Factoring, Lattices and the NP-hardness of the Shortest Vector Problem

Daniele Micciancio

UC San Diego

May 2021

Factoring

Theorem (Fundamental Theorem of Arithmetic)

Every integer greater than 1 can be represented (uniquely) as the product of prime numbers.

(Euclid, Elements Book VII & IX, c. 300 BC)

- Factoring problem: given N find its prime factors
- Special case: factor $N = p \cdot q$
  - Hardest case in practice
  - Basis of the RSA cryptosystem (Rivest, Shamir, Adleman, 1977), (Cooks, 1973)
  - Classic problem in cryptography
- No known polynomial time algorithm
- Efficiently solvable in quantum polynomial time (Shor, 1994)

Shortest Lattice Vectors

Theorem (Convex Body Theorem)

Any symmetric convex body $\mathcal{B} \subset \mathbb{R}^n$ of volume $vol(\mathcal{B}) > 2^n$ contains a nonzero integer vector $\mathrm{x} \in \mathbb{Z}^n \setminus \{0\}$

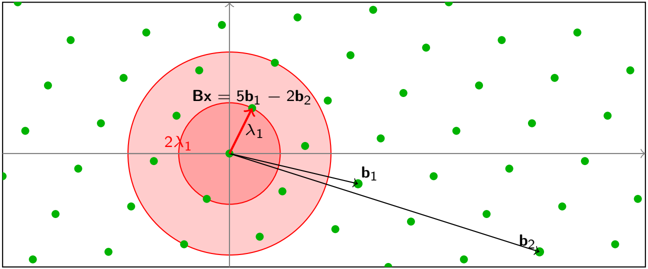
(Minkowski, 1889)

- Equivalent lattice formulation: any lattice $\mathbf{B}\mathbb{Z}^n$ contains a short nonzero vector $\mathbf{Bx}$
- Different convex bodies give different norm bounds:
  - $\|\mathbf{Bx}\|_\infty \leq |\det(\mathbf{B})|^{1/n}$
  - $\|\mathbf{Bx}\|_2 \leq \sqrt{n} \cdot |\det(\mathbf{B})|^{1/n}$
  - ...
- Shortest Vector Problem (SVP): given a lattice basis $\mathbf{B}$, find a short(est) nonzero lattice vector $\mathbf{Bx}$. ($\lambda_1 = \|\mathbf{Bx}\|$.)

Shortest Vector Problem

Definition (Shortest Vector Problem, SVP_γ)

Given a lattice $\mathcal{L}(\mathbf{B})$, find a (nonzero) lattice vector $\mathbf{Bx}$ (with $\mathbf{x} \in \mathbb{Z}^k$) of length (at most) $\|\mathbf{Bx}\| \leq \gamma\lambda_1$



Definition ($\mathrm{DecisionSVP}_\gamma$, informal)

Approximate the value of λ_1, without finding a short vector.

Factoring vs SVP

- Factoring:
  - Unlikely to be NP-hard (subexponential algorithms, quantum polynomial time)
  - Conjectured not in (classic) polynomial time
- SVP (Euclidean norm)
  - LLL (Lenstra, Lestra, Lovasz, 1982) solves it "in practice" in relatively small dimension (< 50)
  - Conjectured to be solvable in polynomial time through the 1980s and early 1990s
  - NP-hardness (under deterministic reductions): still an open problem!

Prime numbers lattice (Schnorr, 1991)

- Use lattice algorithms (e.g., LLL) to factor numbers
- Map the multiplicative structure of the integers to the additive structure of a lattice

$$\mathbf{B} = \begin{bmatrix} \sqrt{\ln p_1} & & \\ & \ddots & \\ & & \sqrt{\ln p_n} \\ \alpha \log p_1 & \cdots & \alpha \log p_n \end{bmatrix}$$

$$\sum_i e_i \log p_i = \log \prod_i p_i^{e_i}$$

- Use LLL to find "smooth congruences"
- Factoring method based on the Quadratic Sieve (Pomerance, 1981). See Leo's talk for details.

