Vertical Versus Horizontal Poincare Inequalities

Assaf Naor

Courant Institute

Bi-Lipschitz distortion

(M, d_M) a metric space and $(X, \|\cdot\|_X)$ a Banach space.

$c_X(M) =$ the infimum over those $D \in (1, \infty]$ for which there exists $f : M \to X$ satisfying

$$\forall\, x, y \in M, \quad d_M(x,y) \leqslant \|f(x) - f(y)\|_X \leqslant D d_M(x,y).$$

$$M \overset{D}{\hookrightarrow} X.$$

The discrete Heisenberg group

- The group $\mathbb{H}$ generated by a, b subject to the relation stating that the commutator of a, b is in the center:

$$ac = ca \qquad \text{and} \qquad bc = cb$$

where

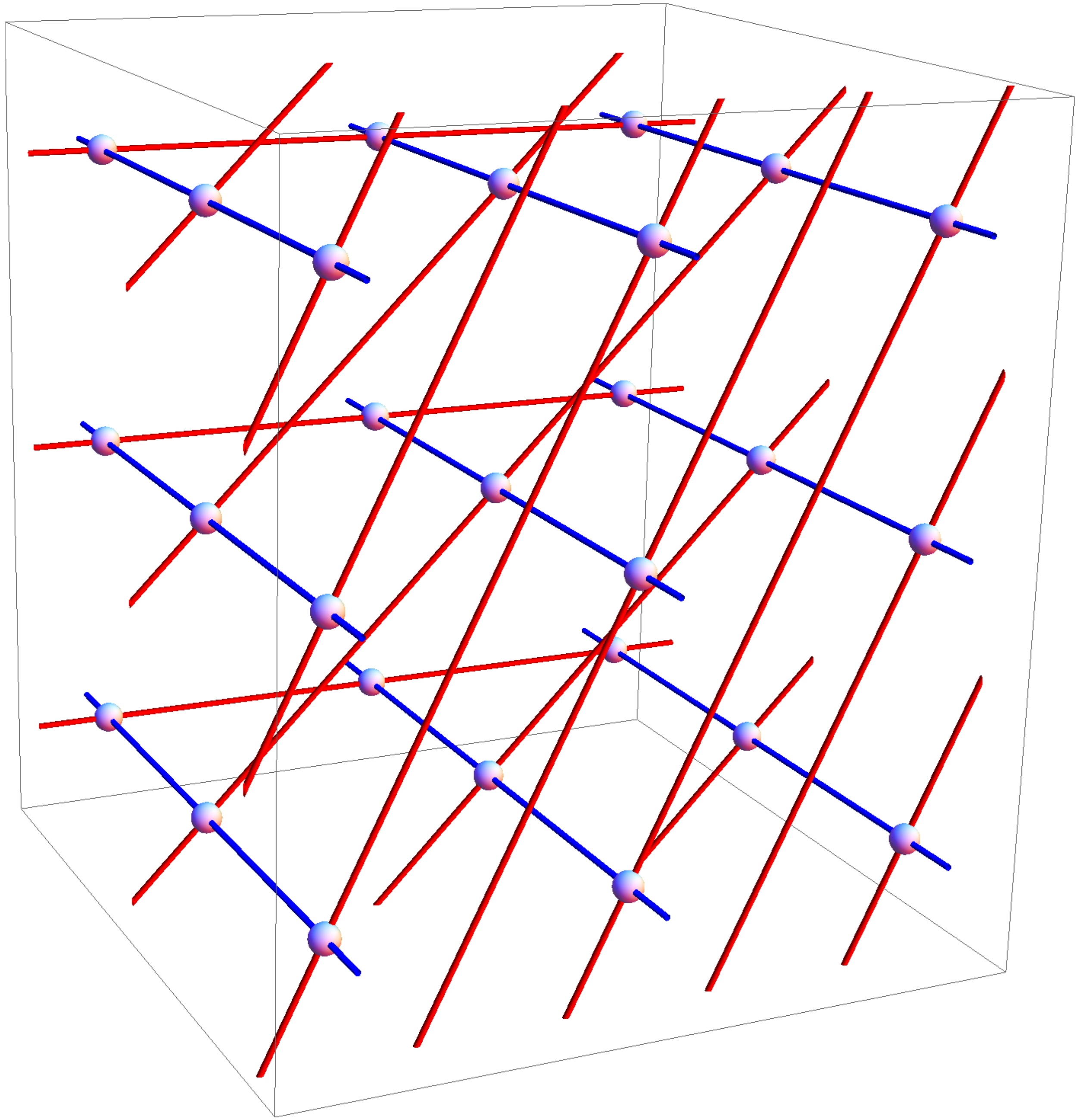
$$c = [a, b] = aba^{-1}b^{-1}$$

Concretely, $\mathbb{H} = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : x, y, z \in \mathbb{Z} \right\}$

$$a = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$c = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The left-invariant word metric on $\mathbb{H}$ corresponding to the generating set $\{a, a^{-1}, b, b^{-1}\}$ is denoted d_W.



Denote

$$\forall\, n \in \mathbb{N}, \qquad B_n = \{x \in \mathbb{H} :\ d_W(x, e_{\mathbb{H}}) \leqslant n\}$$

Basic facts:

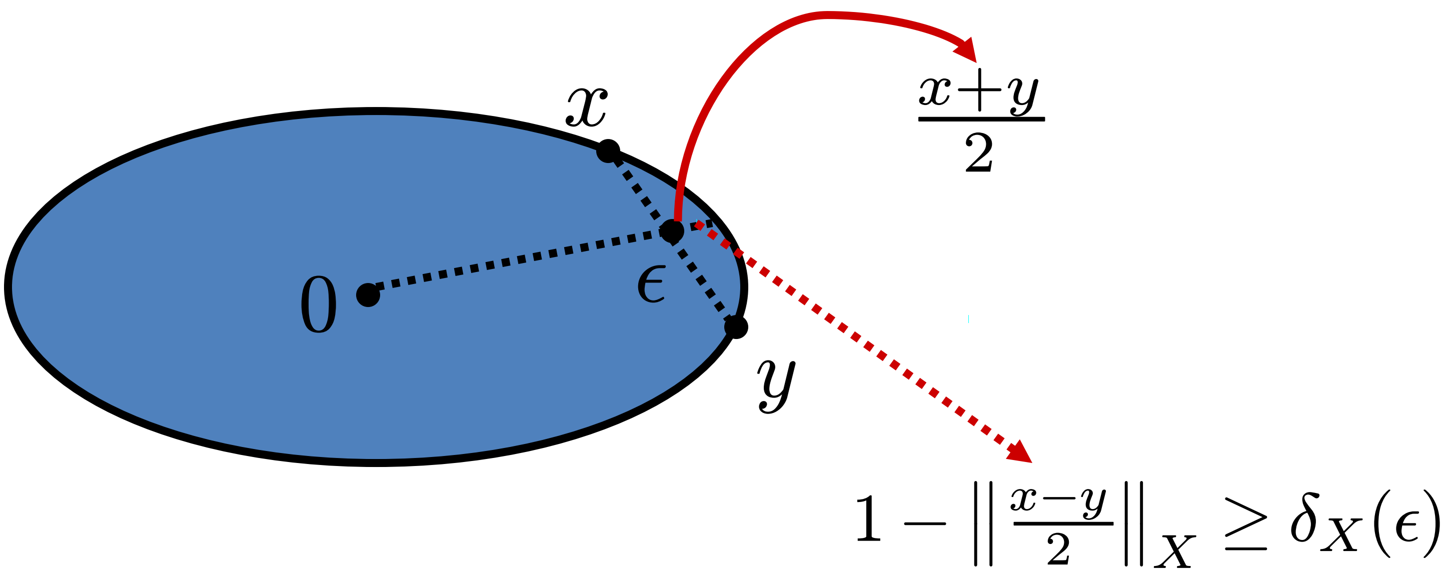
$$\forall\, k \in \mathbb{N}, \qquad d_W(c^k, e_{\mathbb{H}}) \asymp \sqrt{k}$$

$$\forall\, m \in \mathbb{N}, \qquad |B_m| \asymp m^4$$

Uniform convexity

The modulus of uniform convexity of $(X, \|\cdot\|_X)$:

$$\delta_X(\epsilon) = \inf\left\{1 - \left\|\frac{x+y}{2}\right\|_X \ : \ \|x\|_X = \|y\|_X = 1, \quad \|x-y\|_X = \epsilon\right\}$$



- *X* is uniformly convex if $\forall \epsilon \in (0,1), \quad \delta_X(\epsilon) > 0.$
- For $q \in [2,\infty)$, *X* is *q*-convex if it admits an equivalent norm with respect to which $\delta_X(\epsilon) \gtrsim \epsilon^q.$

<u>Theorem (Pisier, 1975)</u>. If *X* is uniformly convex then it is *q*-convex for some $q \in [2,\infty)$.

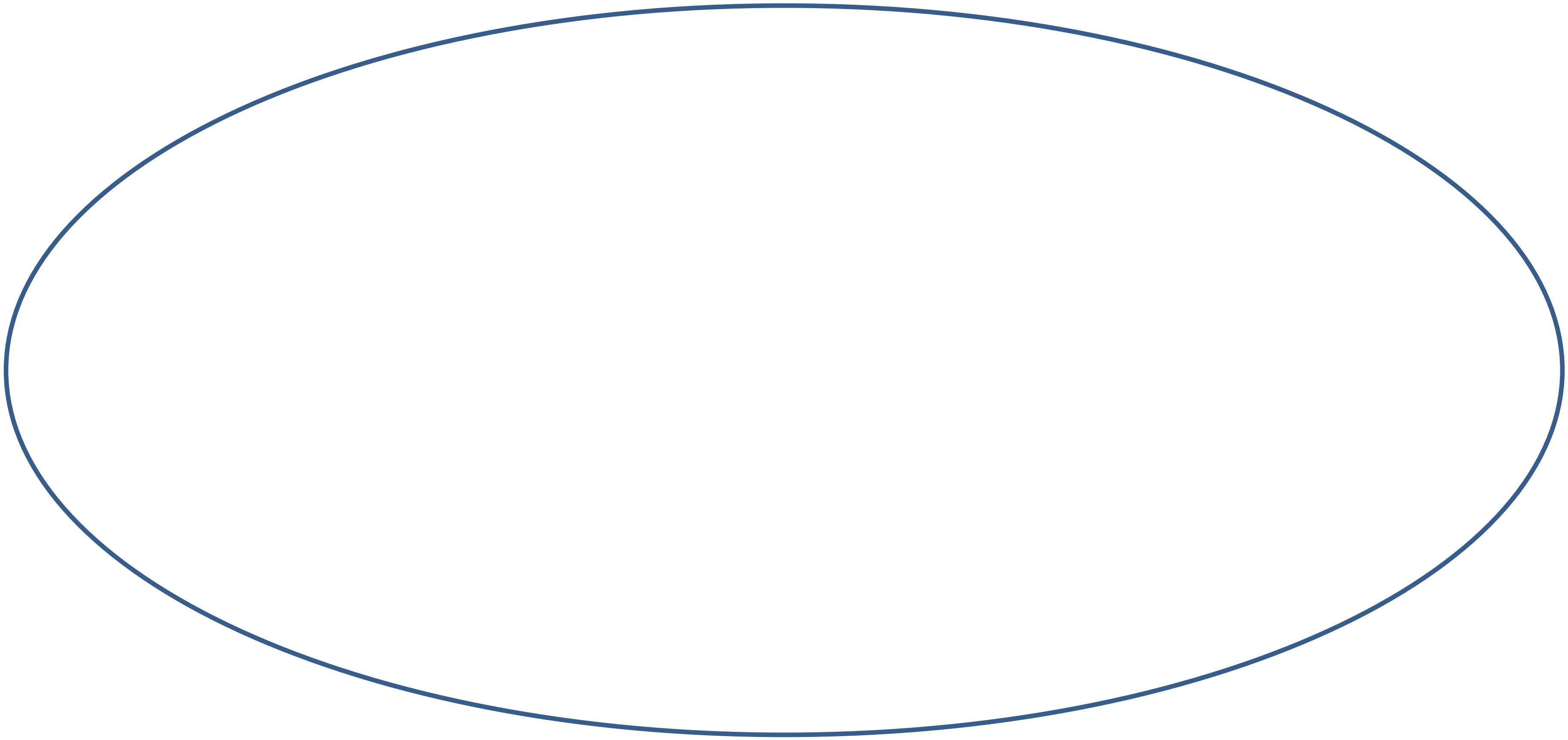
ℓ_p is $\max\{2,p\}$-convex for $p > 1$.

Mostow (1973), Pansu (1989), Semmes (1996)

Theorem. The metric space $(\mathbb{H}, d_W)$ does not admit a bi-Lipschitz embedding into $\mathbb{R}^n$ for any $n \in \mathbb{N}$.

