Online Learning and Online Algorithms: Chasing Convex Bodies

ARIS MARKOGIANNAKIS

ADVANCED ALGORITHMS, NTUA

26/8/2024

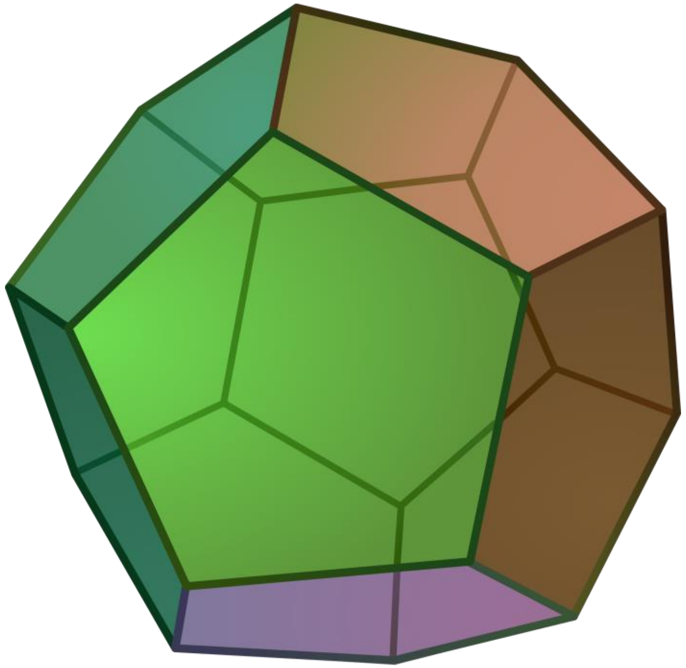
Contents

1. Online Learning and Online Algorithms

- Online learning refers to a setting where an algorithm learns and makes decisions sequentially.
- The algorithm receives data points one by one and must make decisions or predictions based on the current and past observations, without knowing future data.
- Application: Used in environments where data is received incrementally, such as financial markets, adaptive routing in networks, and dynamic pricing.

2. The Chasing Convex Bodies Problem

Convex Body: A compact (closed and bounded) convex set with non-empty interior



2. The Chasing Convex Bodies Problem

For each timestep $t \in \mathbb{N}$, a convex body $K_t \subseteq \mathbb{R}^d$ is given

The player picks a point $x_t \in K_t$

Total Movement Cost: $ALG = \sum_{t=0}^{T-1} ||x_{t+1} - x_t||$

Our goal is to minimize the competitive ratio: $cr(ALG) = \max_{\sigma}\{\frac{ALG(\sigma)}{OPT(\sigma)}\}$

σ an instance of the problem and $OPT(\sigma)$ is the optimal offline cost

Lower bound for competitive ratio: $\Omega(\sqrt{d})$

2. The Chasing Convex Bodies Problem

- Application: Load Balancing in Cloud Computing
- Dynamically allocate computational resources while minimizing energy consumption and latency.
- Varying demand patterns, adapting server allocations in real-time, maintaining system performance within constraints
- Convex Bodies($\boldsymbol{K_t}$) represents feasible regions of server allocations constrained by current system demands.
- At each time step, select an allocation (x_t) that lies within the current convex body and optimizes the competitive ratio

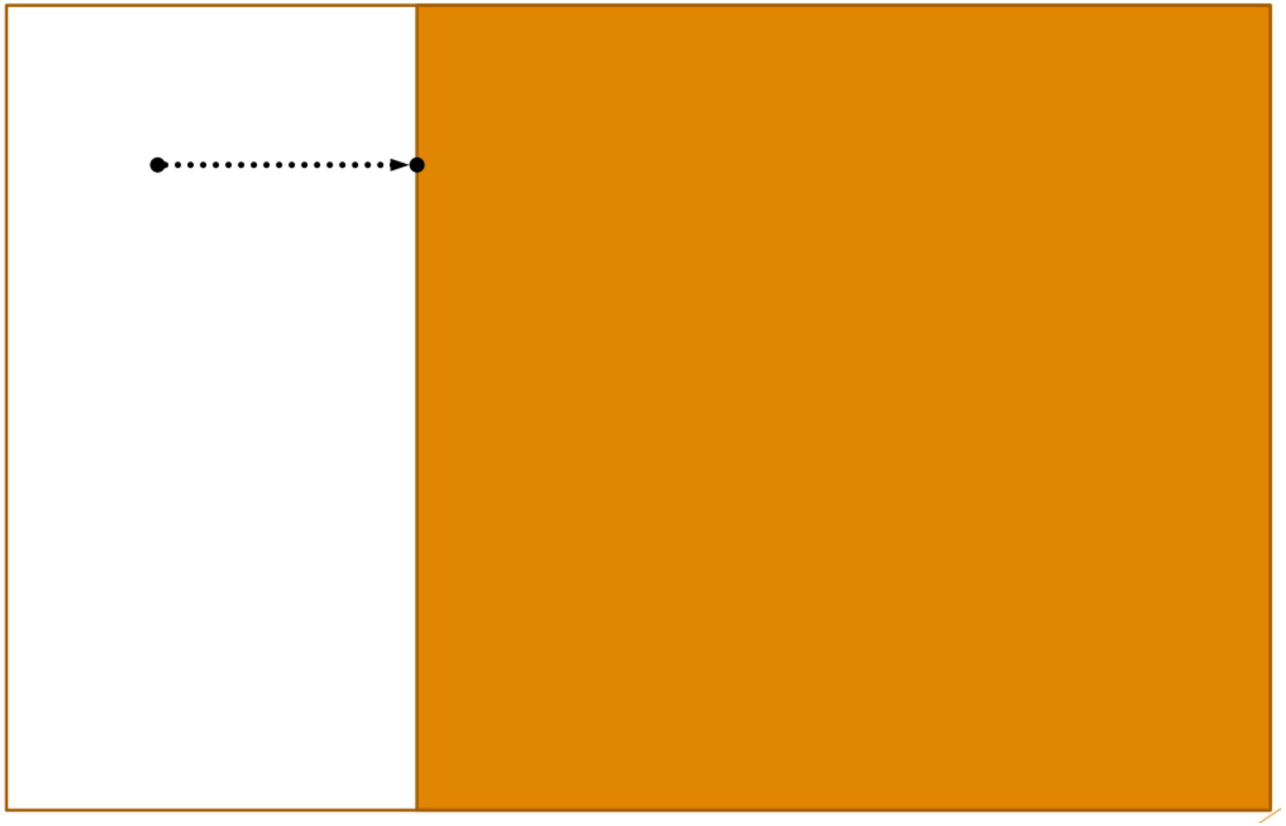
2. The Chasing Convex Bodies Problem

We will focus on the nested version of the problem.