From Factoring Algorithm to NP-hardness proof

- (Schnorr 1991) Use prime number lattice to (heuristically) factor numbers via lattice reduction
- (Adleman 1995) Attempt to give a rigorous proof that factoring reduces to SVP
  - Maybe SVP is not NP-hard
  - Can we prove it is at least as hard as factoring?
  - Attempt to turn Schnorr's algorithm into a formal reduction
- (Ajtai 1998) SVP is NP-hard under randomized reduction
  - Started from Adleman unfinished manuscript
  - Same goal: reduce factoring to SVP via prime number lattice
  - Ended up proving that SVP is NP-hard under randomized reduction
  - Proof is highly technical, uses many additional ideas and technique
- Much follow up work on simplifying and strengthening Ajtai's proof

NP-hardness of SVP

- NP-hard in the ℓ_∞ norm (Van Emde Boas, 1981)
- NP-hardness in ℓ_2: long standing open problem
- NP-hard under randomized reductions [Ajtai 1998]
- Improved to $\gamma < \sqrt{2}$ [Micciancio 1998]
- Improved to any constant γ [Khot 2001]
- Improvements and simplifications [Haviv, Regev 2007]
- Improvements and simplifications [Micciancio 2012]
- All use randomized reductions

Open problem

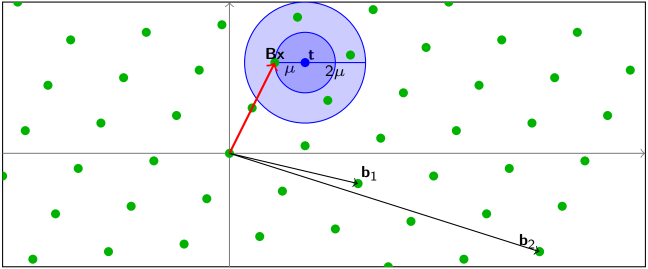
Prove the NP-hardness of SVP in ℓ_2 norm under deterministic reductions

- Randomness used only to construct locally dense lattice.

Closest Vector Problem

Definition (Closest Vector Problem, CVP_γ)

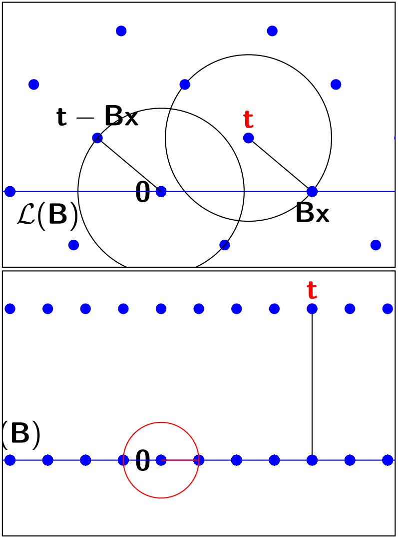
Given a lattice $\mathcal{L}(\mathbf{B})$ and a target point $\mathbf{t}$, find a lattice vector $\mathbf{Bx}$ within distance $\|\mathbf{Bx} - \mathbf{t}\| \leq \gamma\mu$ from the target



NP-hardness of CVP

- NP-hard in any ℓ_p norm (van Emde Boas, 1981)
- CVP': Hard even if solution is in $\mathbf{B}\{0,1\}^n$
- NP-hard to approximate for any constant factor (Arora, Babai, Stern, Sweedyk, 1993) and more (Dinur, Kindler, Raz, Safra, 2003)
- CVP with preprocessing (CVPP):
  - Still NP-hard (Micciancio 2001), even to approximate (Feige, M. 2002), (Regev 2003), (Alekhnovich, Khot, Kindler, Vishnoi, 2011)
  - the lattice **B** is fixed and can be pre-processed arbitrarily
  - NP-hard instance is encoded just in the target vector!
- SVP reduced to CVP (Goldreich, M., Safra, Seifert, 1999)
- Question: Can you reduce CVP to SVP?