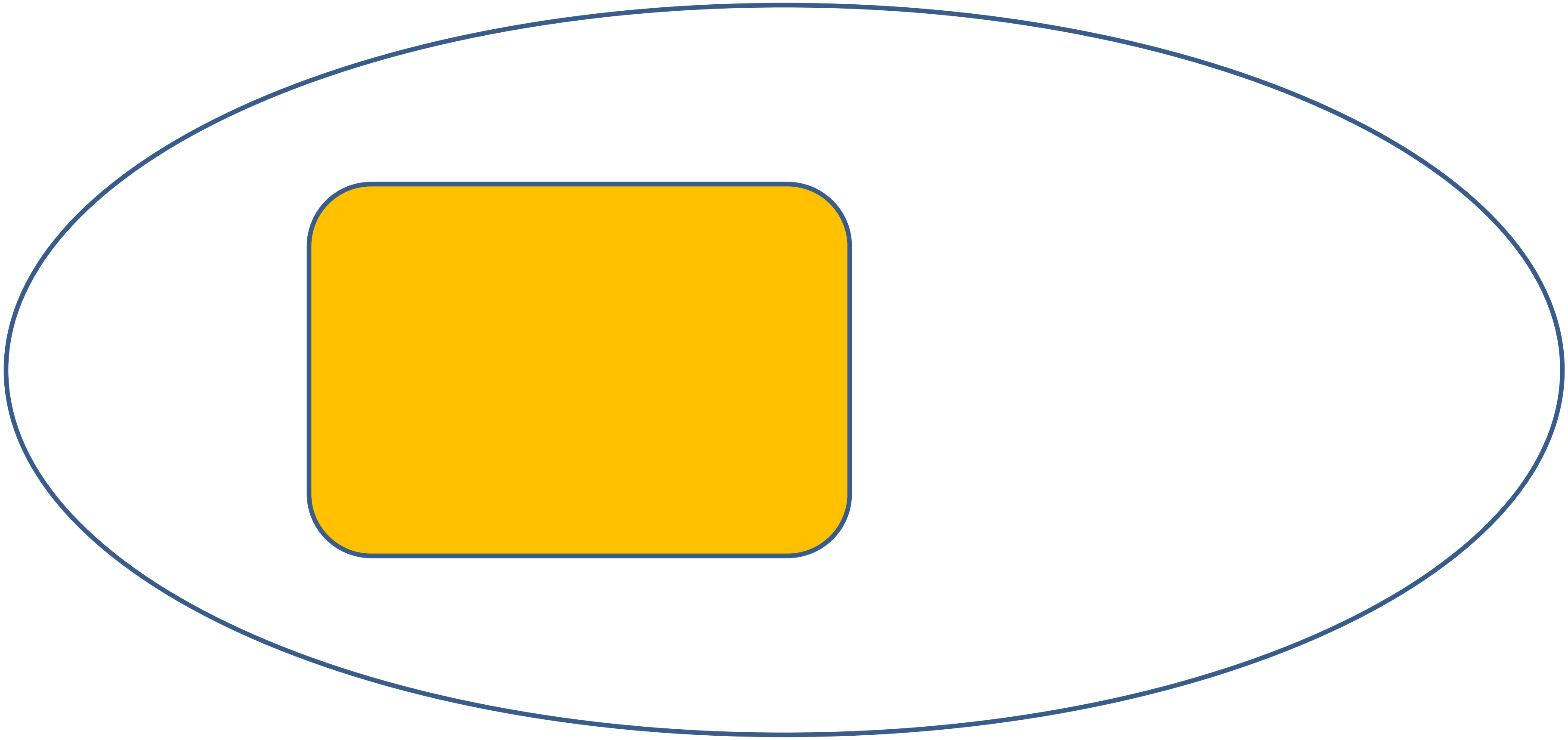
Assouad’s embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



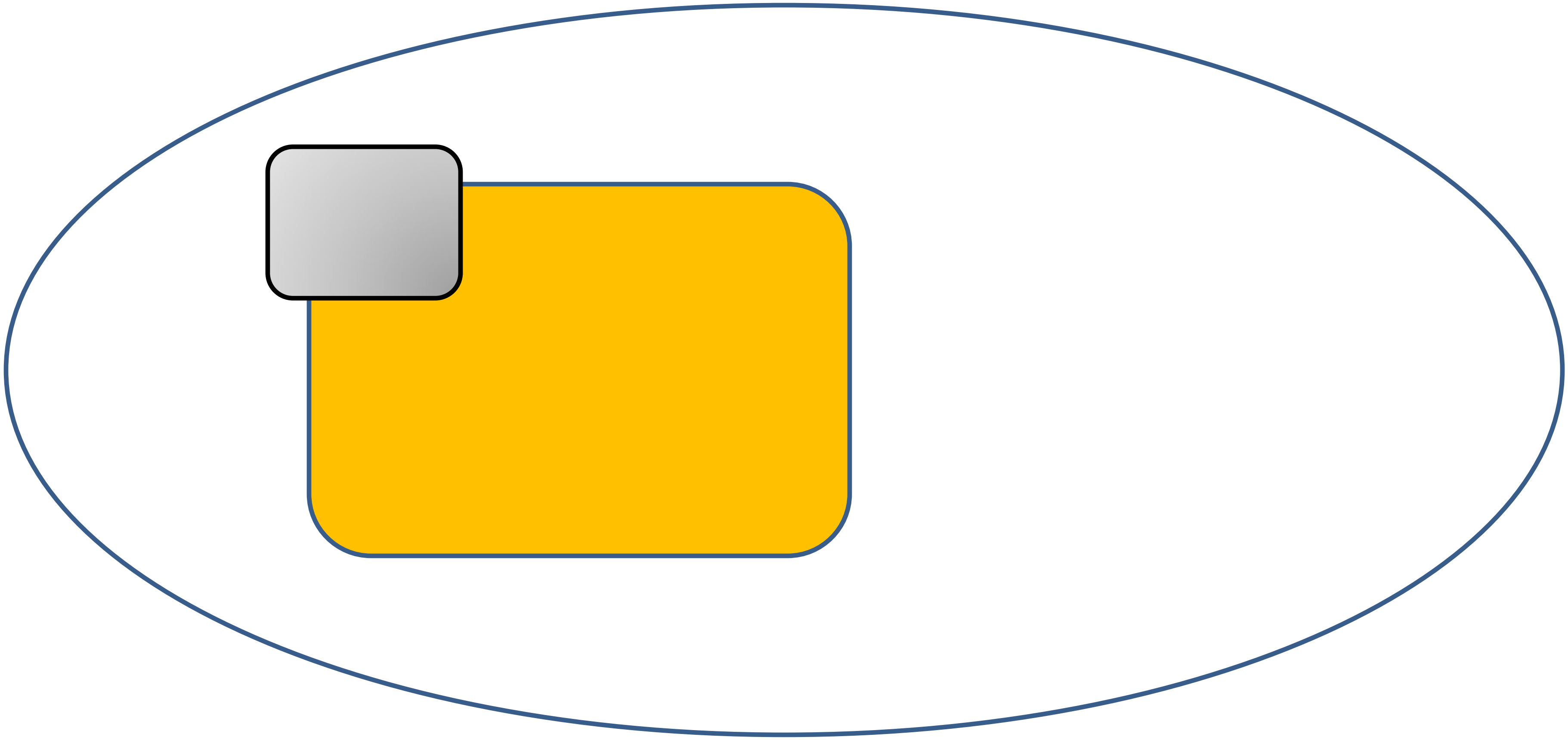
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



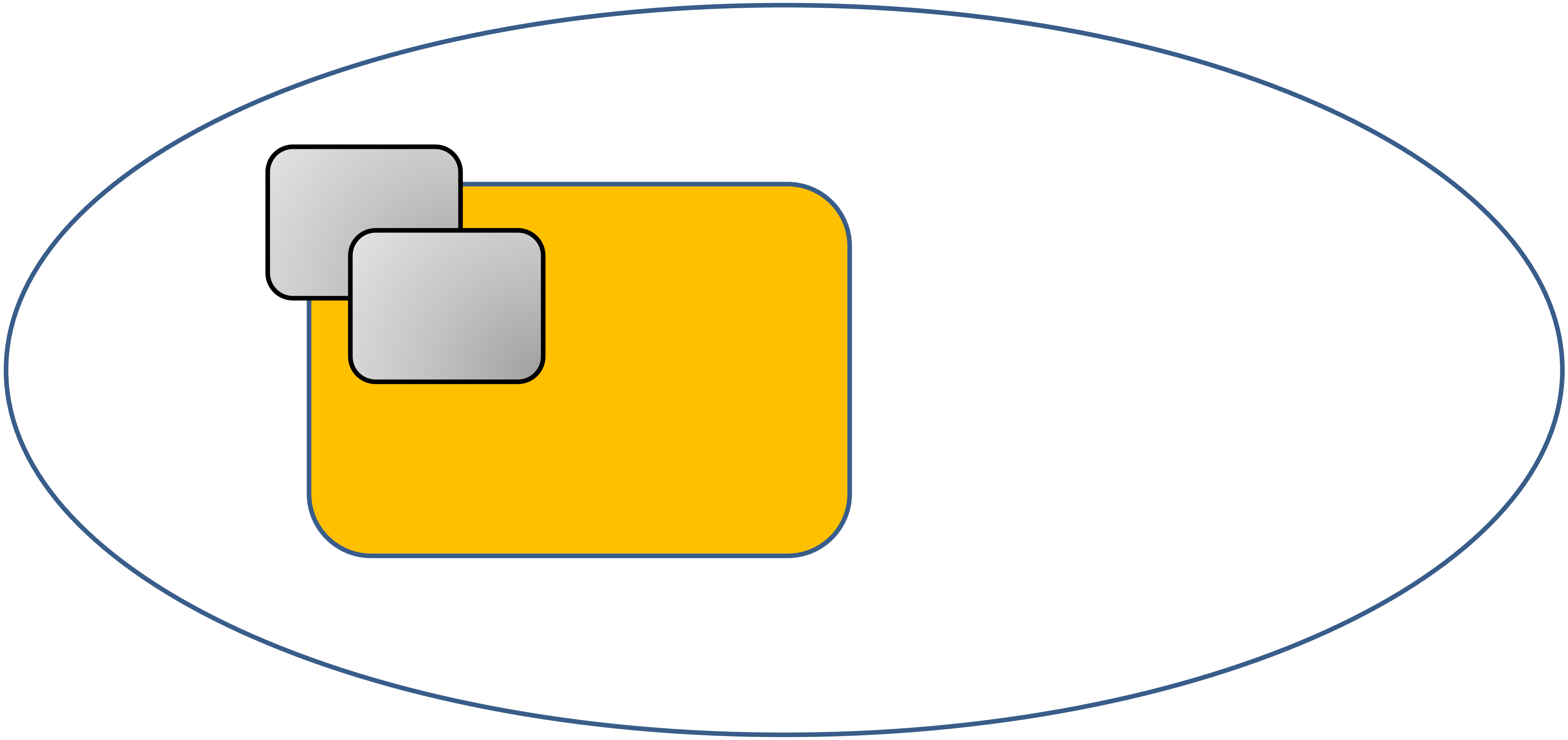
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



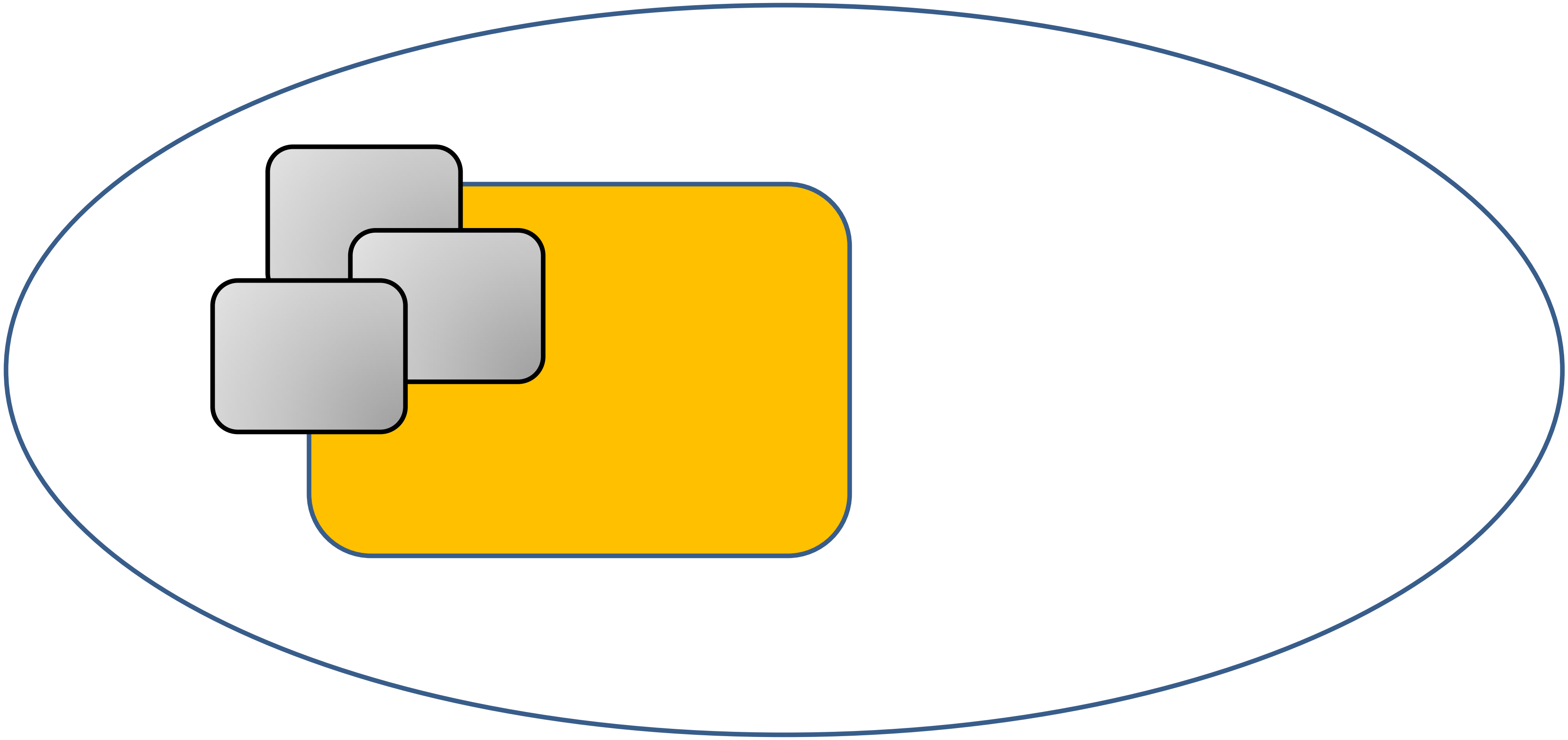
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