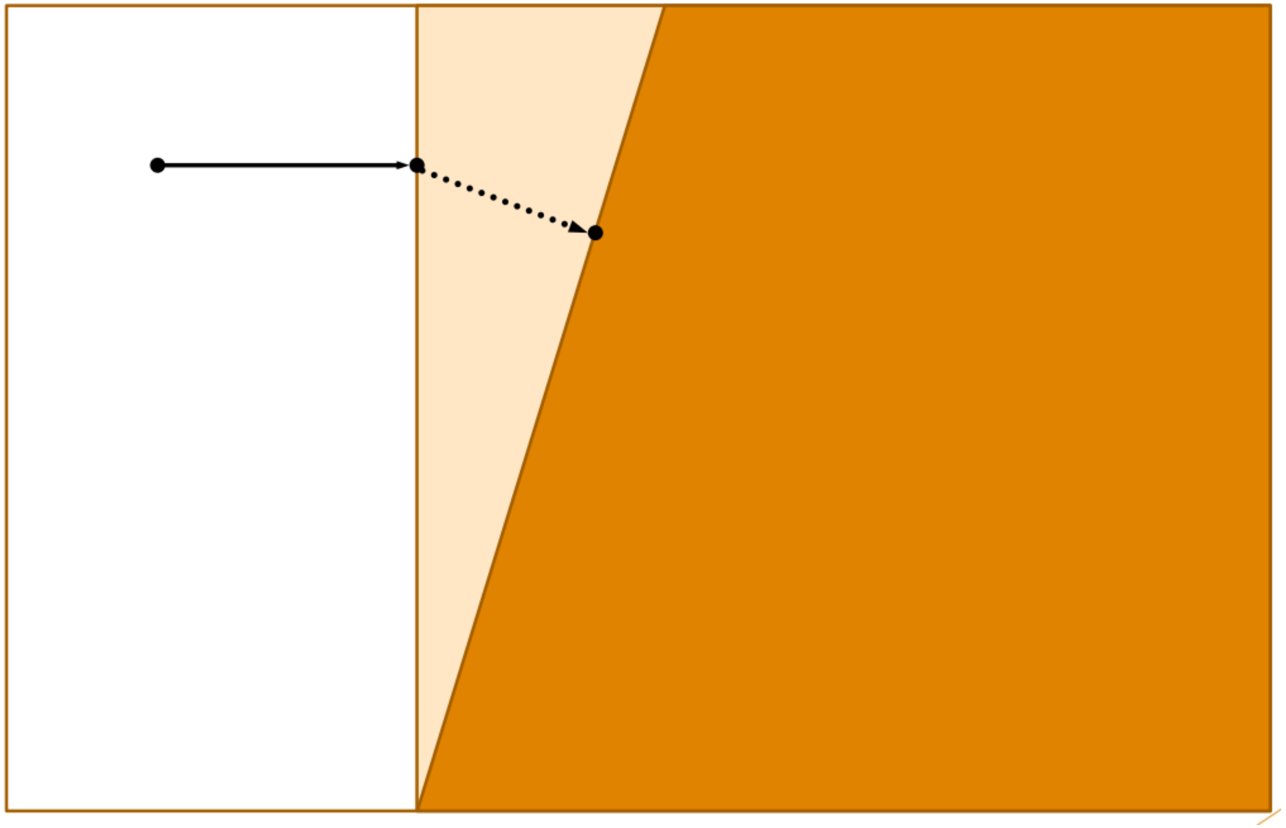
$$K_1 \supseteq K_2 \supseteq \cdots \supseteq K_T$$

In the nested version, each convex body contains the next one, forming a hierarchy of containment.

