $\varepsilon = \varepsilon(n)$, un entero $p = p(n) \geq 2$ y un real $\alpha = \alpha(n) \in (0,1)$. Supongamos que tenemos acceso a un oráculo que $\forall \beta \leq \alpha$, con β desconocido encuentra $\boldsymbol{s}$ a partir de un número polinómico de muestras de la distribución A_{s,Ψ_β}. Entonces dada una retícula n-dimensional Λ, un $r > \sqrt{2}p\eta_\varepsilon(\Lambda)$ y un número polinómico de muestras de la distribución $D_{\Lambda,r}$, existe un algoritmo eficiente que resuelve el problema $CVP^{(p)}_{\Lambda^*,\alpha p/\sqrt{2}r}$.*

Para realizar la demostración describiremos un procedimiento $\boldsymbol{P}$ que dado un punto $\mathbf{x}$ a distancia como mucho $\alpha p/\left(\sqrt{2}r\right)$ de Λ^*, nos devuelva muestras de la distribución $A_{\mathbf{s},\Psi_\beta}$ para algun $\beta \leq \alpha$, donde $\mathbf{s} = (\Lambda^*)^{-1}\kappa_{\Lambda^*}(\mathbf{x})$ mod p. Aplicando el oráculo al resultado de ejecutar este procedimiento $\boldsymbol{P}$ un número polinómico de veces podremos encontrar $\mathbf{s}$.

El procedimiento $\boldsymbol{P}$ es muy sencillo:

- Tomar una muestra $\mathbf{v} \in D_{\Lambda,r}$. Denotaremos $\mathbf{a} = \Lambda^{-1}\mathbf{v} \mod p$.
- Producir $(\mathbf{a}, \langle \mathbf{x}, \mathbf{v} \rangle / p + e \mod 1)$, tomando $e \in \mathbb{R}$ de la distribución normal con desviación típica $\alpha/(2\sqrt{\pi})$.

Solo queda ver que la distribución que se obtiene mediante este procedimiento está a una distancia estadística insignificante de $A_{\mathbf{s},\Psi_\beta}$, (para algún $\beta \leq \alpha$).

Para ello hay que ver que la distribución de la que se toma $\mathbf{a}$ es *esencialmente* la distribución uniforme y que la distribución $\langle \mathbf{x}, \mathbf{v} \rangle / p + e \mod 1$ es *esencialmente* Ψ_β.

Por un lado, al ser la probabilidad de obtener $\mathbf{a} \in \mathbb{Z}_p^n$ proporcional a $\rho_r(p\Lambda + \Lambda\mathbf{a})$; utilizando el lema 27 y que $\eta_\varepsilon(p\Lambda) = p\eta_\varepsilon(\Lambda) < r$ se tiene que la probabilidad de obtener $\mathbf{a}$ está en $(r/p)^n \det(\Lambda^*)(1 \pm \varepsilon)$. Al ser $\epsilon = \epsilon(n)$ una función despreciable, se deduce que la distancia estadística entre la distribución de la que se toma $\mathbf{a}$ y la distribución uniforme es también despreciable.

Ahora, fijado un valor de $\mathbf{a}$, consideramos la distribución de $\langle \mathbf{x}, \mathbf{v} \rangle / p + e \mod 1$. Denotamos $\mathbf{x}' = \mathbf{x} - \kappa_{\Lambda^*}(\mathbf{x})$. Por hipótesis $\|\mathbf{x}'\| \leq \alpha p / \left(\sqrt{2}r\right)$. Se tiene que

$$\langle \mathbf{x}, \mathbf{v} \rangle / p + e \mod 1 = \langle \mathbf{x}'/p, \mathbf{v} \rangle + e + \langle \kappa_{\Lambda^*}(\mathbf{x}), \mathbf{v} \rangle / p \mod 1.$$

Además, como $\Lambda^{-1} = (\Lambda^*)^T$, se tiene que

$$\langle \kappa_{\Lambda^*}(\mathbf{x}), \mathbf{v} \rangle = \left\langle (\Lambda^*)^{-1} \kappa_{\Lambda^*}(\mathbf{x}), \Lambda^{-1}\mathbf{v} \right\rangle,$$

lo que a su vez implica por las definiciones de s y a que

$$\langle \kappa_{\Lambda^*}(\mathbf{x}), \mathbf{v} \rangle / p \mod 1 = \langle \mathbf{s}, \mathbf{a} \rangle / p \mod 1.$$