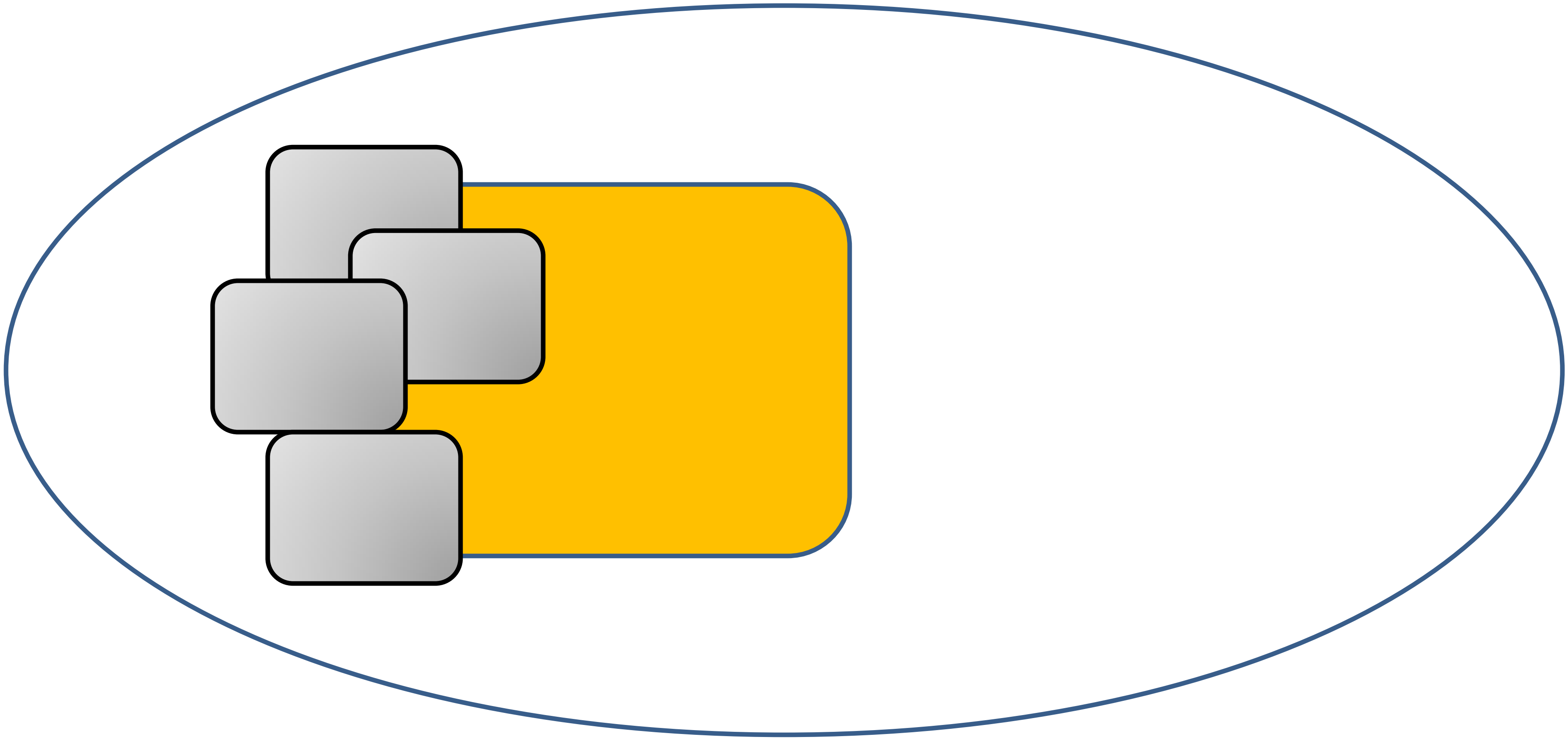
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



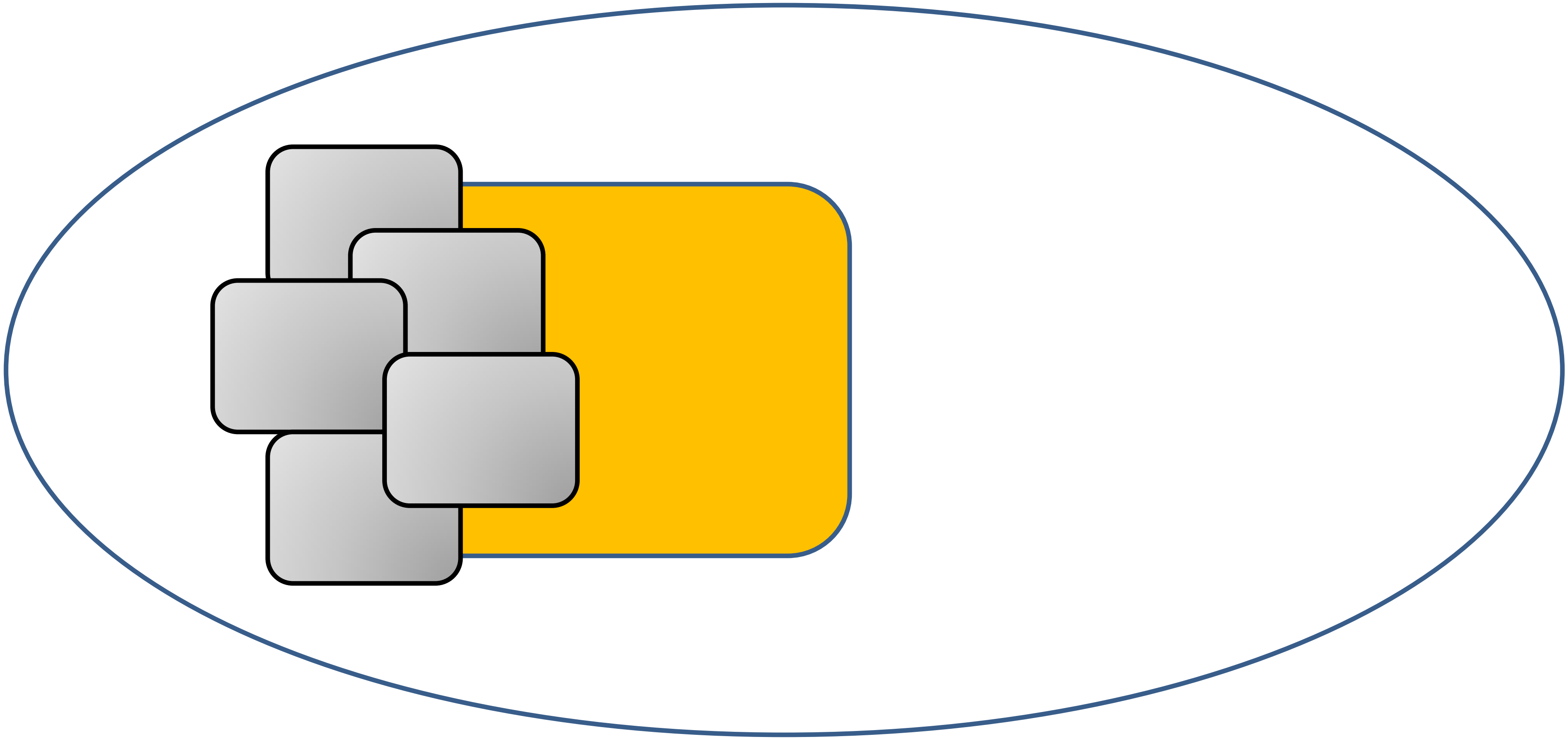
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



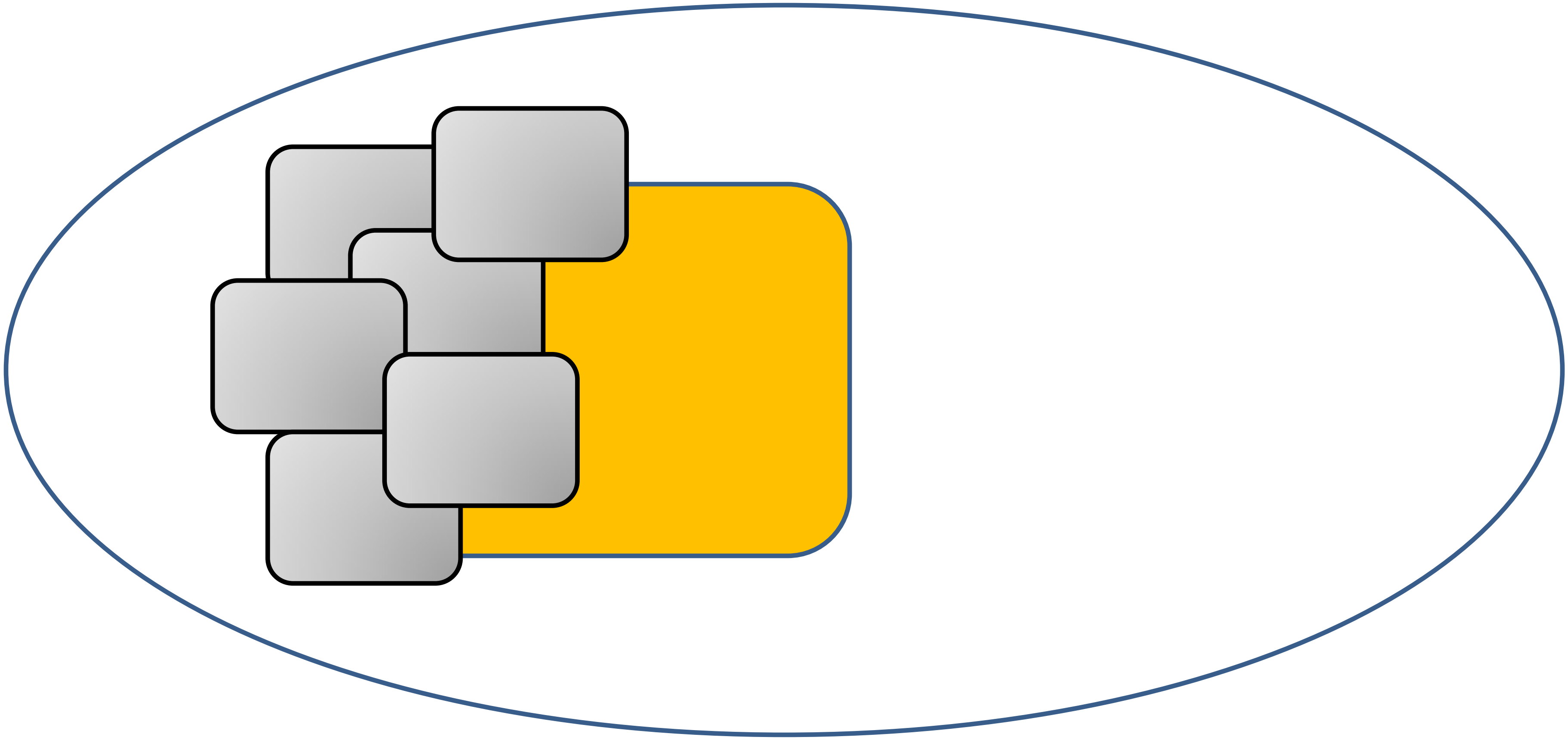
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