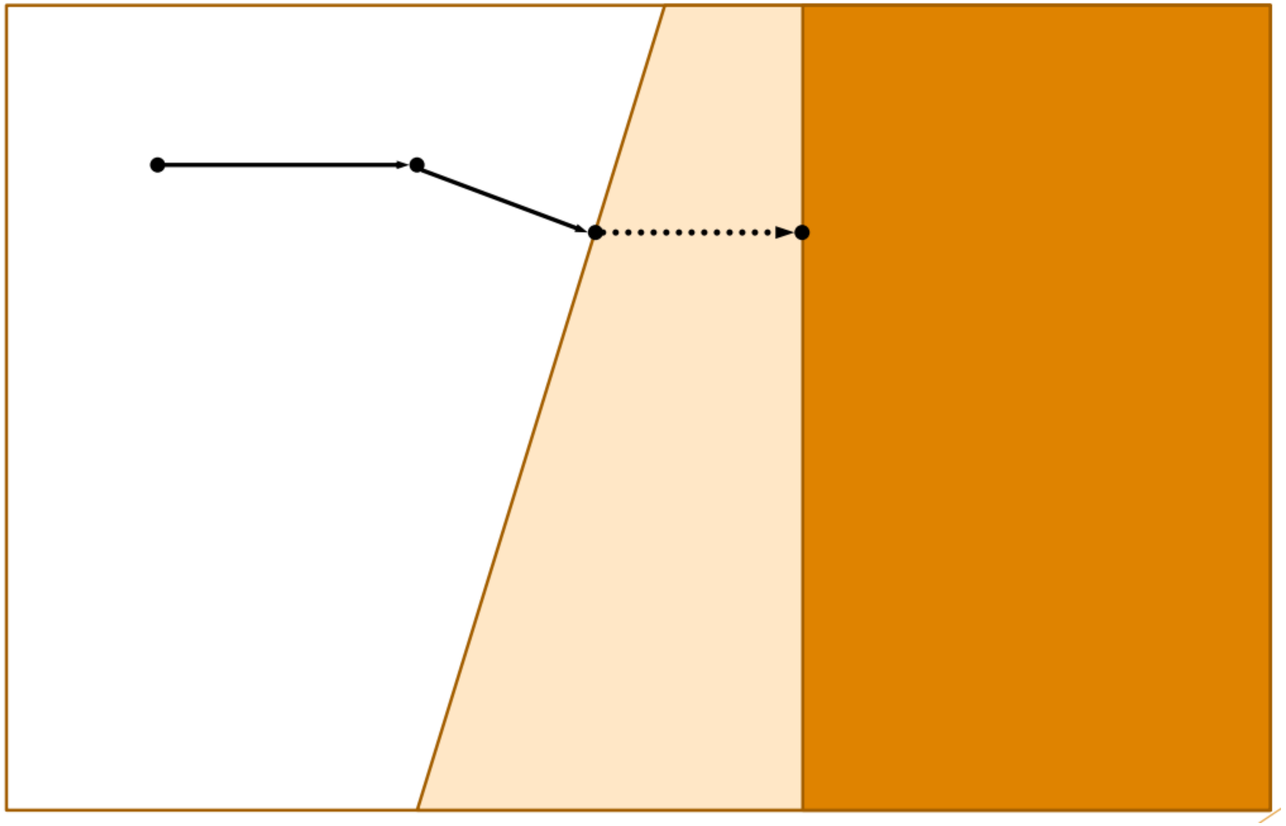
2. The Chasing Convex Bodies Problem



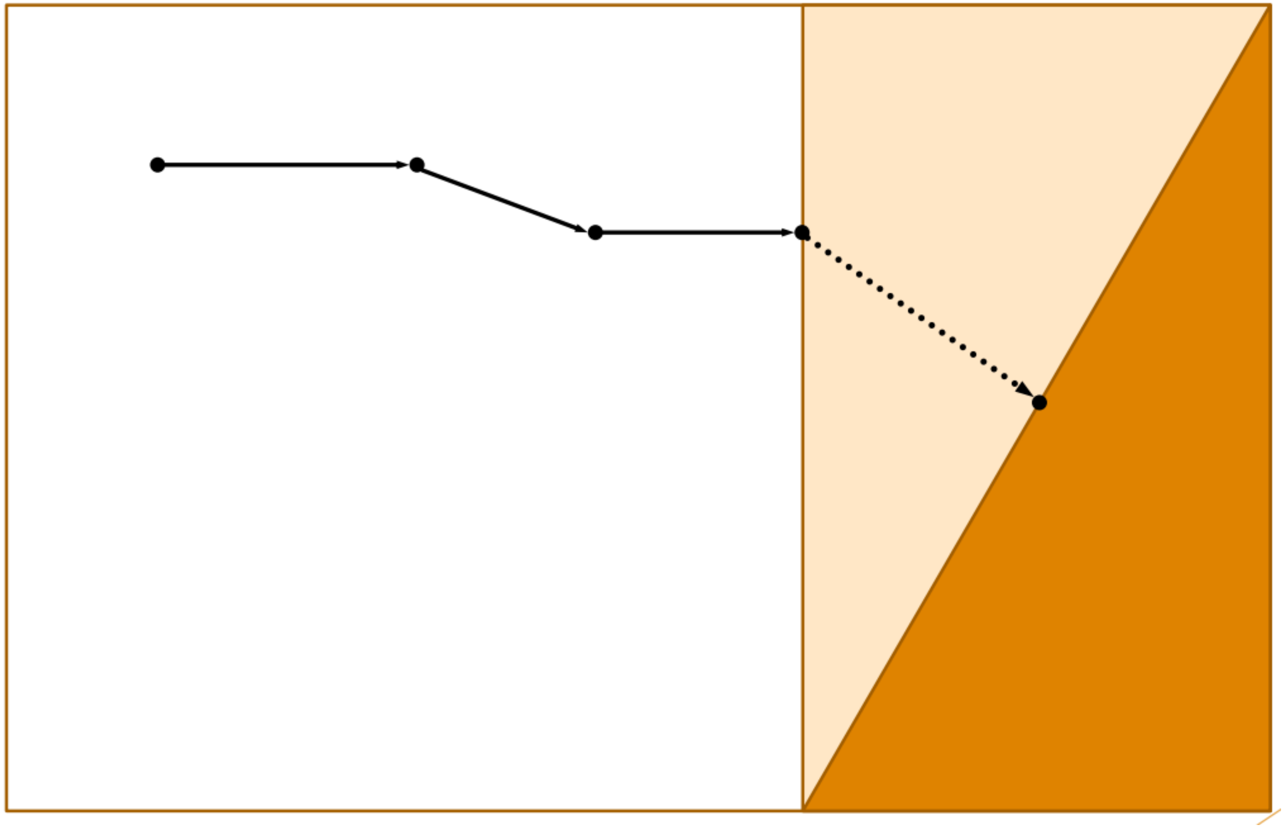
2. The Chasing Convex Bodies Problem



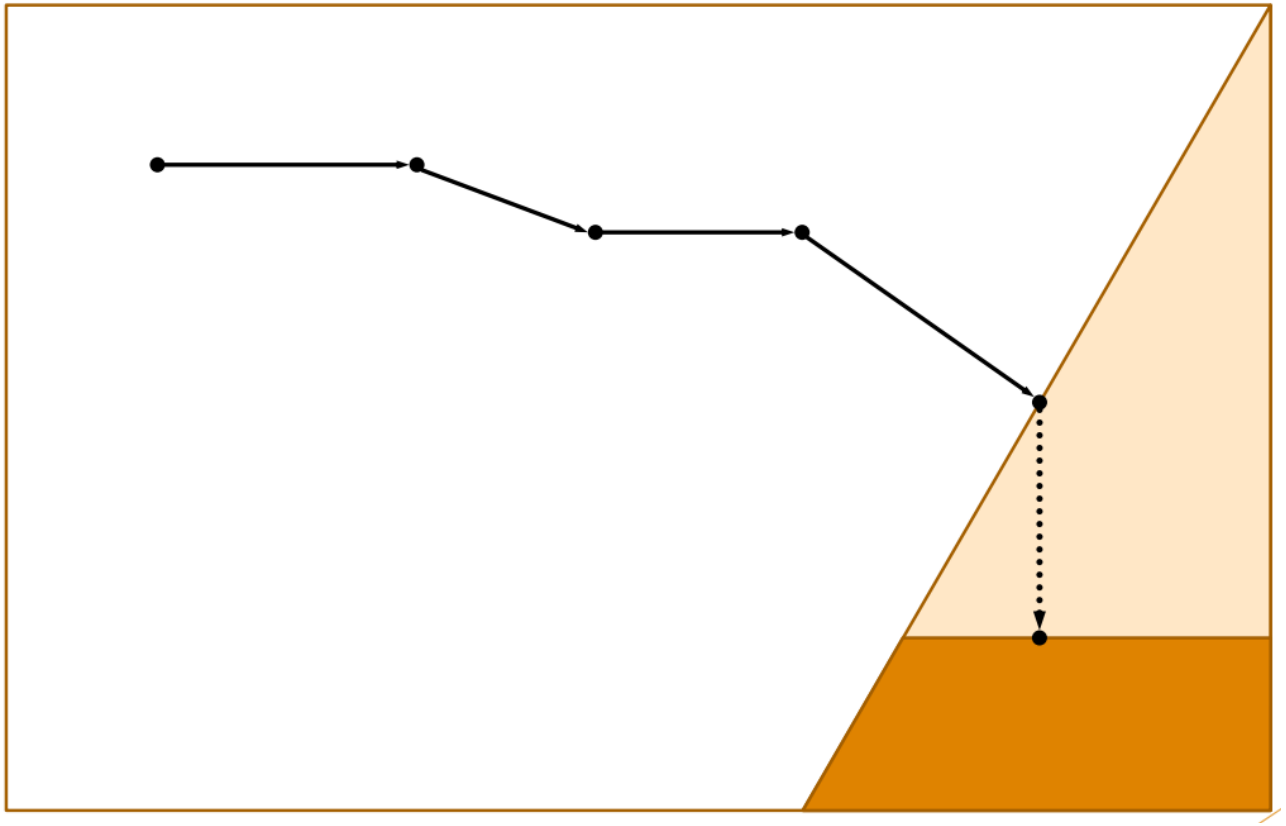
2. The Chasing Convex Bodies Problem