Ahora, al estar condicionando en $\mathbf{a}$, la distribución de $\mathbf{v}$ es $D_{p\Lambda+\Lambda\mathbf{a},r}$. Por tanto, el lema 25 nos permite concluir que el término restante $\langle \mathbf{x}'/p, \mathbf{v} \rangle + e$ se encuentra a una distancia estadística despreciable de Ψ_β con $\beta = \sqrt{(r\|\mathbf{x}'/p\|)^2 + \alpha^2/2} \leq \alpha$, tal y como queríamos.

2. del CVP a las muestras

Para continuar con la segunda parte de la prueba del lema principal (lema 20) deberemos demostrar la existencia de un algoritmo cuántico que a partir de un oráculo del problema CVP genere muestras de la distribución gaussiana discreta. Formalmente:

LEMA 30. *Sea una retícula n-dimensional Λ, un número $d < \lambda_1(\Lambda^*)/2$ y un oráculo que resuelva el problema $CVP_{\Lambda^*,d}$; entonces existe un algoritmo cuántico eficiente que devuelve una muestra de la distribución $D_{\Lambda,\sqrt{n}/\sqrt{2}d}$.*

El algoritmo que tenemos que construir para probar el Lema 30 consta de dos pasos. El primero es la creación de un estado cuántico de n-qubits asociado a la probabilidad gaussiana discreta de amplitud r, con r lo suficientemente grande si se compara on $\lambda_n(\Lambda)$. El segundo es obtener las muestras deseadas a partir de ese estado cuántico.

Para esta sección necesitamos algunos conceptos y resultados de computación cuántica, así como algunos resultados previos sobre retículas.

2.1. Conceptos básicos computación cuántica. Vamos a resumir aquí los conceptos sobre computación cuántica que son necesarios en este TFG.

DEFINICIÓN 31 (un qubit). se trata de la unidad fundamental del modelo de computación cuántica. No es más que el espacio de Hilbert complejo de dimensión 2: $\mathbb{H} = \mathbb{C}^2$. Los dos elementos de la base canónica (también llamada computacional) son los **estados básicos** $|0\rangle$ y $|1\rangle$.

DEFINICIÓN 32 (n qubits). : El espacio de Hilbert asociado a n qubits es el producto tensorial

$$\boldsymbol{H}_n = \boldsymbol{H} \otimes ..^{(n)}.. \otimes \boldsymbol{H}$$

Tiene dimensión 2^n y su base canónica (o computacional) está dada por los elementos

$$\{|x_1\rangle \otimes \cdots \otimes |x_n\rangle\}_{x_j \in \{0,1\}}$$

Para simplificar se denota:

$$|x_1\rangle \otimes \cdots \otimes |x_n\rangle = |x_1 \cdots x_n\rangle = |\mathbf{x}\rangle$$

con $\mathbf{x} \in \{0,1\}^n$.

DEFINICIÓN 33. Un *estado cuántico* no es más que un estado de norma euclídea 1 en el correspondiente espacio de Hilbert.

Como $\{0,1\}^n$ tiene 2^n elementos, podemos expresar un estado cuántico de n-qubits de la siguiente manera

$$\phi = \sum_{j=0}^{2^n-1} a_j |j\rangle, \text{ con } \sum_{j=0}^{2^n-1} |a_j|^2 = 1$$

Si se mide el estado ϕ en la base computacional, se obtiene el resultado j con probabilidad $|a_j|^2$.

DEFINICIÓN 34. Un algortimo cuántico de n qubits no es más que una aplicación unitaria (en particular lineal) en el espacio $\mathbb{H}_n$. Será eficiente si se puede descomponer en una cantidad polinomial (en n) de unitarias que actúan solo en uno o dos qubits.

Nosotros aquí solo necesitaremos dos tipos de algoritmos cuánticos, que son precisamente los dos que aparecen en las principales aplicaciones de la computación cuántica, como el algoritmo de Shor [**3**].