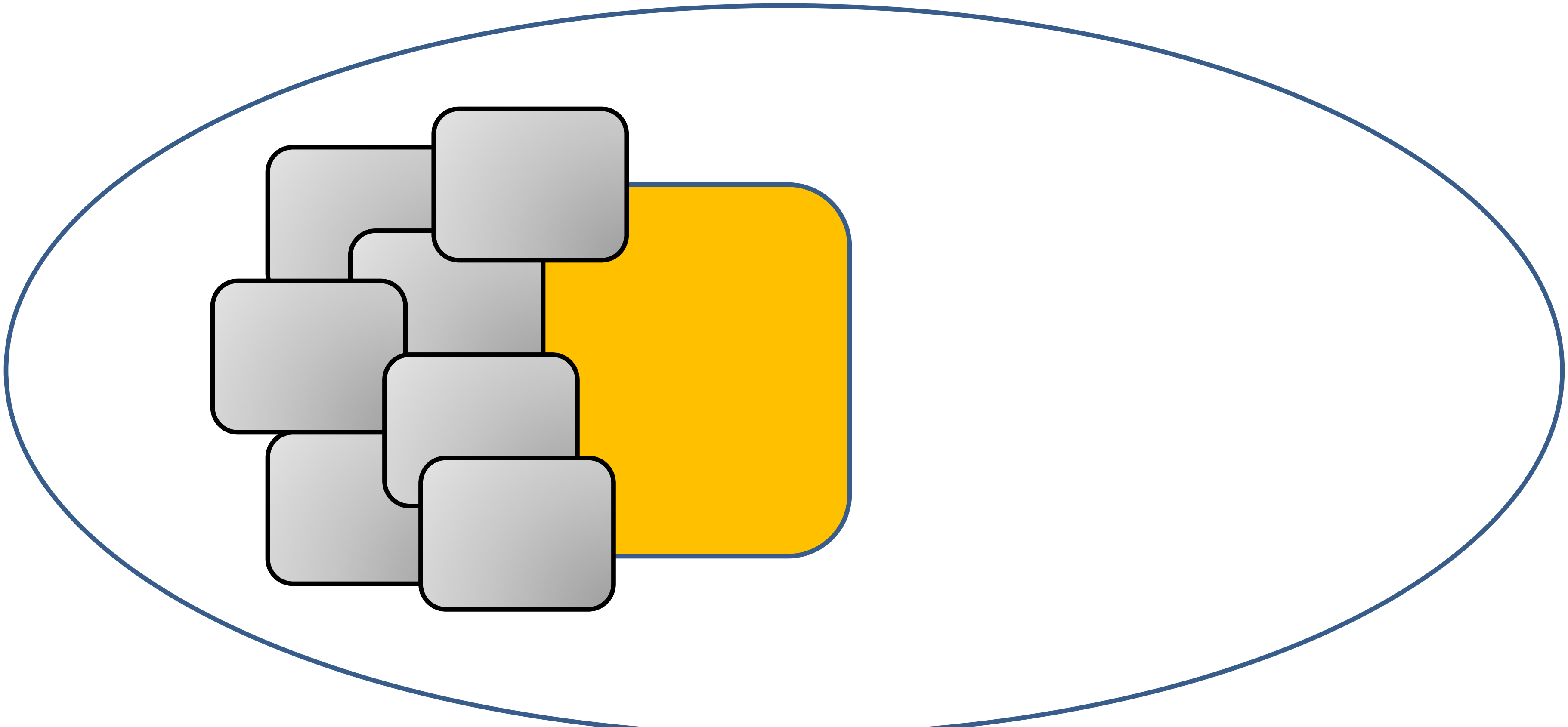
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



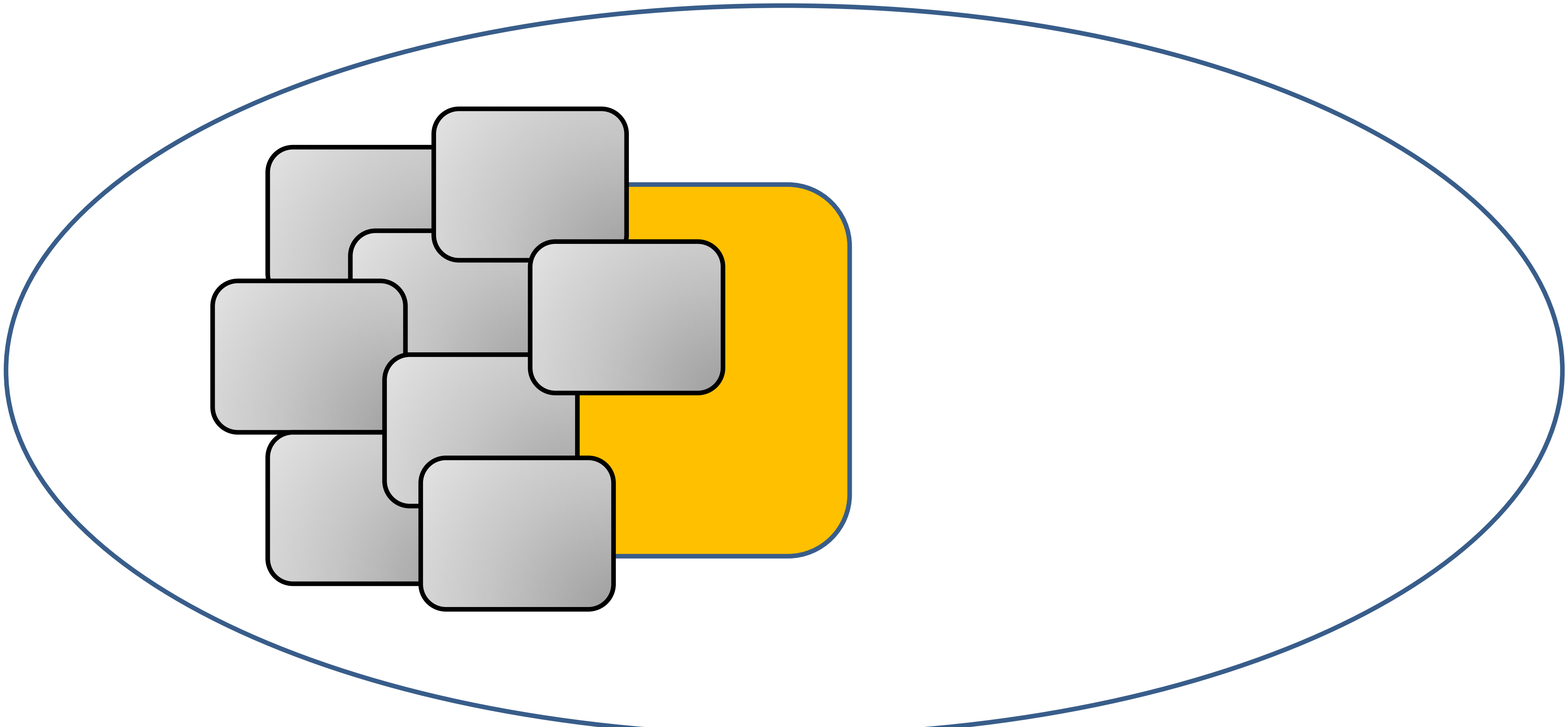
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



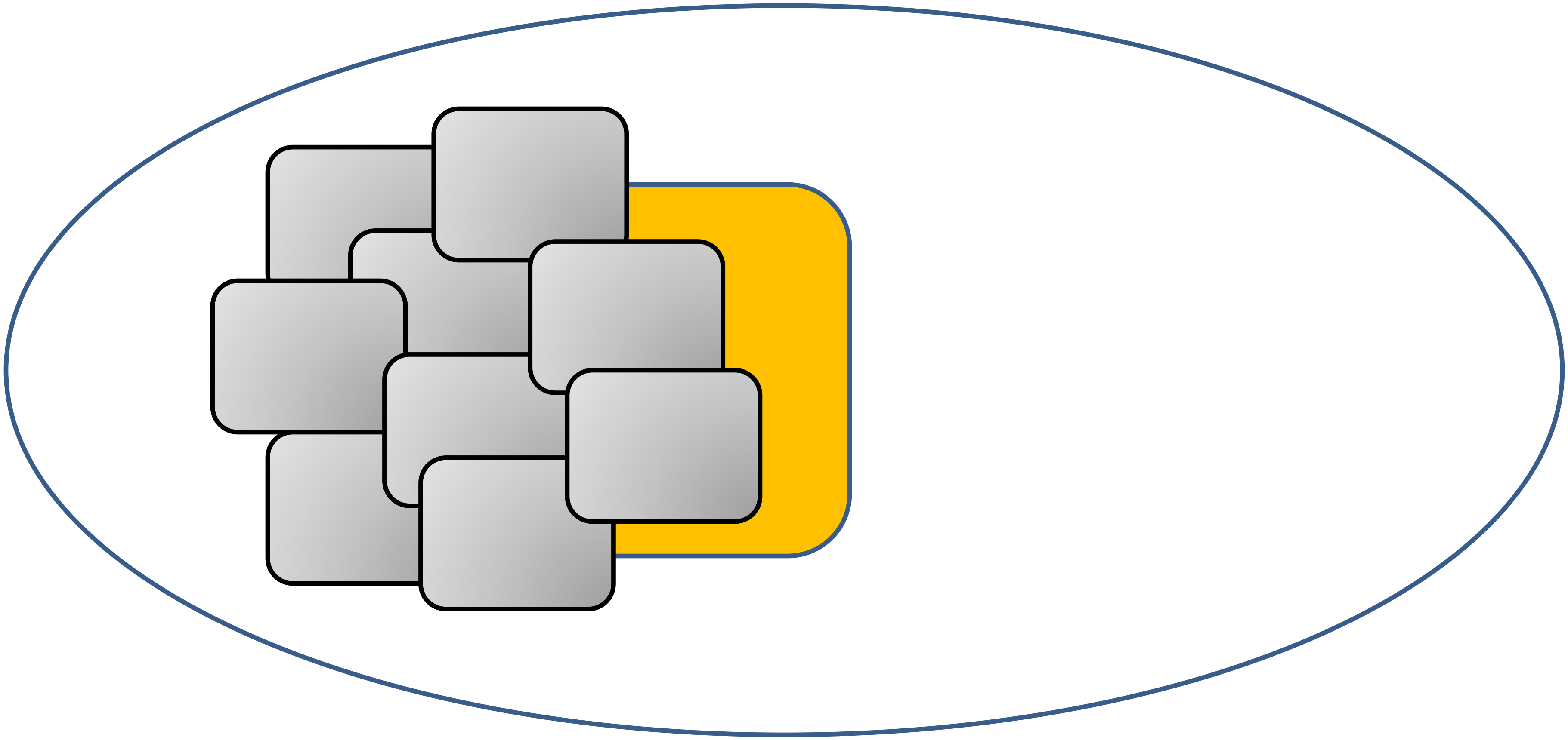
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



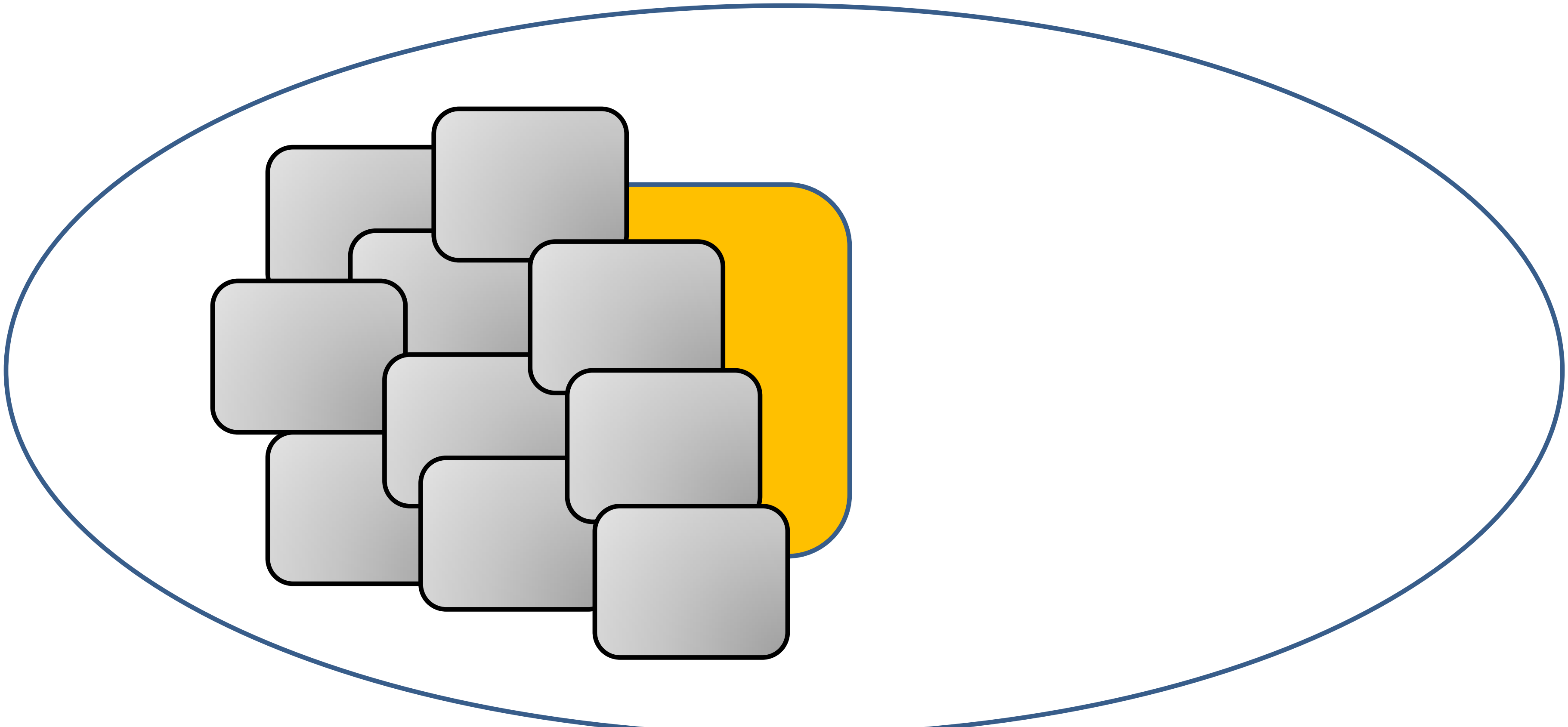
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