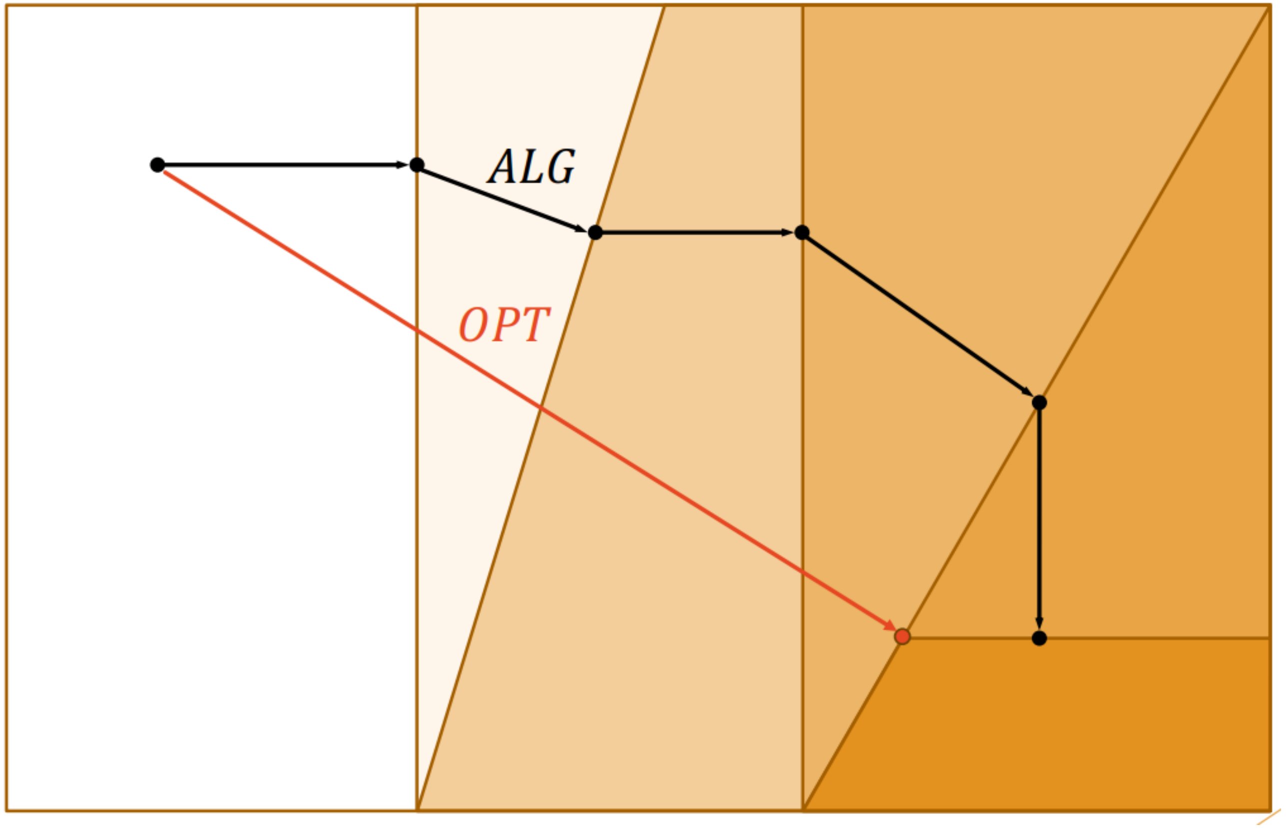
2. The Chasing Convex Bodies Problem



2. The Chasing Convex Bodies Problem



2. The Chasing Convex Bodies Problem



3. Naive Centroid Approach

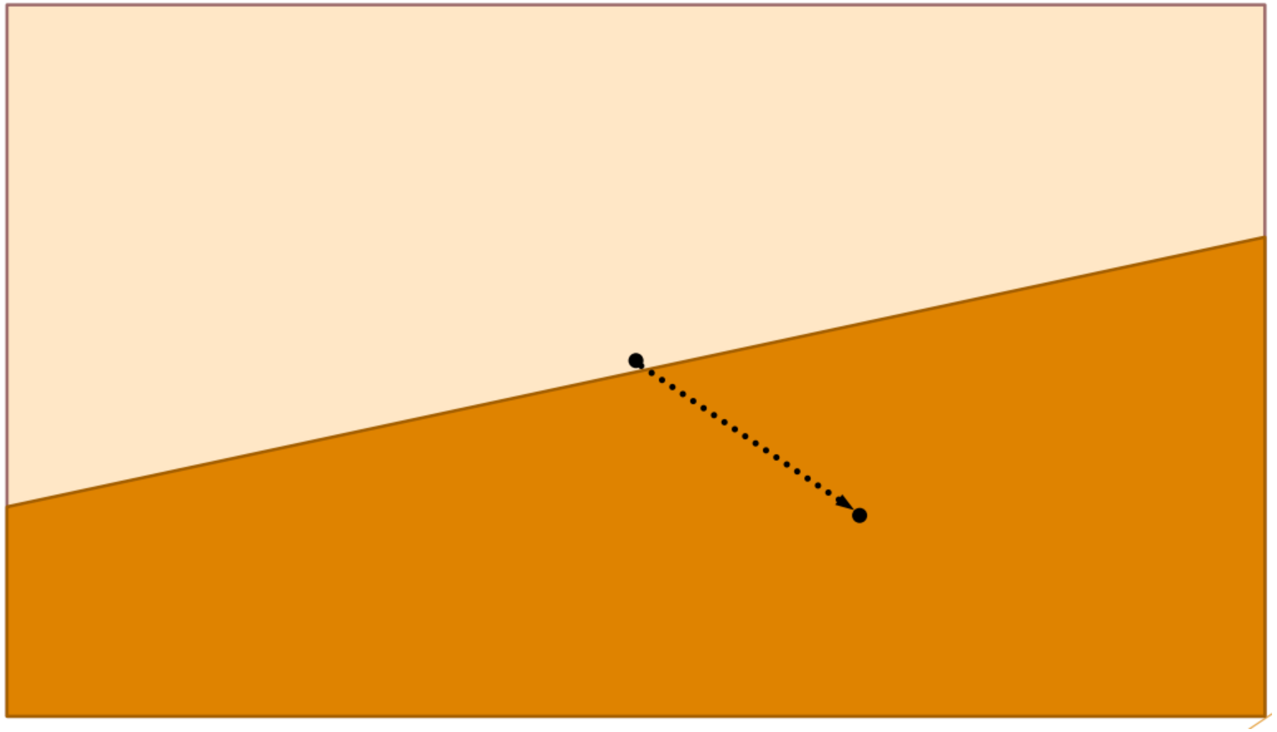
K_t a convex body

Centroid(K_t):=$\frac{1}{Vol(K_T)}\int_{K_T} x\, dx$