Reducing CVP to SVP



- Goal: find lattice point $\mathbf{v} \in \mathcal{L}(\mathbf{B})$ closest to $\mathbf{t}$
- Idea: find shortest vector $\mathbf{w} \in \mathcal{L}([\mathbf{B}, \mathbf{t}])$
- If $\mathbf{w} = \mathbf{t} - \mathbf{Bx}$, then $\mathbf{v} = \mathbf{Bx}$ is closest to $\mathbf{t}$.
- Problem: what if $\lambda(\mathcal{L}(\mathbf{B})) < \mathit{dist}(\mathbf{t}, \mathcal{L}(\mathbf{B}))$?
- Example:

$$\mathcal{L}(\mathbf{B}) = \mathbb{Z}^n \qquad \mathbf{t} = \left(\frac{1}{2}, \ldots, \frac{1}{2}\right)$$

$$\lambda(\mathcal{L}(\mathbf{B})) = 1 < \frac{\sqrt{n}}{2}$$

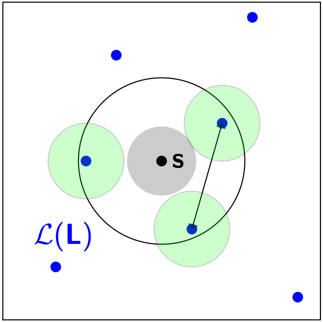
Repairing the reduction

- Goal (CVP'): find lattice point $\mathbf{v} \in \mathbf{B}\{0,1\}^n \subset \mathcal{L}(\mathbf{B})$ closest to $\mathbf{t}$
- Embed $\mathbf{B}$ and $\mathbf{t}$ in higher dimension so that
  - $\lambda(\mathcal{L}(\mathbf{B}))$ gets large
  - $\mathbf{t}$ remains close to $\mathcal{L}(\mathbf{B})$

$$\mathbf{B} \Longrightarrow \begin{bmatrix} \mathbf{BTL} \\ \mathbf{L} \end{bmatrix} \qquad \mathbf{t} \Longrightarrow \begin{bmatrix} \mathbf{t} \\ \mathbf{s} \end{bmatrix}$$

Locally Dense Lattice:

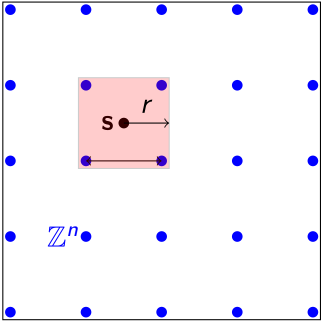
- $\lambda(\mathcal{L}(\mathbf{L})) > d$
- $|\mathcal{L}(\mathbf{L}) \cap \mathcal{B}(\mathbf{s}, r)|$ is large
- $r < d < 2r$
- $\{0,1\}^n \subset \mathbf{T}(\mathcal{L}(\mathbf{L}) \cap \mathcal{B}(\mathbf{s}, r)) \subset \mathbb{Z}^n$



Locally Dense Lattice in ℓ_∞

Trivial Construction:

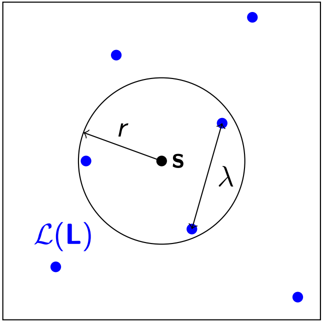
- $\mathcal{L}(\mathbf{L}) = \mathbb{Z}^n$
- $d = \lambda(\mathcal{L}(\mathbf{L})) = 1$
- $\mathbf{s} = (\frac{1}{2}, \ldots, \frac{1}{2})$,
- $r > \frac{1}{2} = d/2$
- $\mathcal{L}(\mathbf{L}) \cap \mathcal{B}_\infty(\mathbf{s}, r) = \{0, 1\}^n$



Locally Dense Lattices in ℓ_2

$$\mathbf{L} = \begin{bmatrix} \sqrt{\ln p_1} & & \\ & \ddots & \\ & & \sqrt{\ln p_n} \\ \alpha \log p_1 & \cdots & \alpha \log p_n \end{bmatrix} \quad \mathbf{s} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \alpha \ln \beta \end{bmatrix} \quad \begin{matrix} \lambda \approx \sqrt{2\beta} \\ r \approx \sqrt{\beta} \end{matrix}$$

- $p_1, \ldots, p_n$ odd primes, $\alpha = \beta^{1-\epsilon}$
- Multiplicative structure of $\prod p_i \in \mathbb{Z}$ maps to additive structure of $\mathcal{L}(\mathbf{B})$
- if $[\beta, \beta + \beta^\epsilon]$ contains many products $\prod_{i \in I} p_i$, then $\mathcal{B}(\mathbf{s}, r)$ contains many lattice vectors $\sum_{i \in I} \mathbf{b}_i$.