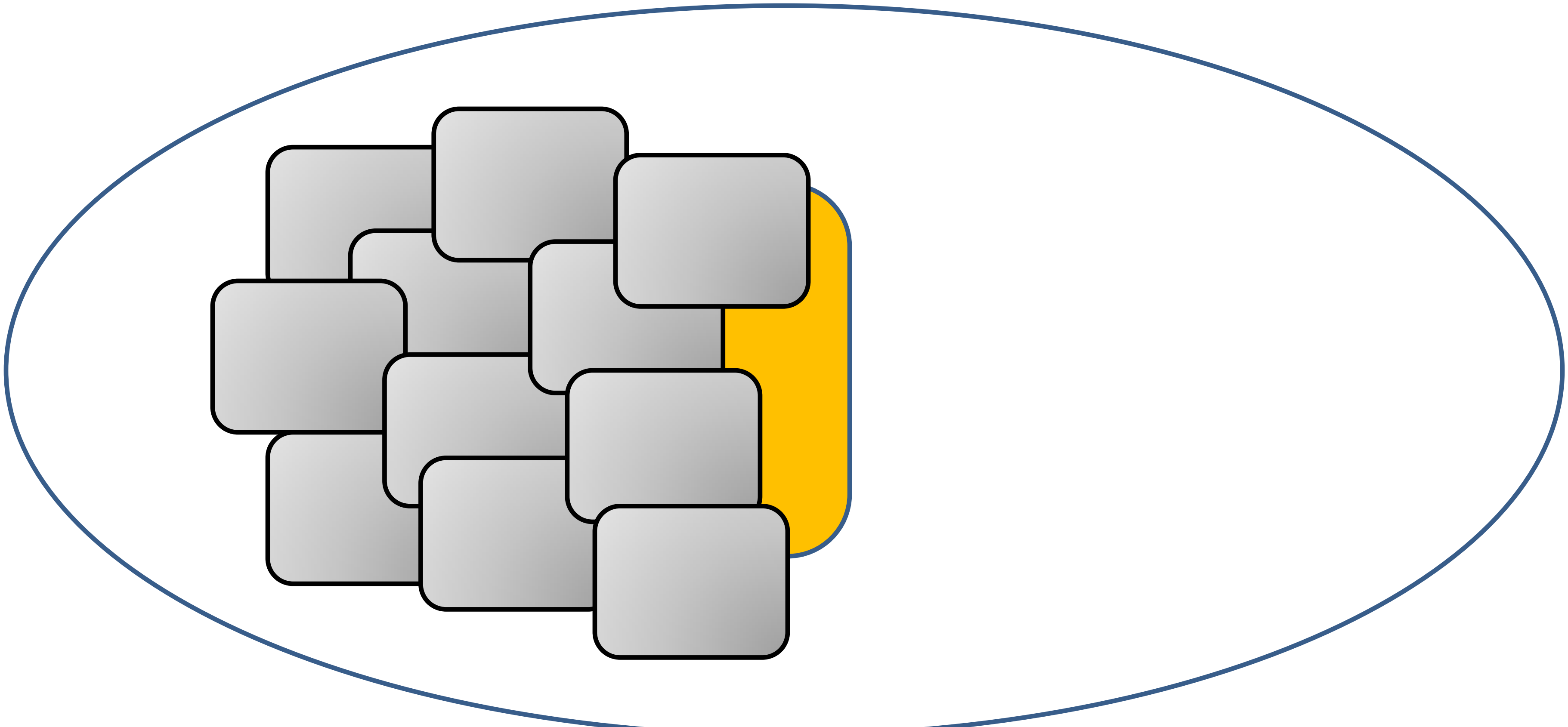
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



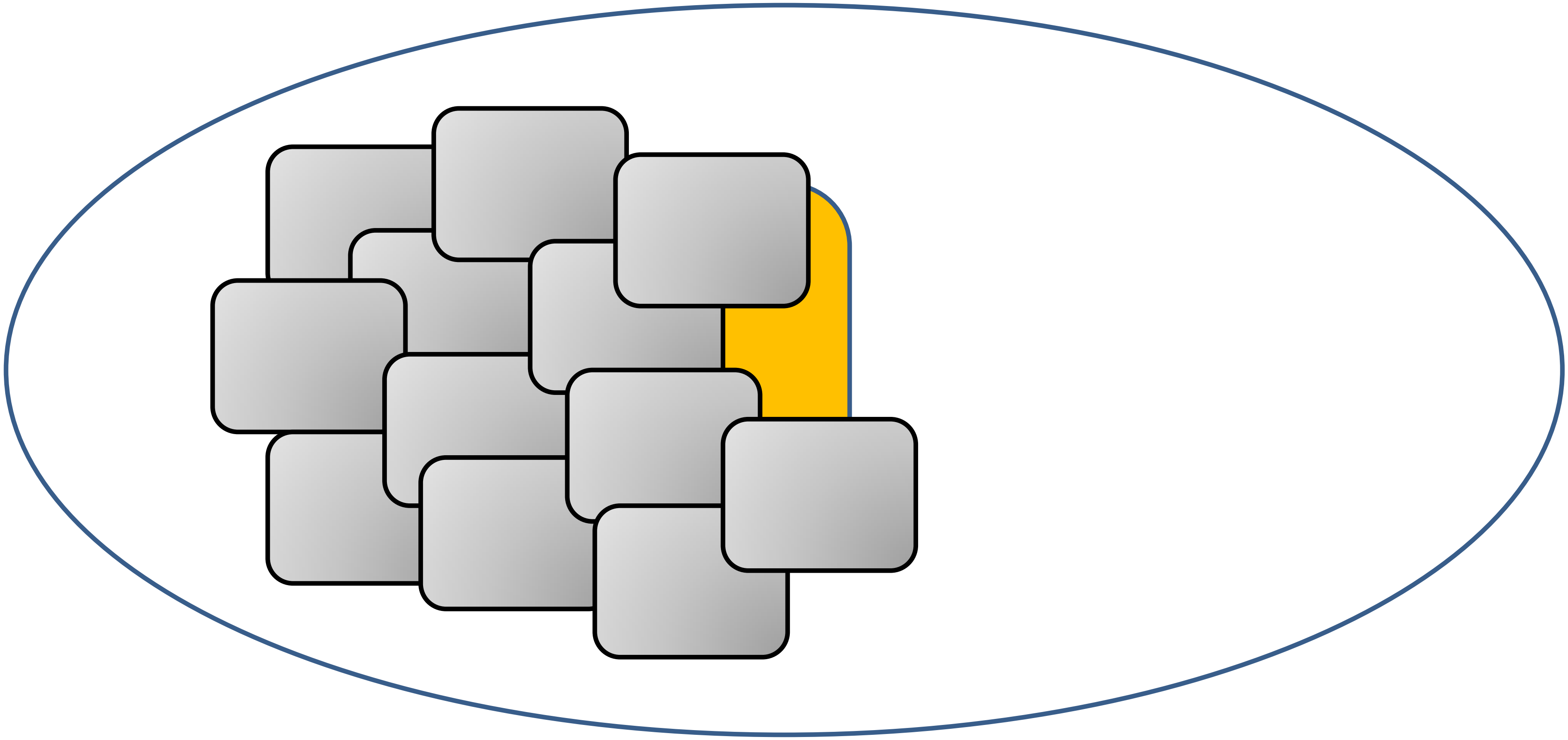
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



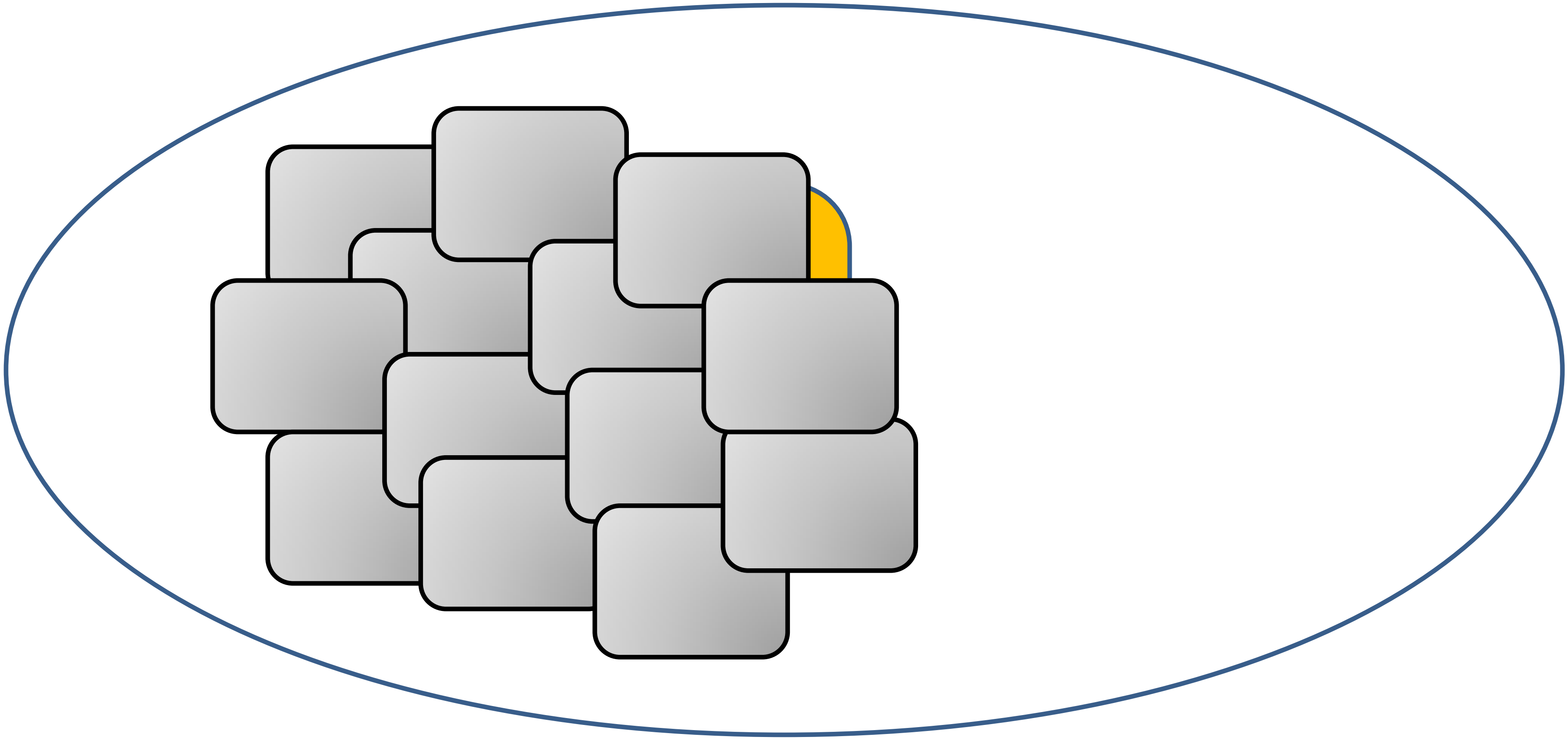
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