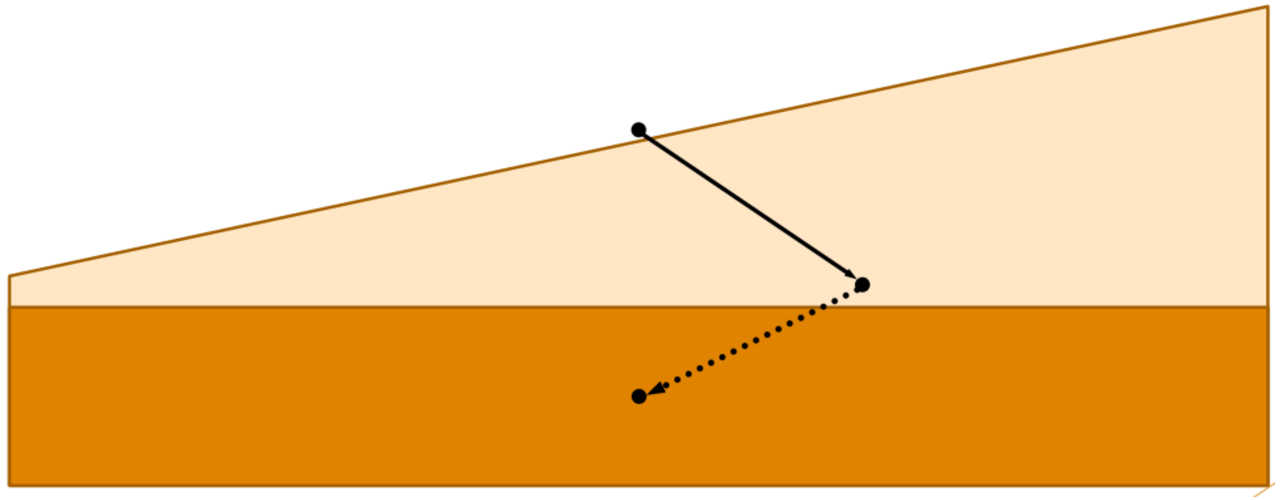
Algorithm: At each timestep t move to the centroid(K_t)

This approach is intuitive and ensures a move to a central point within the convex body, but it provides an unbounded competitive ratio,