Locally Dense Lattices in ℓ_2 (cont.)

Conjecture

For all $\epsilon > 0$, and (large enouh) n, the interval $[n, n + n^\epsilon]$ contains a square free number with prime factors $< \log^{O(1)} n$

- How to choose β:
  - Deterministically, assuming conjecture
  - At random: works with high probability
- Alternative construction based on BCH codes, but still randomized [Micciancio 2012]
- Open problem: Find unconditional deterministic construction

Packing density and Hermite's factor

- Hermite's factor:

$$\gamma(\mathcal{L}) = \left(\frac{\lambda_1(\mathcal{L})}{\det(\mathcal{L})^{1/n}}\right)^2$$

- Minkowski's theorem: $\gamma(\mathcal{L}) \leq O(n)$
- Use lattice $\mathcal{L} \subset \mathbb{R}^n$ to pack $\mathbb{R}^n$ with disjoint balls $\mathbf{v} + \mathcal{B} \cdot r$ of radius $r = \lambda_1/2$ and center $\mathbf{v} \in \mathcal{L}$
- Packing density:

$$\mathrm{vol}(\mathcal{B} \cdot r) = \frac{\mathrm{vol}(\mathcal{B})(\lambda_1/2)^n}{\det(\mathcal{L})} = \mathrm{vol}(\mathcal{B}) \left(\frac{\sqrt{\gamma(\mathcal{L})}}{2}\right)^n$$

- Minkowki's theorem: density cannot be higher than 1
- Dense lattices: $\gamma(\mathcal{L})$ close to Minkowski's bound $O(n)$

Global Density vs Local Density

- Fix a radius $r = \lambda_1 / c$ for some constant $c \geq 1$
- Global density: expected number of lattice points in $\mathbf{s} + \mathcal{B} \cdot r$ when $\mathbf{s} \in \mathbb{R}^n$ is chosen uniformly at random (modulo $\mathcal{L}$)
  - Must be < 1 if $c > 2$
  - Can be > 1 if $c < 2$
  - Can be exponentially large if $c < \sqrt{2}$
- The global density of a lattice is precisely $\mathrm{vol}(\mathcal{B}r) / \det(\mathcal{L})$
- If $\gamma(\mathcal{L})$ is close to Minkowski's bound, and $c > 0.5$, then the global density is exponentially large
- There exists a "locally dense" center $\mathbf{s}$ such that $\mathbf{s} + \mathcal{B} \cdot r$ contains exponentially many lattice points

How to find a “locally dense” center?

- Goal: find a center $\mathbf{s}$ such that $\mathbf{s} + \mathcal{B}r$ contains many lattice points, for some $r < \lambda_1/\sqrt{2}$
- Choose $\mathbf{s}$ at random within $\mathcal{B}r \subset \mathbb{R}^n$
- $\mathbf{0}$ is always in $\mathbf{s} + \mathcal{B}r$
- By symmetry, $\mathbf{s} \in \mathbb{R}^n/\mathcal{L}$ is chosen with probability proportional to the number of lattice points in $\mathbf{s} + \mathcal{B}r$

The geometry of the prime numbers lattice

Prime number lattice:

$$\mathbf{B} = \begin{bmatrix} \sqrt{\ln p_1} & & \\ & \ddots & \\ & & \sqrt{\ln p_n} \\ \alpha \log p_1 & \cdots & \alpha \log p_n \end{bmatrix}$$

"Complexity of Lattice Problems" (M., Goldwasser, 2002), Prop. 5.9

Theorem (Lemma 5.3)

$$\lambda \geq 2 \ln \alpha$$

Theorem (Prop. 5.9)

$$\det(\mathrm{B}) = \sqrt{\left(1 + \alpha^2 \sum_k \ln p_k\right) \prod_k \ln p_k}$$