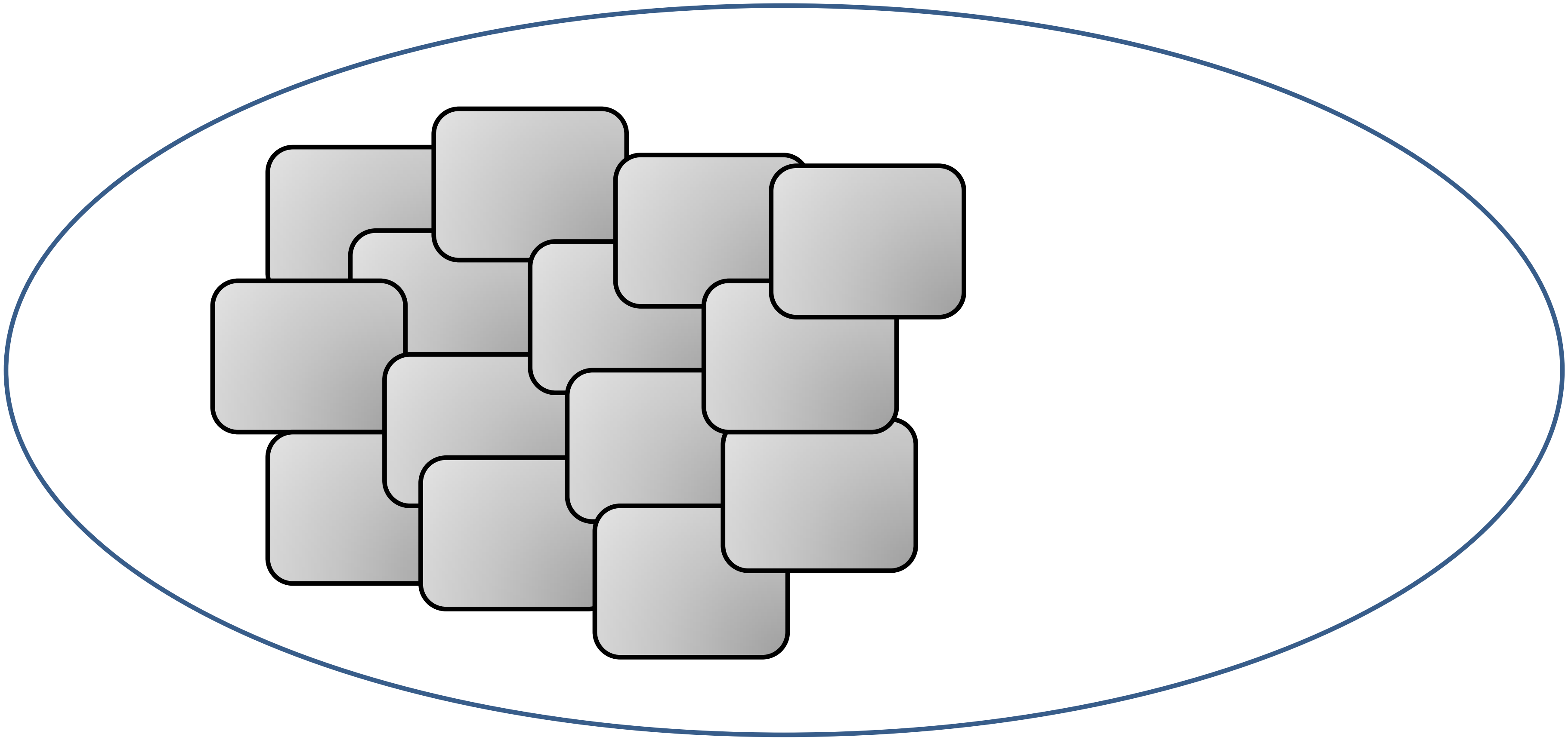
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



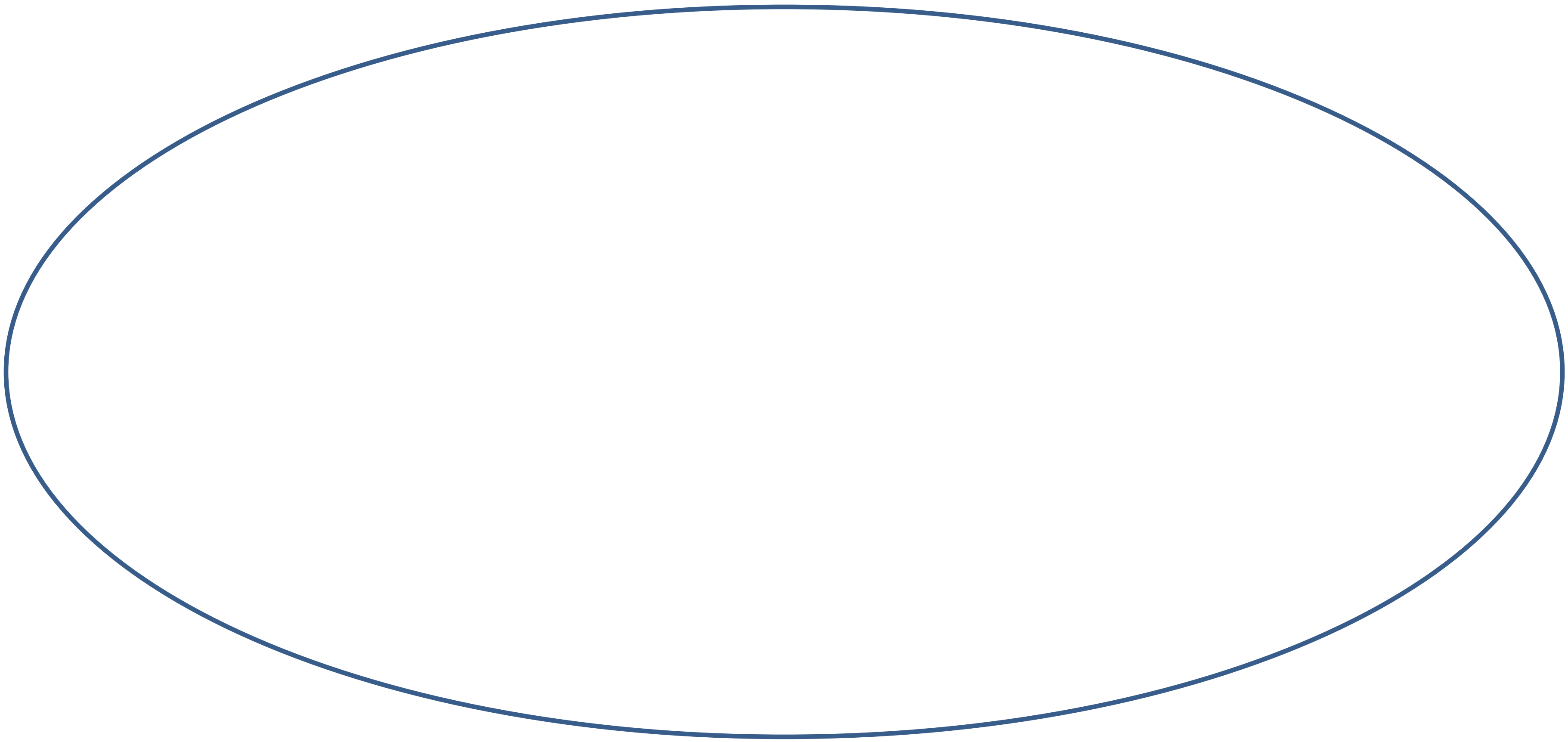
Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



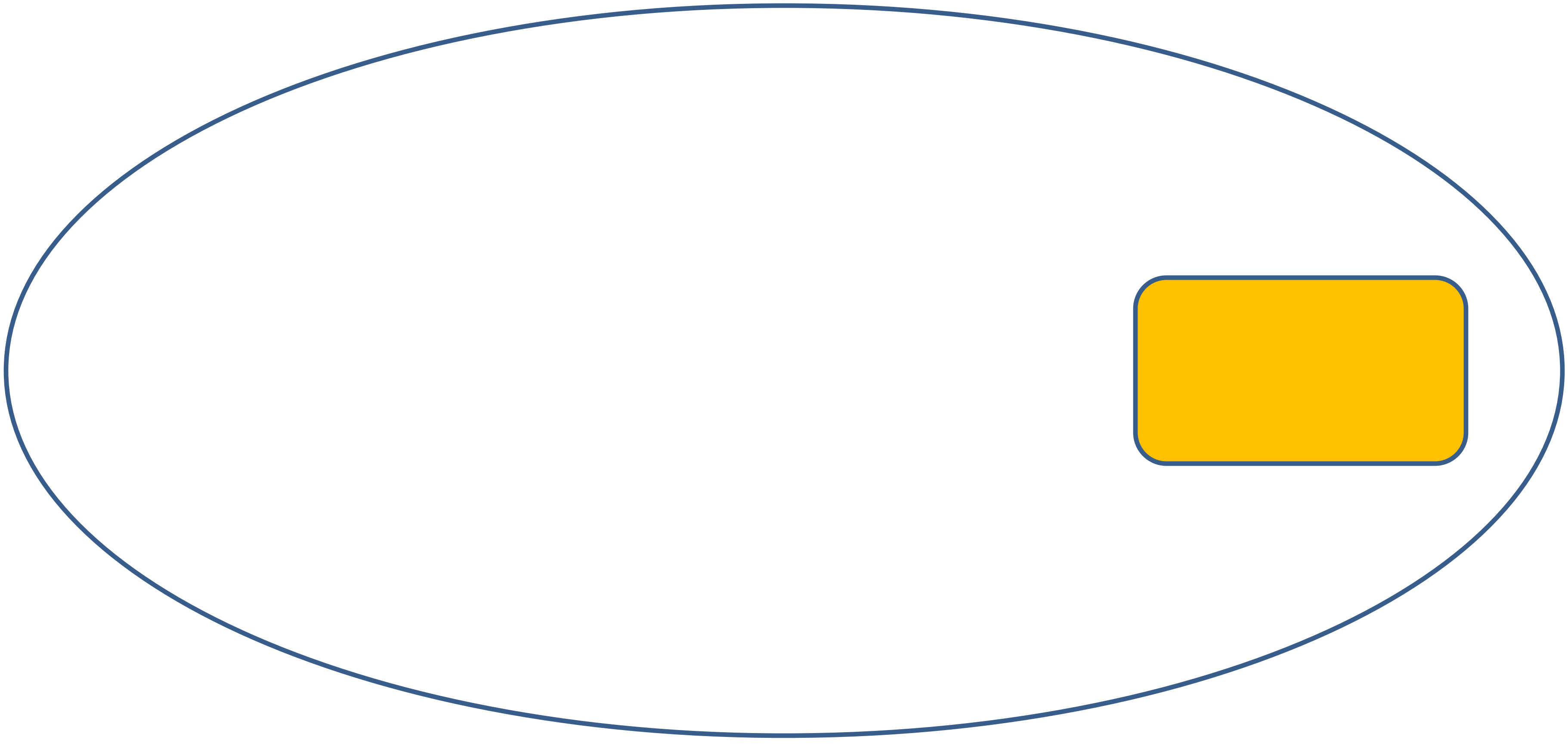
Assouad's embedding theorem

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



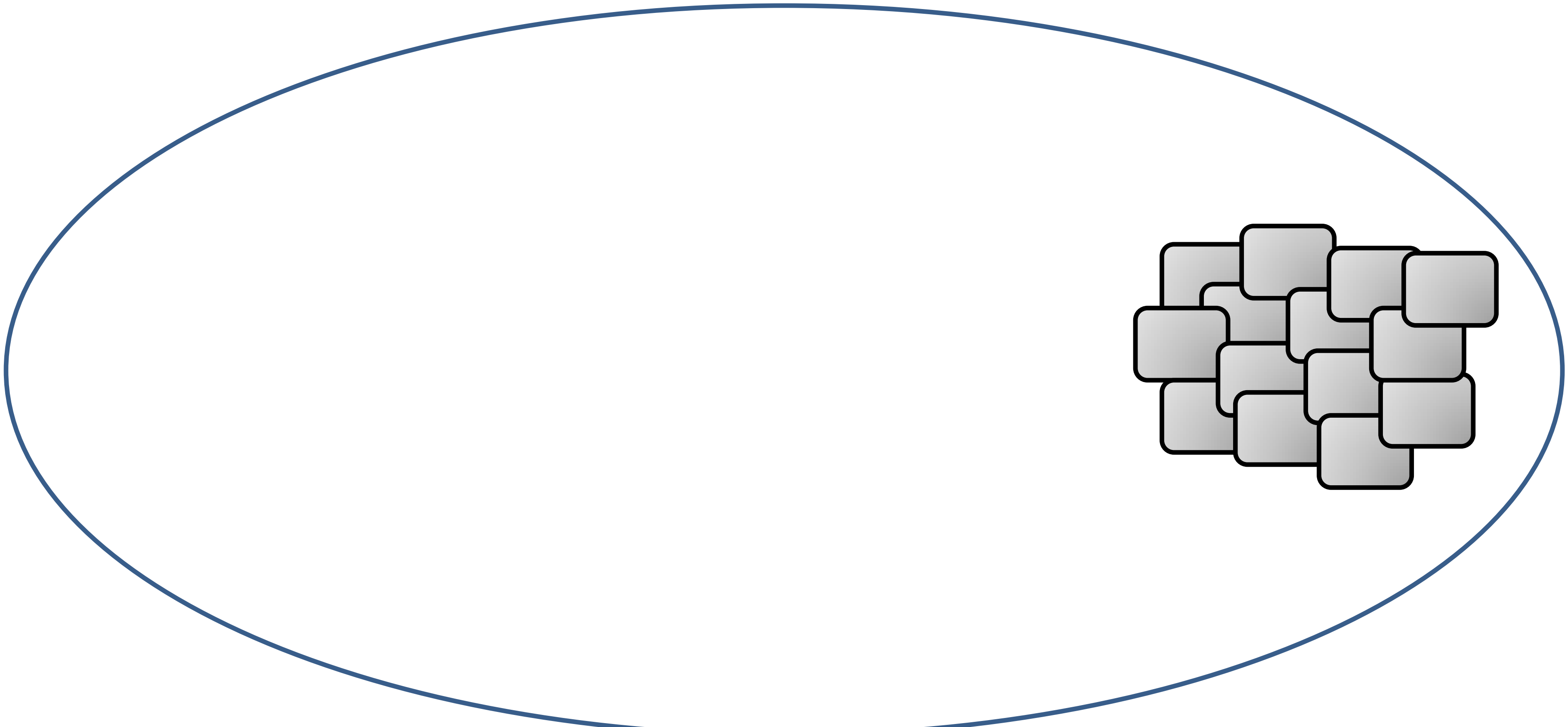
Assouad's embedding theorem

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



Assouad's embedding theorem (1983)

- A metric space (M, d_M) is K-doubling if any ball can be covered by K-balls of half its radius.