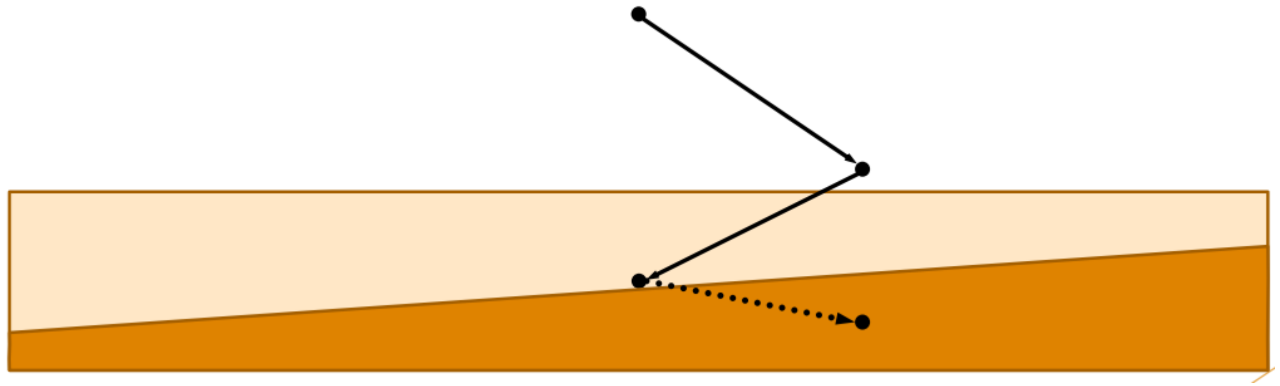
3. Naive Centroid Approach



3. Naive Centroid Approach



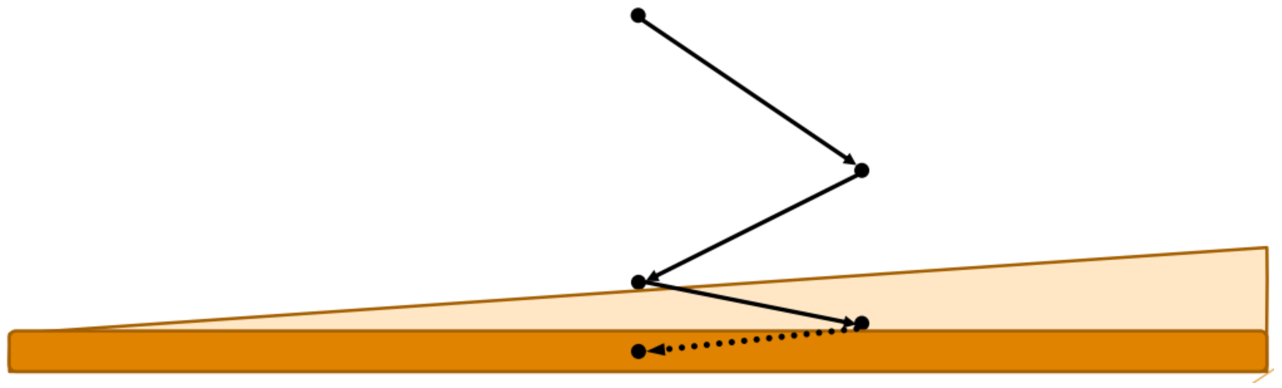
3. Naive Centroid Approach



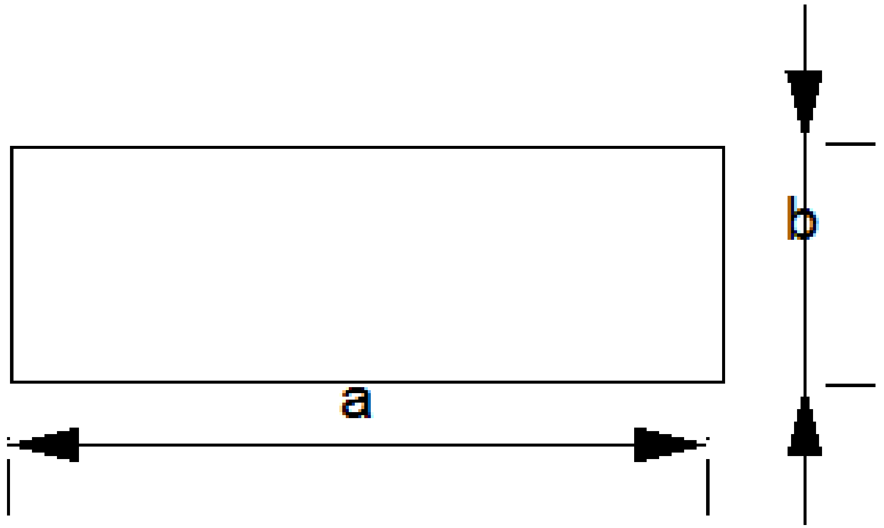
3. Naive Centroid Approach

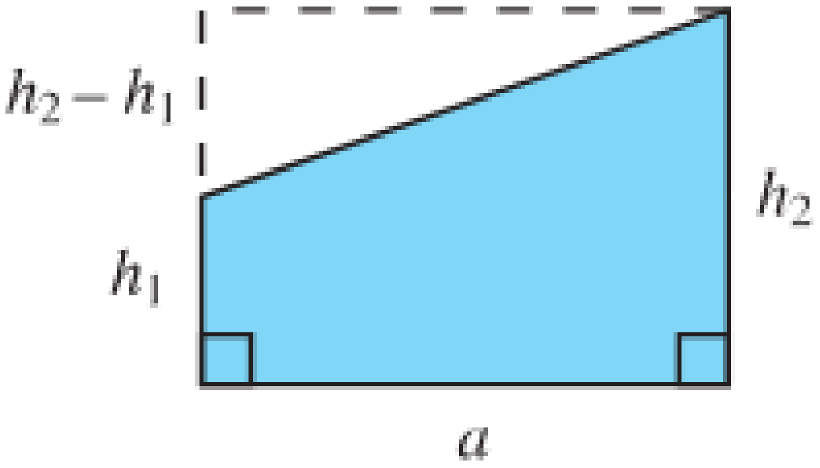
Not competitive
☹

Diameter constant



3. Naive Centroid Approach





$$x_{centroid} = \frac{a}{2}$$

$$x_{centroid} = \frac{a(h_1+2h_2)}{3(h_1+h_2)}$$

3. Naive Centroid Approach

- Lines: $x = -1,\ x = -1,\ y = 0.$
- For $t = 1, 3, 5 \dots :\ y = (\frac{1}{2})^t$, rectangle, x-coordinate of centroid is $\mathbf{0}$
- For $t = 2, 4, 6, \dots :\ y = \left(\frac{1}{2}\right)^{t+1}(x + 3)$, trapezium, x-coordinate of centroid is $\frac{1}{9}$
- The x-coordinate of the centroid will oscillate between $\mathbf{0}$ and $\frac{1}{9}$ showing that the total distance traveled by the algorithm can be made arbitrarily large