El primero es lo que denominamos **computación en superposición** [**3**].

TEOREMA 35 (Computación en superposición). *Dado un algoritmo eficiente clásico f, existe un algoritmo eficiente cuántico U_f (que es por tanto un operador unitario) que lleva $|x,0\rangle$ a $|x,f(x)\rangle$ para todo $|x\rangle$ en la base computacional. Al ser U_f lineal, el mismo algoritmo cuántico lleva por tanto $\sum_{x\in\{0,1\}^n} a_x|x,0\rangle$ a*

$$\sum_{x\in\{0,1\}^n} a_x|x,f(x)\rangle$$

El segundo es la **transformada cuántica de Fourier**, que es eficiente y está definida como sigue:

LEMA 36 (Transformada cuática de Fourier). *la transformada cuántica de Fourier de n-qubits es el operador unitario* $\boldsymbol{F}_n : \boldsymbol{H}_n \to \boldsymbol{H}_n$ *definido unívocamente a través de*

$$F_n|j\rangle = \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} \sigma_n^{jk}|k\rangle, \quad 0 \leq j \leq 2^n - 1$$

donde $\sigma_n = e^{\frac{2\pi i}{2^n}}$.

2.2. Lemas previos sobre retículas.

DEFINICIÓN 37 (paralelepípedo fundamental). El paralelepípedo P_Λ para una retícula Λ determinada por los vectores que la generan,

$$\Lambda = Gen(b_1, b_2, ..., b_n) = \{a_1b_1 + ... + a_nb_n : a_1, ..., a_n \in \mathbb{Z}\}$$

es aquel cuyos 2^n vértices son de la forma $a_1b_1 + ... + a_nb_n$ con $a_i \in \{0, 1\}\, \forall i$.

LEMA 38 (Banaszczyk). *Sea* B_n *la bola euclídea de radio uno, una retícula* Λ*, y cualquier* $r > 0$*,*

$$\rho_r(\Lambda \setminus \sqrt{nr}B_n) < 2^{-2n} \cdot \rho_r(\Lambda)$$

con $\Lambda \setminus \sqrt{nr}B_n$*, el conjunto de puntos de* Λ *con norma mayor que* $\sqrt{nr}$.

LEMA 39 (Banaszczyk). *Para cualquier retícula* Λ *de rango n se cumple que*

$$1 \leq \lambda_1(\Lambda) \cdot \lambda_n(\Lambda^*) \leq n$$

LEMA 40. *:* $\forall s, t, l > 0$ *y* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *con* $\|\mathbf{x}\| \leq t$ *y* $\|\mathbf{x} - \mathbf{y}\| \leq l$ *se tiene que*

$$\rho_s(\mathbf{y}) \geq (1 - \pi(2lt + l^2)/s^2)\rho_s(\mathbf{x})$$

.

DEMOSTRACIÓN. Empleando la desigualdad $e^{-z} \geq 1 - z$ se tiene que

$$\rho_s(\mathbf{y}) = e^{-\pi\|\mathbf{y}/s\|^2} \geq e^{-\pi(\|\mathbf{x}\|/s + l/s)^2} = e^{-\pi(2l\|\mathbf{x}\|/s^2 + (l/s)^2)}.$$

Por tanto

$$\rho_s(\mathbf{y}) \geq (1 - \pi(2lt + l^2)/s^2)\rho_s(\mathbf{x}).$$

□

Podemos ya centrarnos en la demostración del Lema 30

2.3. Primer paso: creación del estado cuántico.

Este primer paso se resume en el siguiente lema, cuya demostración se reduce esencialmente al algoritmo de Grover-Rudolph [**5**] de generación de estados cuánticos asociados a distribuciones de probabilidad.