Theorem (Assouad, 1983). Suppose that (M, d_M) is K-doubling and $\epsilon \in (0,1)$. Then

$$(M, d_M^{1-\epsilon}) \xhookrightarrow{D(K,\epsilon)} \mathbb{R}^{N(K,\epsilon)}.$$

Theorem (N.-Neiman, 2010). In fact

$$(M, d_M^{1-\epsilon}) \xhookrightarrow{D(K,\epsilon)} \mathbb{R}^{N(K)}.$$

David-Snipes, 2013: Simpler deterministic proof.

$$(M, d_M^{1-\epsilon}) \xhookrightarrow{D(K,\epsilon)} \mathbb{R}^{N(K,\epsilon)}.$$

<ins>Obvious question</ins>: Why do we need to raise the metric to the power $1-\epsilon$?

$$(M, d_M^{1-\epsilon}) \xhookrightarrow{D(K,\epsilon)} \mathbb{R}^{N(K,\epsilon)}.$$

Obvious question: Why do we need to raise the metric to the power $1-\epsilon$?

Since in $(\mathbb{H}, d_W)$ we have $\forall m \in \mathbb{N}, \quad |B_m| \asymp m^4$, the metric space $(\mathbb{H}, d_W)$ is O(1)-doubling.

By Mostow-Pansu-Semmes, $(\mathbb{H}, d_W) \not\hookrightarrow \mathbb{R}^N$.

Proof of non-embeddability into $\mathbb{R}^n$

By a limiting argument and a non-commutative variant of Rademacher's theorem on the almost-everywhere differentiability of Lipschitz functions (Pansu differentiation) we have the statement

"If the Heisenberg group embeds bi-Lipschitzly into $\mathbb{R}^n$ then it also embeds into $\mathbb{R}^n$ via a bi-Lipschitz mapping that is a group homomorphism."

A non-Abelian group cannot be isomorphic to a subgroup of an Abelian group!

Heisenberg non-embeddability

- Mostow-Pansu-Semmes (1996).
- Cheeger (1999).
- Pauls (2001).
- Lee-N. (2006).
- Cheeger-Kleiner (2006).
- Cheeger-Kleiner (2007).
- Cheeger-Kleiner (2008).
- Cheeger-Kleiner-N. (2009).
- Austin-N.-Tessera (2010).
- Li (2013).

Heisenberg non-embeddability

- Mostow-Pansu-Semmes (1996).
- Cheeger (1999).
- Pauls (2001).
- Lee-N. (2006).
- Cheeger-Kleiner (2006).
- Cheeger-Kleiner (2007).
- Cheeger-Kleiner (2008).
- Cheeger-Kleiner-N. (2009).
- Austin-N.-Tessera (2010).
- Li (2013).

$\mathbb{H}$ does not embed into any uniformly convex space.

Heisenberg non-embeddability

- Mostow-Pansu-Semmes (1996).
- Cheeger (1999).
- Pauls (2001).
- Lee-N. (2006).
- Cheeger-Kleiner (2006).
- Cheeger-Kleiner (2007).
- Cheeger-Kleiner (2008).
- Cheeger-Kleiner-N. (2009).
- Austin-N.-Tessera (2010).
- Li (2013).

[ANT (2010)]: Hilbertian case

$$c_{\ell_2}(B_n, d_W) \asymp \sqrt{\log n}.$$

[ANT (2010)]: Hilbertian case

$$c_{\ell_2}(B_n, d_W) \asymp \sqrt{\log n}.$$

A limiting argument combined with [Aharoni-Maurey-Mityagin (1985), Gromov (2007)] shows that it suffices to treat embeddings that are 1-cocycles associated to an action by affine isometries. By [Guichardet (1972)] it further suffices to deal with coboundaries. This is treated by examining each irreducible representation separately.

[ANT (2010)]: *q*-convex case

If $(X, \|\cdot\|_X)$ is *q*-convex then

$$c_X(B_n, d_W) \gtrsim_X \left(\frac{\log n}{\log\log n}\right)^{1/q}.$$

[ANT (2010)]: *q*-convex case, continued

Qualitative statement: There is no bi-Lipschitz embedding of the Heisenberg group into an ergodic Banach space *X* via a 1-cocycle associated to an action by affine isometries.

X is ergodic if for every linear isometry $T : X \to X$ and every $x \in X$ the sequence

$$\frac{1}{n}\sum_{j=1}^{n-1} T^j x$$

converges in norm.

[ANT (2010)]: *q*-convex case, continued

N.-Peres (2010): In the case of *q*-convex spaces, it suffices to treat 1-cocycle associated to an affine action by affine isometries.

For combining this step with the use of ergodicity, uniform convexity is needed, because by [Brunel-Sucheston (1972)], ultrapowers of *X* are ergodic if and only if *X* admits an equivalent uniformly convex norm.

[ANT (2010)]: *q*-convex case, continued

Conclusion of proof uses algebraic properties of cocycles combined with rates of convergence for the mean ergodic theorem in *q*-convex spaces.

Li (2013): A quantitative version of Pansu's differentiation theorem. Suboptimal bounds.

Almost matching embeddability

Assouad (1983): If a metric space (M, d_M) is O(1)-doubling then there exists $k \in \mathbb{N}$ and 1-Lipschitz functions $\{\phi_j : M \to \mathbb{R}^k\}_{j \in \mathbb{Z}}$ such that for $x, y \in M$,

$$d_M(x,y) \in [2^{j-1}, 2^j] \implies \|\phi_j(x) - \phi_j(y)\|_2 \gtrsim d_M(x,y).$$

Almost matching embeddability

Assouad (1983): If a metric space (M, d_M) is O(1)-doubling then there exists $k \in \mathbb{N}$ and 1-Lipschitz functions $\{\phi_j : M \to \mathbb{R}^k\}_{j \in \mathbb{Z}}$ such that for $x, y \in M$,

$$d_M(x,y) \in [2^{j-1}, 2^j] \implies \|\phi_j(x) - \phi_j(y)\|_2 \gtrsim d_M(x,y).$$

So, define $f : B_n \to \bigoplus_{j=1}^{O(\log n)} \mathbb{R}^k$ by $f(x) = \bigoplus_{j=1}^{O(\log n)} \phi_j(x)$.

For $p \in [2, \infty)$ the bi-Lipschitz distortion of *f* is of order $(\log n)^{1/p}$.

Lafforgue-N., 2012

Theorem. For every *q*-convex space $(X, \|\cdot\|_X)$, every $f:\mathbb{H}\to X$ and every $n\in\mathbb{N}$,

$$\sum_{k=1}^{n^2}\sum_{x\in B_n}\frac{\|f(xc^k)-f(x)\|_X^q}{k^{1+q/2}}$$

$$\lesssim_X \sum_{x\in B_{21n}}\left(\|f(xa)-f(x)\|_X^q+\|f(xb)-f(x)\|_X^q\right).$$

The proof of this inequality relies on real-variable Fourier analytic methods. Specifically, a vector-valued Littlewood-Paley-Stein inequality due to Martinez, Torrea and Xu (2006), combined with a geometric argument.

For embeddings into ℓ_p one can use the classical Littlewood-Paley inequality instead.

Sharp non-embeddability

If $\forall x, y \in B_{22n}, \quad d_W(x,y) \leqslant \|f(x)-f(y)\|_X \leqslant D d_W(x,y),$

$$\sum_{x \in B_{21n}} \left(\|f(xa)-f(x)\|_X^q + \|f(xb)-f(x)\|_X^q\right)$$

$$\lesssim D^q |B_{21n}| \asymp D^q n^4,$$

and

$$\sum_{k=1}^{n^2} \sum_{x \in B_n} \frac{\|f(xc^k)-f(x)\|_X^q}{k^{1+q/2}} \geqslant \sum_{k=1}^{n^2} \sum_{x \in B_n} \frac{d_W(xc^k, x)^q}{k^{1+q/2}}$$

$$\gtrsim \sum_{k=1}^{n^2} \sum_{x \in B_n} \frac{k^{q/2}}{k^{1+q/2}} \asymp |B_n| \log n \asymp n^4 \log n.$$

$$\sum_{k=1}^{n^2} \sum_{x \in B_n} \frac{\|f(xc^k) - f(x)\|_X^q}{k^{1+q/2}}$$

$$\lesssim_X \sum_{x \in B_{21n}} \left(\|f(xa) - f(x)\|_X^q + \|f(xb) - f(x)\|_X^q\right),$$

so,

$$n^4 \log n \lesssim_X D^q n^4 \implies D \gtrsim_X (\log n)^{1/q}.$$

$$c_X(B_n, d_W) \gtrsim_X (\log n)^{1/q}.$$

Sharp distortion computation

$$p \in (1,2] \implies c_{\ell_p}(B_n, d_W) \asymp_p \sqrt{\log n}.$$

$$p \in [2,\infty) \implies c_{\ell_p}(B_n, d_W) \asymp_p (\log n)^{1/p}.$$

The Sparsest Cut Problem

<u>Input</u>: Two symmetric functions

$$C, D : \{1, \ldots, n\} \times \{1, \ldots, n\} \to [0, \infty).$$

<u>Goal</u>: Compute (or estimate) in polynomial time the quantity

$$\Phi^*(C, D) = \min_{\emptyset \neq S \subsetneq \{1,\ldots,n\}} \frac{\sum_{i,j=1}^{n} C(i,j)|\mathbf{1}_S(i) - \mathbf{1}_S(j)|}{\sum_{i,j=1}^{n} D(i,j)|\mathbf{1}_S(i) - \mathbf{1}_S(j)|}.$$

The Goemans-Linial Semidefinite Program

The best known algorithm for the Sparsest Cut Problem is a continuous relaxation called the Goemans-Linial SDP (~1997).

Theorem (Arora, Lee, N., 2005). The Goemans-Linial SDP outputs a number that is guaranteed to be within a factor of

$$(\log n)^{\frac{1}{2}+o(1)}$$

of $\Phi^*(C,D)$.

Minimize $\sum_{i,j=1}^{n} C(i,j)\|v_i - v_j\|_2^2$

over all $v_1, \ldots, v_n \in \mathbb{R}^n$,

subject to the constraints

$$\sum_{i,j=1}^{n} D(i,j)\|v_i - v_j\|_2^2 = 1,$$

and

$\forall\, i, j, k \in \{1, \ldots, n\}$,

$$\|v_i - v_j\|_2^2 \leqslant \|v_j - v_k\|_2^2 + \|v_k - v_j\|_2^2.$$

The link to the Heisenberg group

Theorem (Lee-N., 2006): The Goemans-Linial SDP has an integrality gap of at least $c_{\ell_1}(B_n, d_W)$.

Cheeger-Kleiner-N., 2009: There exists a universal constant *c>0* such that

$$c_{\ell_1}(B_n, d_W) \geqslant (\log n)^c.$$

Cheeger-Kleiner, 2007, 2008: Non-quantitative versions that also reduce matters to ruling out a certain more structured embedding.

Quantitative estimate controls phenomena that do not have qualitative counterparts.

Khot-Vishnoi (2005): The Goemans-Linial SDP has integrality gap at least $(\log \log n)^c$.

How well does the G-L SDP perform?

Conjecture: $c_{\ell_1}(B_n, d_W) \gtrsim \sqrt{\log n}.$

Remark: In a special case called *Uniform Sparsest Cut* (approximating graph expansion) the G-L SDP might perform better. The best known performance guarantee is $\lesssim \sqrt{\log n}$ [Arora-Rao-Vazirani, 2004] and the best known integrality gap lower bound is

$$e^{c\sqrt{\log \log n}}$$

[Kane-Meka, 2013].

Vertical perimeter versus horizontal perimeter

Conjecture: For every smooth and compactly supported $f:\mathbb{R}^3\to\mathbb{R}$,

$$\left(\int_0^\infty\left(\int_{\mathbb{R}^3}|f(x,y,z+t)-f(x,y,z)|dxdydz\right)^2\frac{dt}{t^2}\right)^{\frac12}$$
$$\lesssim\int_{\mathbb{R}^3}\left(\left|\frac{\partial f}{\partial x}(x,y,z)\right|+\left|\frac{\partial f}{\partial y}(x,y,z)+x\frac{\partial f}{\partial z}(x,y,z)\right|\right)dxdydz.$$

Lemma: A positive solution of this conjecture implies that $c_{\ell_1}(B_n,d_W)\gtrsim\sqrt{\log n}$.