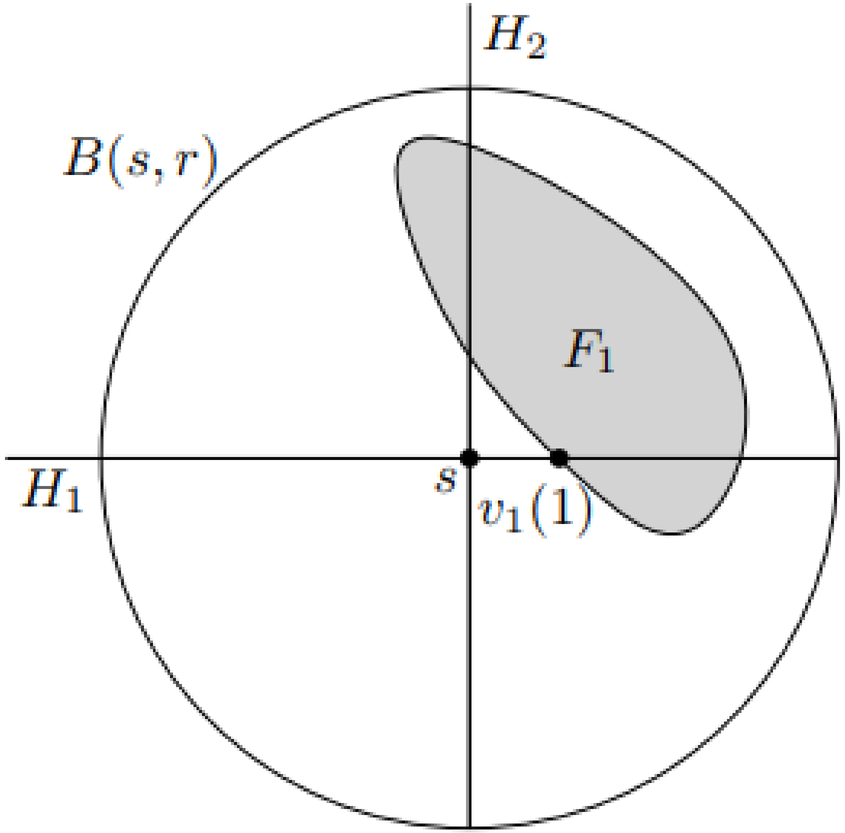
4. Recursive Greedy Method

- The first $f(d)$-competitive algorithm for Chasing Nested Convex Bodies
- Competitive ratio: $O(6^d(d!)^2)$
- Tries to tackle the problem of the “naïve centroid” approach in which the algorithm was moving back and forth
- It is called recursive because it solves “subproblems” a lower dimension

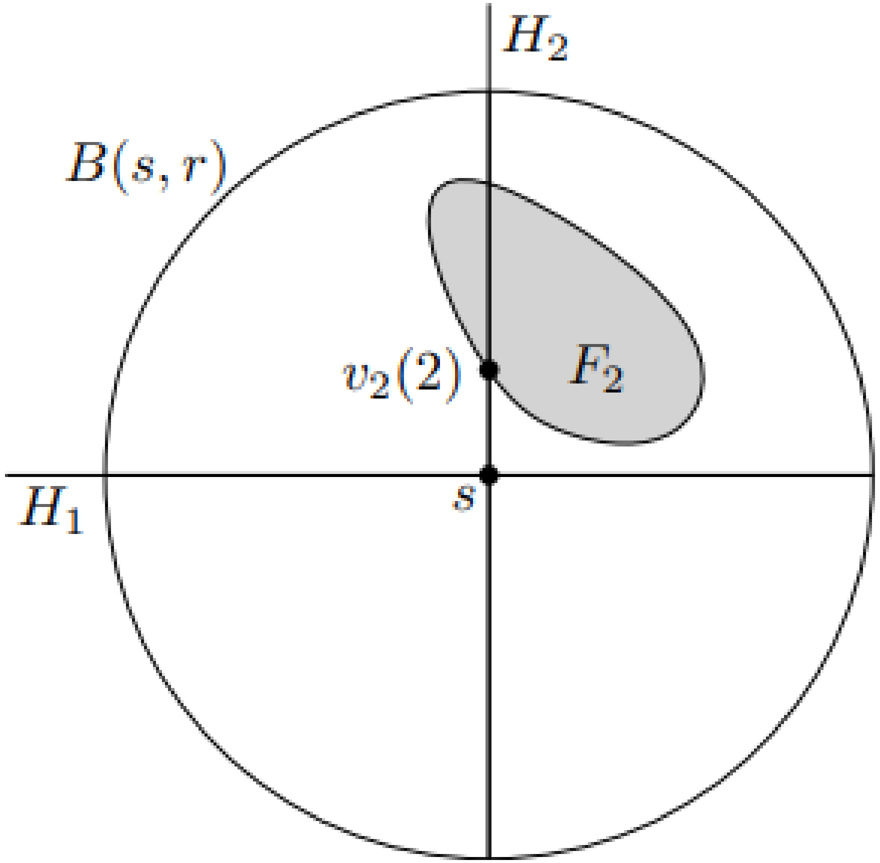
4. Recursive Greedy Method

- We will call the algorithm $Chase_d$.
- The algorithm runs in phases.
- Consider a phase that starts at center s and radius r.
- Let $H_1, \ldots, H_d$ denote the axis-aligned hyperplanes passing through s.
- **Hyperplane step**: When a request K_i arrives find the smallest index k such that H_k intersects K_i and run $Chase_{d-1}$ at H_k, until a request K_j arrives that cannot be served by H_k. Then try the hyperplanes with index $k' > k$ to serve the request K_j.
- **Recentering step:** If a request K_i arrives and none of the remaining hyperplanes intersects it, compute the smallest ball $B(s', r')$ that contains K_i, move to s' and start a new phase.