Density of the prime numbers lattice

- $\lambda \geq 2 \ln \alpha$
- $\det(\mathbf{B}) = \sqrt{(1 + \alpha^2 \sum_k \ln p_k) \prod_k \ln p_k}$
- Hermite factor is maximized setting $p_1, \ldots, p_n$ to the first n prime numbers, and $\alpha \approx e^{n/2}$
- Hermite factor $\gamma = \Omega(n/\log n)$ close to Minkowski's bound
- The prime number lattice is globally dense
- Lattice points in a small ball centered around $(0, \ldots, 0, \alpha b)$ corresponds to subset-products of $\{p_1, \ldots, p_n\}$ close to b
- Lattice density corresponds to density of square-free p_n-smooth numbers in small intervals

Smooth numbers and derandomization

Conjecture

For all sufficiently large n, the interval $[n, n + n^\epsilon]$ contains at least one square-free $(\log n)^{O(1)}$-smooth number.

- If the smooth number conjecture is true, then SVP is NP-hard under deterministic reductions.
- Conjecture is easy to prove for $\epsilon = 1$
- $\epsilon = 0.5$ is considered a serious barrier in mathematics
- SVP NP-hardness needs conjecture for $\epsilon \ll 0.5$
- Can we find some other locally dense lattice contruction?

Locally Dense Lattices from BCH codes

- $\mathbb{F} = \{0, 1\}$: finite field with 2 elements
- $\mathbb{F}^n$ vector space with Hamming metric
- Linear codes $C[n, k, d]$: k-dimensional subspaces of $\mathbb{F}^n$ with minimum dinstance d
- (Extended) BCH codes $\mathbb{F}^n = C_0 \supset C_1 \supset \cdots \supset C_h$, where $C_i[n, k_i, d_i]$ for $d_i \geq 4^i$ and $k_i \geq n - (\log n)(4^i/2 - 1)$
- Barnes-Sloane lattice (Construction D)

$$\mathcal{L} = \sum_i C_i \cdot 2^{h-i}$$

Theorem

The Barnes-Sloane lattice satisfies $\lambda \geq 2^h$ *and* $\det \leq n^{\frac{2}{3}4^h}$.

NP-hardness of SVP using Barnes-Sloane lattice

- (Micciancio 2012) Barnes-Sloane lattice to give alternate proof that SVP is NP-hard under randomized reductions (with one sided error)
- Selection of the dense center still required randomization
- New proof uses special tensoring properties of this lattice to show that SVP is NP-hard to approximate within any constant factor
- NP-hardness proofs based on the prime number lattice stopped working for approximation factors $> \sqrt{2}$
- Other techniques to prove NP-hardess for any constant factor introduced more randomness and two-side error

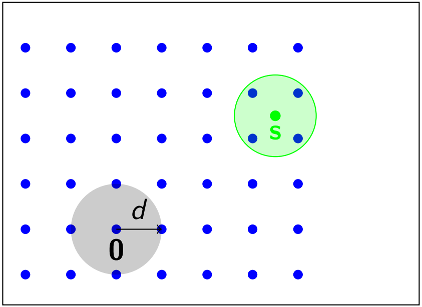
Locally Dense Codes

A locally dense code consists of

- A linear code $L[h, m, d]$
- A radius $r < d$
- A center $\mathbf{s}$ such that

$$X = \mathcal{B}(\mathbf{s}, r) \cap L$$

has size $|X| \geq 2^k$



Often required also a linear transformation $\mathbf{T}$ such that

$$\mathbf{T}(X) = \{0, 1\}^k$$

Minimum Distance Problem (MDP)

- SVP for codes: find the shortest codeword in a linear code
- NP-hard to solve exactly [Vardy 1996]
- NP-hard to approximate (for any $\gamma \geq 1$) under randomized reductions [Dumer, M., Sudan 1999] using locally dense codes
- Derandomized in [Cheng, Wan 2009] using powerful mathematical tools (Weil's character sum bound on affine line)
- Simplified and extended to asymptotically good codes [Khot, Austrin 2011], but using additional techniques
- Deterministic reduction using locally dense codes [Micciancio, 2014]