LEMA 41. *Dada una retícula n-dimensional* $\Lambda \subseteq \mathbb{Z}^n$ *y* $r > 2^{2n}\lambda_n(\Lambda)$ *existe un algoritmo eficiente que genera un estado cuántico que está dentro de la distancia euclídea* $2^{-\Omega(n)}$ *del estado cuántico normalizado*

$$\sum_{\mathbf{x}\in\Lambda} \sqrt{\rho_r(\mathbf{x})} \mid \mathbf{x}\rangle = \sum_{\mathbf{x}\in\Lambda} \rho_{\sqrt{2}r}(\mathbf{x}) \mid \mathbf{x}\rangle$$

2.4. Segundo paso: muestrear usando el estado cuántico y el oráculo. El segundo paso tiene que objetivo acabar la demostración del Lema 30, que por conveniencia del lector reproducimos aquí.

LEMA. *Sea una retícula n-dimensional Λ, un número $d < \lambda_1(\Lambda^*)/2$ y un oráculo que resuelva el problema $CVP_{\Lambda^*,d}$; entonces existe un algoritmo cuántico eficiente que devuelve una muestra de la distribución $D_{\Lambda,\sqrt{n}/\sqrt{2}d}$.*

DEMOSTRACIÓN. Suponemos que $d = \sqrt{n}$, sin perdida de generalidad escalando los datos de entrada. Sea un entero $R \geq 2^{3n}\lambda_n(\Lambda^*)$. Como, dada Λ, R puede ser calculado en tiempo polinómico, suponemos que $\log R$ es polinómico con respecto al tamaño de nuestra entrada.

Consideramos la retícula dual reescalada, que tiene por matriz Λ^*/R. Por el Lema 41 aplicado a Λ^*/R y $r = \frac{1}{\sqrt{2}}$ podemos construir de forma eficiente (salvo error exponencialmente pequeño) el estado

$$\sum_{\mathbf{x}\in\Lambda^*/R} \rho(\mathbf{x}) \mid \mathbf{x}\rangle$$

Por el lema 38 este estado es también exponencialente cercano a

$$\sum_{\mathbf{x}\in\Lambda^*/R, \|\mathbf{x}\|<\sqrt{n}} \rho(\mathbf{x}) \mid \mathbf{x}\rangle$$

Utilizamos ahora la *computación en superposición* (ver Sección 2.1) asociada a $f(x) = x \mod P(\Lambda^*)$, con $P(\Lambda^*)$ el paralepípedo fundamental de Λ^* (Definición 37). Obtenemos así el estado

$$\sum_{\mathbf{x}\in\Lambda^*/R, \|\mathbf{x}\|<\sqrt{n}} \rho(\mathbf{x}) \mid \mathbf{x}, \mathbf{x} \mod P(\Lambda^*)\rangle$$

Ahora es el paso clave en el que usamos el oráculo CVP para, a partir de x mód $P(\Lambda^*)$ recuperar x. Si lo hacemos como computación en superposición, llegamos al estado

$$\sum_{\mathbf{x}\in\Lambda^*/R, \|\mathbf{x}\|<\sqrt{n}} \rho(\mathbf{x}) \mid \mathbf{x} \mod P(\Lambda^*)\rangle$$

Hacemos ahora la siguiente observación, cuya demostración postponemos hasta el final de la sección.

OBSERVACIÓN 42. Sea un entero $R \geq 1$ y una retícula n-dimensional L tal que $\lambda_1(L) > 2\sqrt{n}$, entonces la distancia euclídea entre los estados cuánticos normalizados siguientes es $2^{-\Omega(n)}$

$$\mid \vartheta_1\rangle = \sum_{\mathbf{x}\in L/R, \|\mathbf{x}\|<\sqrt{n}} \rho(\mathbf{x}) \mid \mathbf{x} \mod P(\Lambda)\rangle$$

$$\mid \vartheta_2\rangle = \sum_{\mathbf{x}\in L/R} \rho(\mathbf{x}) \mid \mathbf{x} \mod P(\Lambda) = \sum_{\mathbf{x}\in (L/R)\,\cap\, P(L)} \sum_{\mathbf{y}\in L} \rho(\mathbf{x}-\mathbf{y}) \mid \mathbf{x}\rangle.$$

Tenemos por tanto un estado exponencialmente próximo a

$$\sum_{\mathbf{x}\in(\Lambda^*/R)\cap P(\Lambda)} \sum_{\mathbf{y}\in\Lambda} \rho(\mathbf{x}-\mathbf{y}) \mid \mathbf{x}\rangle.$$