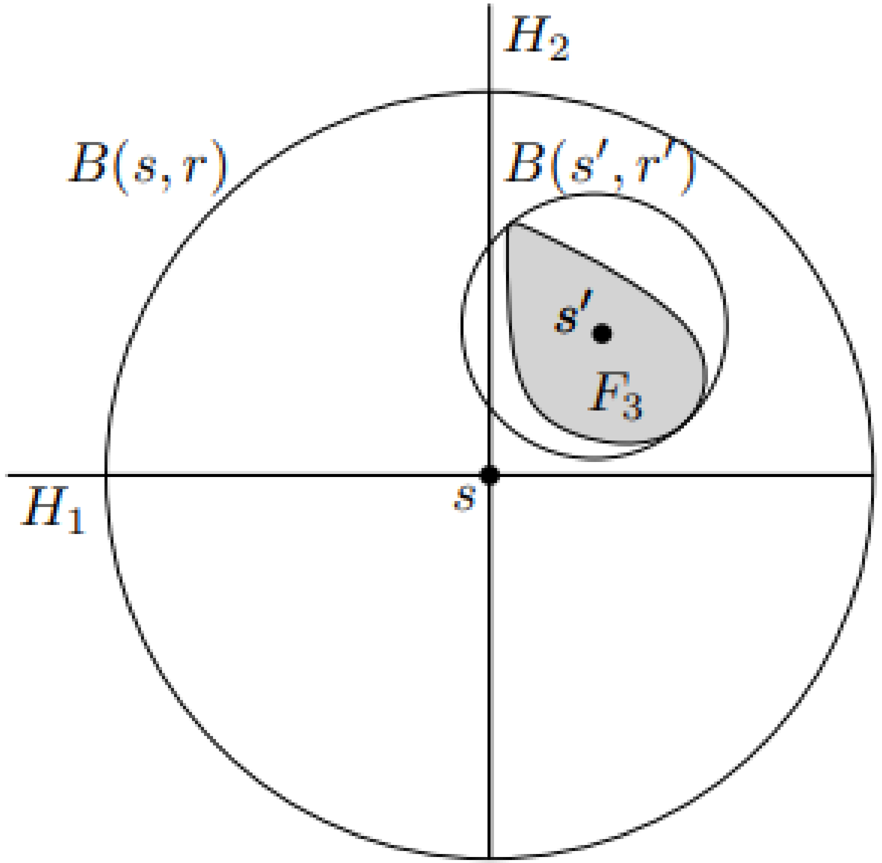
4. Recursive Greedy Method



4. Recursive Greedy Method



4. Recursive Greedy Method



4. Recursive Greedy Method

- Let the algorithm has cost $g(d) \cdot r$, and that we are at phase j
- Movement due to hyperplane steps is $g(d-1) \cdot r_j$, because on each of the d hyperplanes H_k, we move at most $g(d-1) \cdot r_j$ by following $Chase_{d-1}$.
- Movement due to switching hyperplanes. We switch hyperplanes at most $d-1$ times, so this is at most $(d-1) \cdot 2r_j$.
- Movement due to recentering is at most $2r_j$.
- Thus, the total distance traveled in phase j is at most:

$$d(g(d-1)+2) \cdot r_j \leq 3d \cdot g(d-1) \cdot r_j, \text{ because } g(d-1) \geq 1$$

4. Recursive Greedy Method

- It is proven that the radii of the enclosing balls decrease geometrically across phases: $\mathrm{r_j} \leq r_{j-1}\left(1-\frac{1}{d}\right)^{1/2}$ and we have $r_1 = r$.
- $3d \cdot g(d-1) \cdot \frac{r}{\left(1-\frac{1}{d}\right)^{\frac{1}{2}}} \leq 3d \cdot g(d-1) \cdot 2dr = 6d^2 \cdot g(d-1) \cdot r$, by using Taylor series
- By solving $g(d) \cdot r = 6d^2 \cdot g(d-1) \cdot r$ with base case $g(1) = 1$ we get $g(d) = 6^d (d!)^2$

5. Steiner Point Method

- Imagine you are standing at the origin of a coordinate system and looking in various directions $\theta \in \mathbb{S}^{d-1}$
- For each direction θ, find the point on the boundary of K that is the furthest in that direction.
- Now, imagine you could average out all these furthest points by considering all possible directions.
- The result of this averaging gives you the Steiner point $s(K)$

5. Steiner Point Method

For a convex body $K \subset \mathbb{R}^d$, its Steiner Point $s(K)$ is defined in these two equivalent ways:

1. For any direction $\theta \in \mathbb{S}^{d-1} \subset \mathbb{R}^d$, let $f_K(\theta) = argmax_{x \in K}(\theta \cdot x)$ be the extremal point in K in direction θ. Then compute the average of this extremal point for a random direction: $s(K) = \int_{\theta \in \mathbb{S}^{d-1}} f_K(\theta) d\theta$
2. For any direction $\theta \in \mathbb{S}^{d-1} \subset \mathbb{R}^d$, let $h_K(\theta) = max_{x \in K}(\theta \cdot x)$ be the support function for K in direction θ, and compute: $s(K) = d \int_{\theta \in \mathbb{S}^{d-1}} h_K(\theta) \theta d\theta$

5. Steiner Point Method

- Algorithm: At each timestep t move to the $s(K_t)$
- It is a memoryless Algorithm
- It works only for Euclidean Spaces
- Competitive ratio: $O(\min(d, \sqrt{d\ logT}))$, where T is the number of timesteps
- We will prove the $O(d)$ part of the competitive ratio

5. Steiner Point Method

Proof: Let $B_1 \supseteq K_1 \supseteq \cdots \supseteq K_T \in \mathbb{R}^d$ (by scaling) a sequence of convex bodies

Total cost of the movement is:

$$\sum_{t=1}^{T-1} ||s(K_t) - s(K_{t+1})||_2 = \sum_{t=1}^{T-1} d| \int_{\theta \in \mathbb{S}^{d-1}} \big(h_{K_t}(\theta) - h_{K_{t+1}}(\theta)\big)\theta \, d\theta| \text{ (by definition)}$$

$$\leq d \sum_{t=1}^{T-1} | \int_{\theta \in \mathbb{S}^{d-1}} \left(h_{K_t}(\theta) - h_{K_{t+1}}(\theta)\right) \, d\theta| \text{ (}\theta\text{ is unit vector)}$$

$$= d \int_{\theta \in \mathbb{S}^{d-1}} \sum_{t=1}^{T-1} |h_{K_t}(\theta) - h_{K_{t+1}}(\theta)| d\theta$$

5. Steiner Point Method

Now, as $K_t \supseteq K_{t+1}$, we have that $h_{K_t}(\theta) \geq h_{K_{t+1}}(\theta)$ and so:

$$d \int_{\theta \in \mathbb{S}^{d-1}} \sum_{t=1}^{T-1} \left| h_{K_t}(\theta) - h_{K_{t+1}}(\theta) \right| = d \int_{\theta \in \mathbb{S}^{d-1}} \sum_{t=1}^{T-1} (h_{K_t}(\theta) - h_{K_{t+1}}(\theta)) d\theta$$

$$= d(\int_{\theta \in \mathbb{S}^{d-1}} h_{K_1}(\theta) d\theta - \int_{\theta \in \mathbb{S}^{d-1}} h_{K_T}(\theta) d\theta) \text{ (telescoping sum)}$$

The first integral is at most one because $K_1 \subseteq B_1$. As $h_K(\theta) + h_K(-\theta) \geq 0$ for any convex body K, the second integral is non-negative. So we conclude that

$$\sum_{t=1}^{T-1} ||s(K_t) - s(K_{t+1})||_2 \leq d$$

References

- Nikhil Bansal, Martin Böhm, Marek Eliáš, Grigorios Koumoutsos, Seeun William Umboh (2017). *Nested Convex Bodies are Chaseable*. https://arxiv.org/abs/1707.05527
- Sébastien Bubeck, Bo'az Klartag, Yin Tat Lee, Yuanzhi Li, Mark Sellke (2018). *Chasing Nested Convex Bodies Nearly Optimally*. https://arxiv.org/abs/1811.00999