Building Locally Dense Codes

- Start from a binary linear code $C[n, k, d]$ with $d/n > 1/\sqrt{6}$.
  - Many classic constructions achieve $d \approx n/2$. E.g., concatenate Reed-Solomon codes over $\mathbb{F}_{2^h}$ with Hadamard code.
- Use C to define a binary code $L[4n^2, k(k+1)/2, 6d^2]$
  - Represent $4n^2$-dim vectors by four $n \times n$ matrices

$$(\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3, \mathbf{W}_4)$$

  - Consider ball of radius $r = n^2 < 6d^2$ centered around

$$(\mathbf{O}, \mathbf{O}, \mathbf{O}, \mathbf{U})$$

  where $\mathbf{U} = \mathbf{u}\mathbf{u}^\top$ is the all 1 matrix.
- If $d \approx n/2$, then $r \approx \frac{2}{3}(6d^2)$

Construction

- The code L is the set of all codewords

$$\mathbf{W} = (\mathbf{Y}, \mathbf{Y} + \mathbf{u}\mathbf{y}^\top, \mathbf{Y} + \mathbf{y}\mathbf{u}^\top, \mathbf{Y} + \mathbf{u}\mathbf{y}^\top + \mathbf{y}\mathbf{u}^\top)$$

  where $\mathbf{Y} = \mathbf{C}\mathbf{X}\mathbf{C}^\top$ for some symmetric matrix $\mathbf{X} = \mathbf{X}^\top \in \mathbb{F}_2^{k \times k}$ and $\mathbf{y} = \text{diagonal}(\mathbf{Y}) = \mathbf{C} \cdot \text{diagonal}(\mathbf{X})$.
- Notice: $\mathbf{y}$,columns($\mathbf{Y}$), rows($\mathbf{Y}$) $\in C[n, k, d]$
- L has block length $4n^2$
- $\mathbf{W}$ is linear in $\mathbf{X}$
- The dimension is $k(k+1)/2$
- To be proved:
  - the minimum distance is at least $6d^2$
  - there are 2^k codewords within distance n^2 from $(\mathbf{O}, \mathbf{O}, \mathbf{O}, \mathbf{U})$

Decoding Dense Lattices

- Bounded Distance Decoding: CVP when target point **t** is within the unique decoding radius $\lambda/2$
- (Ducas, Pierrot, 2019) give efficient BDD algorithm for prime numbers lattice,
- (Mook, Peikert, 2020) give efficient BDD (and list decoding) algorithm for Barnes-Sloane lattice
- Both lattices previously used for proving NP-hardness of SVP.
  - Is there any connection?
  - Can the BDD algorithms be used to find the locally dense centers?
  - Can you efficiently solve CVP in these or other locally dense lattices?
  - Can you solve BDD/CVP in lattices achieving $\gamma(\mathcal{L}) = \Omega(n)$? (E.g., Mordell-Weil lattices)

Open Problems

- Reduce factoring to approximate SVP for approximation factors $\gamma > \sqrt{n}$:
  - $\sqrt{n}$-approximate SVP is in $NP \cap coNP$, and unlikely to be NP-hard
  - Is $\sqrt{n}$-approximate SVP at least as hard as factoring?
- Derandomization of Locally Dense Lattice construction
  - Implies NP-hardness of SVP under deterministic reduction, a long standing open problem
  - Several deterministic dense lattice constructions
  - some are based on linear codes
  - Randomness only used to find dense center
  - Locally Dense Codes have been successfully derandomized

Want to know more?

- *"The shortest vector problem is NP-hard to approximate to within some constant"*, Micciancio, SIAM J. Computing, 2001.
- *"Inapproximability of the Shortest Vector Problem: Toward a deterministic reduction"*, Micciancio, Theory of Computing, 2012
- *"Locally Dense Codes"*, Micciancio, Computational Complexity Conference, 2014
- *"Polynomial time bounded distance decoding near Minkowski's bound in discrete logarithm lattices"*, Ducas, Pierrot, Des. Codes Cryptogr. 2019
- *"Lattice (List) Decoding Near Minkowski's Inequality"*, Mook, Peikert, arXiv 2020