Theorem (Lafforgue-N., 2012): For every *p>1,*

$$\left(\int_0^\infty \left(\int_{\mathbb{R}^3} |f(x,y,z+t)-f(x,y,z)|^p dxdydz\right)^{2/p} \frac{dt}{t^2}\right)^{1/2}$$

$$\lesssim_p \left(\int_{\mathbb{R}^3} \left(\left|\frac{\partial f}{\partial x}(x,y,z)\right|^p + \left|\frac{\partial f}{\partial y}(x,y,z) + x\frac{\partial f}{\partial z}(x,y,z)\right|^p\right) dxdydz\right)^{1/p}.$$

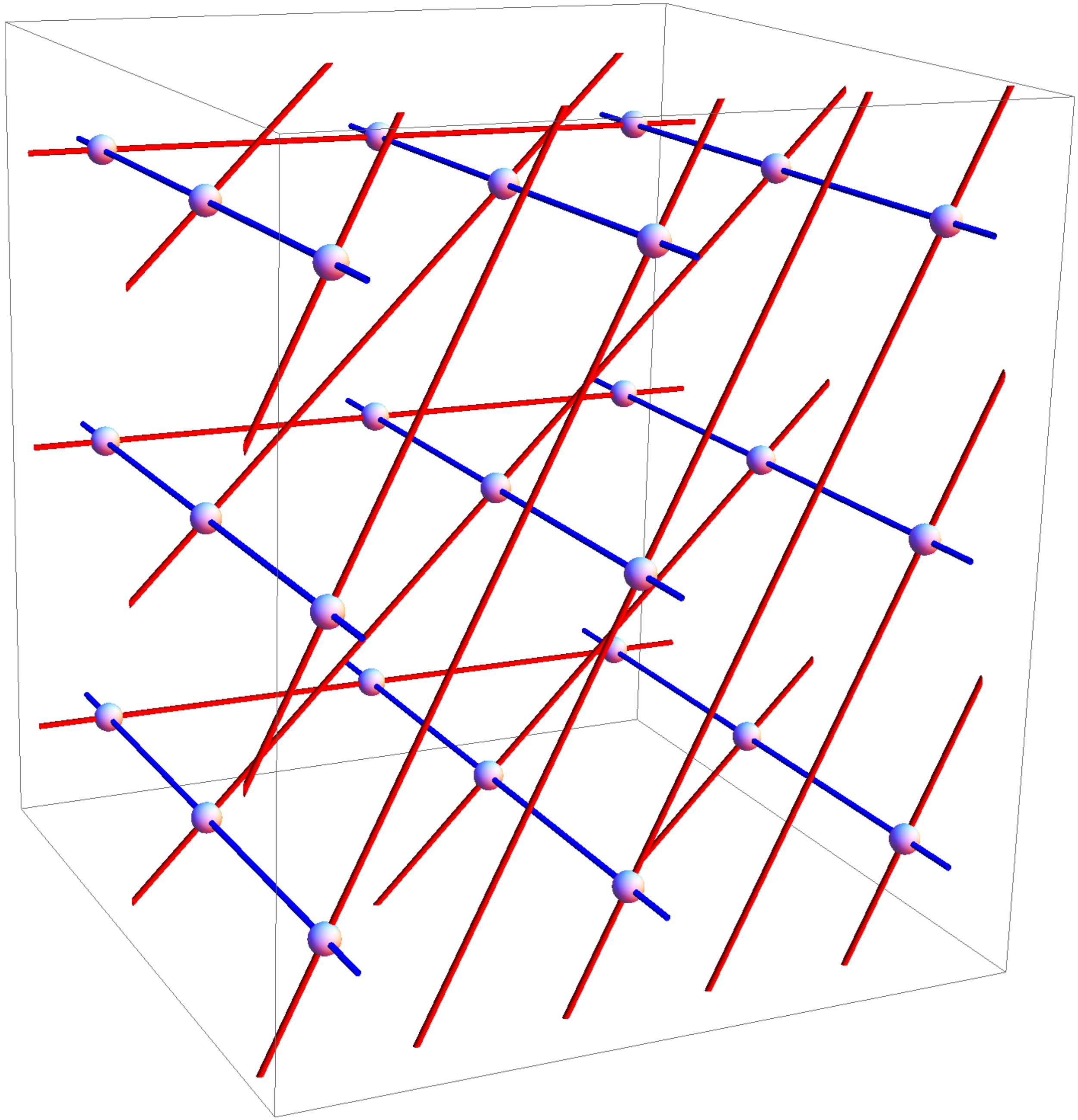
Equivalent form of the conjecture

Let *A* be a measurable subset of $\mathbb{R}^3$. For *t>0* define

$$v_t(A) = \mathrm{vol}\Big(\{(x,y,z) \in A : \ (x,y,z+t) \notin A\}\Big).$$

Then

$$\int_0^\infty \frac{v_t(A)^2}{t^2} dt \lesssim \mathrm{PER}(A)^2.$$



Proof of the vertical versus horizontal Poincare inequality

Equivalent statement: Suppose that $(X,\|\cdot\|_X)$ is q-convex and $f:\mathbb{R}^3\to X$ is smooth and compactly supported. Then

$$\left(\int_0^\infty\int_{\mathbb{R}^3}\frac{\|f(x,y,z+t)-f(x,y,z)\|_X^q}{t^{1+q/2}}dxdydz\right)^{\frac{1}{q}}$$

$$\lesssim_X\left(\int_{\mathbb{R}^3}\left(\left\|\frac{\partial f}{\partial x}(x,y,z)\right\|_X^q+\left\|\frac{\partial f}{\partial y}(x,y,z)+x\frac{\partial f}{\partial z}(x,y,z)\right\|_X^q\right)dxdydz\right)^{\frac{1}{q}}.$$

Proof of the equivalence: partition of unity argument + classical Poincare inequality for the Heisenberg group.

$$a = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} f\left(\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \epsilon & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right)$$

$$= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} f\left(\begin{pmatrix} 1 & x+\epsilon & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \right) = \frac{\partial f}{\partial x}.$$

$$a = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0} f\left(\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \epsilon \\ 0 & 0 & 1 \end{pmatrix} \right)$$

$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} f\left(\begin{pmatrix} 1 & x & z+\epsilon x \\ 0 & 1 & y+\epsilon \\ 0 & 0 & 1 \end{pmatrix} \right) = \frac{\partial f}{\partial y} + x\frac{\partial f}{\partial z}.$$

The Poisson semigroup

$$P_t(x) = \frac{1}{\pi(t^2 + x^2)}.$$

$$Q_t(x) = \frac{\partial}{\partial t} P_t(x) = \frac{x^2 - t^2}{\pi(t^2 + x^2)^2}.$$

Vertical convolution

For $\psi \in L_1(\mathbb{R})$,

$$\psi * f(x, y, z) = \int_{\mathbb{R}} \psi(u) f(x, y, z - u) du \in X.$$

Heisenberg gradient

$$\nabla_{\mathbb{H}} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} + x\frac{\partial f}{\partial z}\right) : \mathbb{R}^3 \to X \oplus X.$$

Proposition:

$$\left(\int_0^\infty \int_{\mathbb{R}^3} \frac{\|f(x,y,z+t) - f(x,y,z)\|_X^q}{t^{1+q/2}} dxdydz\right)^{\frac{1}{q}}$$

$$\lesssim \left(\int_0^\infty t^{q-1} \left\|Q_t * \nabla_{\mathbb{H}} f\right\|_{L_q(\mathbb{R}^3, X\oplus X)}^q dt\right)^{\frac{1}{q}}.$$

Littlewood-Paley

By Martinez-Torrea-Xu (2006), the fact that *X* is *q*-convex implies

$$\left(\int_0^\infty t^{q-1}\,\|Q_t * \nabla_{\mathbb{H}} f\|^q_{L_q(\mathbb{R}^3, X\oplus X)}\,dt\right)^{\frac{1}{q}} \lesssim \|\nabla_{\mathbb{H}} f\|_{L_q(\mathbb{R}^3, X\oplus X)}.$$

So, it remains to prove the proposition.

By a variant of a classical argument (using Hardy's inequality and semi-group properties),

$$\left(\int_0^\infty \int_{\mathbb{R}^3} \frac{\|f(x,y,z+t)-f(x,y,z)\|_X^q}{t^{1+q/2}} dxdydz\right)^{\frac{1}{q}}$$

$$\lesssim \left(\int_0^\infty t^{\frac{q}{2}-1} \|Q_t * f\|_{L_q(\mathbb{R}^3,X)}^q \, dt\right)^{\frac{1}{q}}.$$

So, we need to show that

$$\left(\int_0^\infty t^{\frac{q}{2}-1}\,\|Q_t * f\|_{L_q(\mathbb{R}^3,X)}^q\,dt\right)^{\frac{1}{q}} \lesssim \left(\int_0^\infty t^{q-1}\,\|Q_t * \nabla_{\mathbb{H}} f\|_{L_q(\mathbb{R}^3,X\oplus X)}^q\,dt\right)^{\frac{1}{q}}.$$

Key lemma: For every t>0,

$$\|Q_t * f - Q_{2t} * f\|_{L_q(\mathbb{R}^3,X)} \lesssim \sqrt{t}\,\|Q_t * \nabla_{\mathbb{H}} f\|_{L_q(\mathbb{R}^3,X\oplus X)}$$

The desired estimate

$$\left(\int_0^\infty t^{\frac{q}{2}-1}\,\|Q_t * f\|_{L_q(\mathbb{R}^3,X)}^q\,dt\right)^{\frac{1}{q}} \lesssim \left(\int_0^\infty t^{q-1}\,\|Q_t * \nabla_{\mathbb{H}} f\|_{L_q(\mathbb{R}^3,X\oplus X)}^q\,dt\right)^{\frac{1}{q}}$$

Follows from key lemma by the telescoping sum

$$Q_t * f = \sum_{m=1}^{\infty} \left(Q_{2^{m-1}t} - Q_{2^m t} * f\right).$$

Proof of key lemma

Since $P_{2t} = P_t * P_t$ we have $Q_{2t} = P_t * Q_t$.

So, by identifying $\mathbb{R}^3$ with $\mathbb{H}$, for every $h \in \mathbb{R}^3$,

$$
\begin{aligned}
& Q_t * f(h) - Q_{2t} * f(h) \\
& = Q_t * f(h) - P_t * Q_t * f(h) \\
& = \int_{\mathbb{R}} P_t(u) \left(Q_t * f(h) - Q_t * f(hc^{-u})\right) du.
\end{aligned}
$$

For every *s>0* consider the commutator path

$$\gamma_s : [0, 4\sqrt{s}] \to \mathbb{R}^3,$$

$$\gamma_s(\theta) = \begin{cases} a^{\theta} & \text{if } 0 \leqslant \theta \leqslant \sqrt{s}, \\ a^{\sqrt{s}} b^{\theta - \sqrt{s}} & \text{if } \sqrt{s} \leqslant \theta \leqslant 2\sqrt{s}, \\ a^{\sqrt{s}} b^{\sqrt{s}} a^{-\theta + 2\sqrt{s}} & \text{if } 2\sqrt{s} \leqslant \theta \leqslant 3\sqrt{s}, \\ a^{\sqrt{s}} b^{\sqrt{s}} a^{-\sqrt{s}} b^{-\theta + 3\sqrt{s}} & \text{if } 3\sqrt{s} \leqslant \theta \leqslant 4\sqrt{s}. \end{cases}$$

So, $\gamma_s(0) = 0 = e_{\mathbb{H}}$ and

$$\gamma_s(4\sqrt{s}) = \left[a^{\sqrt{s}}, b^{\sqrt{s}}\right] = [a, b]^s = c^s.$$

Hence,

$$Q_t * f(h) - Q_t * f(hc^{-u})$$
$$= \int_0^{4\sqrt{u}} \frac{d}{d\theta} Q_t * f\left(hc^{-u}\gamma_u(\theta)\right) d\theta.$$

By design, $\frac{d}{d\theta}Q_t * f\left(hc^{-u}\gamma_u(\theta)\right)$ is one of

$$\partial_a Q_t * f(hc^{-u}\gamma_u(\theta)) = Q_t * \partial_a f(hc^{-u}\gamma_u(\theta))$$

or

$$\partial_b Q_t * f(hc^{-u}\gamma_u(\theta)) = Q_t * \partial_b f(hc^{-u}\gamma_u(\theta)),$$

where $\partial_a = \partial_x$ and $\partial_b = \partial_y + x\partial_z$.

We used here the fact that since Q_t is convolution along the center, it commutes with $\partial_a, \ \partial_b.$

We saw that

$$Q_t * f(h) - Q_{2t} * f(h))$$

$$= \int_{\mathbb{R}} P_t(u) \left(Q_t * f(h) - Q_t * f(hc^{-u}) \right) du$$

$$= \int_{\mathbb{R}} P_t(u) \int_0^{4\sqrt{u}} \frac{d}{d\theta} Q_t * f \left(hc^{-u} \gamma_u(\theta) \right) d\theta du.$$

Now the key lemma follows from the triangle inequality and the fact that $\displaystyle\int_0^{\infty} \sqrt{u} P_t(u) du \asymp \sqrt{t}.$