El siguiente paso será aplicar la transformada cuántica de Fourier. Para ello identificamos $(\Lambda^*/R)\cap P(\Lambda^*)$ con $\mathbb{Z}_R^n$, con lo que podemos reescribir nuestro estado como

$$\sum_{\mathbf{s}\in\mathbb{Z}_R^n}\sum_{\mathbf{r}\in\mathbb{Z}^n}\rho\left(\Lambda^*\mathbf{s}/R-\Lambda^*\mathbf{r}\right)\mid\mathbf{s}\rangle.$$

La transformada cuántica de Fourier nos devuelve un estado proporcional a

$$\begin{aligned}
&\sum_{\mathbf{t}\in\mathbb{Z}_R^n}\sum_{\mathbf{s}\in\mathbb{Z}_R^n}\sum_{\mathbf{r}\in\mathbb{Z}^n}\rho\left(\Lambda^*\mathbf{s}/R-\Lambda^*\mathbf{r}\right)\exp\left(2\pi i\left\langle\mathbf{s},\mathbf{t}\right\rangle/R\right)|t\rangle\\
&\quad=\sum_{\mathbf{t}\in\mathbb{Z}_R^n}\sum_{\mathbf{s}\in\mathbb{Z}_R^n}\rho\left(\Lambda^*\mathbf{s}/R\right)\exp\left(2\pi i\left\langle\mathbf{s},\mathbf{t}\right\rangle/R\right)|t\rangle\\
&\quad=\sum_{\mathbf{t}\in\mathbb{Z}_R^n}\sum_{\mathbf{x}\in\Lambda^*/R}\rho\left(\mathbf{x}\right)\exp\left(2\pi i\left\langle(\Lambda^*)^{-1}\mathbf{x},\mathbf{t}\right\rangle\right)|t\rangle\\
&\quad=\sum_{\mathbf{t}\in\mathbb{Z}_R^n}\sum_{\mathbf{x}\in\Lambda^*/R}\rho\left(\mathbf{x}\right)\exp\left(2\pi i\left\langle\mathbf{x},\Lambda t\right\rangle\right)|t\rangle\\
&\quad\overset{(\star)}{=}\det(R\Lambda)\sum_{\mathbf{t}\in\mathbb{Z}_R^n}\sum_{\mathbf{y}\in R\Lambda}\rho(\mathbf{y}-\Lambda\mathbf{t})|t\rangle\\
&\quad=\sum_{\mathbf{x}\in(R\Lambda)\cap\Lambda}\sum_{\mathbf{y}\in R\Lambda}\rho(\mathbf{y}-\mathbf{x})\mid\mathbf{x}\rangle,
\end{aligned}$$

donde en el paso $(\star)$ hemos utilizado la fórmula de sumación de Poisson, el hecho de que $\hat{\rho}=\rho$ y la fórmula para la transformada de Fourier de una función trasladada. Es importante hacer notar que

$$\lambda_1(R\Lambda)=R\lambda_1(\Lambda)\geq R/\lambda_n(\Lambda^*)\geq 2^{3n}\ ,$$

con lo que la Observación 42 nos garantiza que el estado que tenemos está exponencialmente cerca de

$$\sum_{\mathbf{x}\in\Lambda,\|\mathbf{x}\|<\sqrt{n}}\rho(\mathbf{x})\mid\mathbf{x}\mod P(R\Lambda)\rangle$$

Basta medir ahora en la base computacional para obtener $\mathbf{x}\mod P(R\Lambda)$ para algún $\mathbf{x}$ con $\|\mathbf{x}\|<\sqrt{n}$. Si tenemos en cuenta que $\mathbf{x}\mod P(R\Lambda)$ está como mucho a distancia $\sqrt{n}$ de $R\Lambda$ y $\lambda_1(R\Lambda)\geq 2^{3n}$ podemos recuperar $\mathbf{x}$ utilizando el algoritmo NPA de Babai ya comentado anteriormente.

Para acabar basta ver que la distribución de $\mathbf{x}$ es exponencialmente cercana a $D_{\Lambda,1/\sqrt{2}}$. Por un lado, para cada $\mathbf{x}\in\Lambda,\|\mathbf{x}\|<\sqrt{n}$, la probabilidad de obtener $\mathbf{x}$ es proporcional a $\rho(\mathbf{x})^2=\rho_{1/\sqrt{2}}(\mathbf{x})$. Además, por el lema 38 toda la distribución de probabilidad de $D_{\Lambda,1/\sqrt{2}}$, menos una fracción exponencialmente pequeña, se acumula en puntos de norma menor que $\sqrt{n}$. □

Acabamos la sección, y la demostración del teorema principal del trabajo, con la demostración pendiente de la observación 42.

Demostración. (Obs. 42) Pensemos en $\mid\vartheta_1\rangle$ y $\mid\vartheta_2\rangle$ como vectores de $\mathbb{R}^n$. Sea $\boldsymbol{Z}$ la norma euclídea de $\mid\vartheta_1\rangle$. Se quiere demostrar que la distancia entre $\mid\vartheta_1\rangle$ y $\mid\vartheta_2\rangle$ es, como mucho $2^{-\Omega(n)}$ (es decir, que la distancia euclídea entre estos estados es exponencialmente pequeña). Para ello primero debemos obtener un buen estimador para $\boldsymbol{Z}$.

Como $\lambda_1(L) > 2\sqrt{n}$, sabemos que cada 'ket' de la definición de $\mid \vartheta_1\rangle$ aparecerá en el sumatorio tan solo una vez.

$$\boldsymbol{Z} = \sum_{\mathbf{x}\in L/R, \|\mathbf{x}\|<\sqrt{n}} \rho(\mathbf{x})^2 = \rho(\sqrt{2}L/R \cap \sqrt{2n}B_n)$$

Aplicando ahora el Lema 38 a la retícula $\sqrt{2}L/R$ obtenemos

$$(1 - 2^{-2^n})\rho(\sqrt{2}L/R) \leq \boldsymbol{Z} \leq \rho(\sqrt{2}L/R). \tag{5}$$

Acotamos ya la distancia euclídea de los dos vectores, utilizando que $\|\cdot\|_2 \leq \|\cdot\|_1$.

$$\| \mid \vartheta_1\rangle\langle\vartheta_1 \mid - \mid \vartheta_2\rangle\langle\vartheta_2 \mid \|_2 \leq \| \mid \vartheta_1\rangle\langle\vartheta_1 \mid - \mid \vartheta_2\rangle\langle\vartheta_2 \mid \|_1 = \sum_{\mathbf{x}\in\Lambda/R, \|\mathbf{x}\|\geq\sqrt{n}} \rho(\mathbf{x}) \leq$$

$$\leq 2^{-2n}\rho(\Lambda/R) \quad \text{Por el Lema 38}$$

$$\leq 2^{-2n}2^{n/2}\rho(\sqrt{2}\Lambda/R) \quad \text{Trivial por la definición de } \rho$$

$$\leq 2^n\rho(\sqrt{2}\Lambda/R)$$

que, unido a (5)) concluye la demostración. □

Bibliografía

[1] S. Khot, Hardness of approximating the shortest vector problem in lattices, J. ACM. 52 (5): 789–808 (2005).

[2] O. Regev, On lattices, learning with errors, random linear codes, and cryptography. Journal of the ACM (JACM), 56(6), 1-40 (2009).

[3] M. Nielsen, I. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, 2000.

[4] C. Gentry, Fully Homomorphic Encryption Using Ideal Lattices. In Proc. 41st ACM Symp. on Theory of Computing (STOC), pág. 169-178 (2019).

[5] L. Grover and T. Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions. Arxiv:quant-ph/0208112, 2002.

[6] L. Chen et al. Report on Post-Quantum Cryptography. National Institute of Standards and Technology, 2016. http://dx.doi.org/10.6028/NIST.IR.8105

[7] G. Alagic et al, Status Report on the First Round of the NIST Post-Quantum Cryptography Standardization Process. National Institute of Standards and Technology, 2019. 27p. https://doi.org/10.6028/NIST.IR.8240

[8] F. Arute et al. Quantum supremacy using a programmable superconducting processor. Nature 574.7779 (2019): 505-510.