Computational Homology Project (CHomP)

http://chomp.rutgers.edu/

Konstantin Mischaikow

Mathematics and BioMaPS Institute
Rutgers University

mischaik@math.rutgers.edu

Computing Homology

1. Homology
2. Geometric Complexes
3. Geometric Preprocessors
4. Maps

1. Homology

vector space

X a topological space $\Longrightarrow$ $H_*(X)$ an ~~abelian group~~

$$H_k(X, \mathbb{Z}_2) \cong \mathbb{Z}_2^{\beta_k}$$

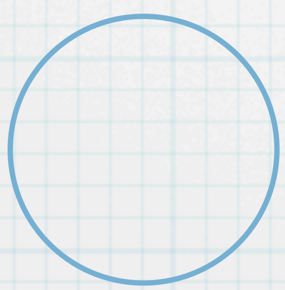
Betti number

$f : X \to Y$ continuous map $\Longrightarrow$ $f_* : H_*(X) \to H_*(Y)$

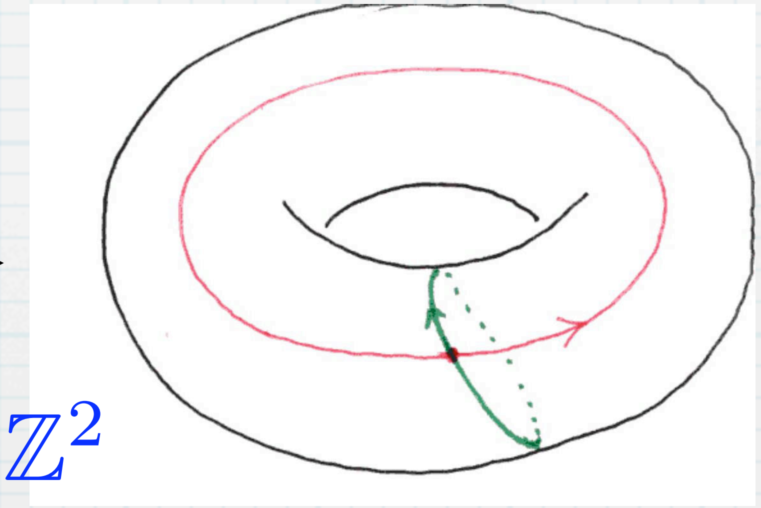
~~group homomorphism~~

linear map

$H_1(S^1) \cong \mathbb{Z}$



$\xrightarrow{f}$



$H_1(T) \cong \mathbb{Z}^2$

$f_1 : \mathbb{Z} \to \mathbb{Z}^2$

With an appropriate choice of basis $f_1 = [a, b]$

Computing Homology

A free chain complex $\mathcal{C} = \{C_k, \partial_k\}_{k\in\mathbb{Z}}$ consists of free abelian groups C_k, called chains, and homomorphisms $\partial_k : C_k \to C_{k-1}$, called boundary operators, such that

$$\partial_k \circ \partial_{k+1} = 0.$$

The cycles are the subgroups $Z_k := \ker \partial_k$.
The boundaries are the subgroups $B_k := \text{image}\, \partial_{k+1}$.
The k-th Homology group is

$$H_k(\mathcal{C}) := \frac{Z_k}{B_k}$$

Theorem: (**Smith Normal Form**) Given a matrix $A : \mathbb{Z}^n \to \mathbb{Z}^m$ there exist invertible matrices $R : \mathbb{Z}^n \to \mathbb{Z}^n$ and $Q : \mathbb{Z}^m \to \mathbb{Z}^m$ such that

$$B := Q^{-1}AR = \left[\begin{array}{ccccc|c} b_1 & & & & & \\ & b_2 & & 0 & & \\ & & \cdot & & & \\ & & & \cdot & & 0 \\ & 0 & & & \cdot & \\ & & & & b_t & \\ \hline & & 0 & & & 0 \end{array}\right],$$

where b_i are positive integers, $b_i = 1$ for $i = 1, 2, \ldots, s$, and b_i divides b_{i+1} for $i = 1, 2, \ldots, t-1$.

Let $X \subset \mathbb{R}^d$ be a ~~cubical~~ simplicial set. The boundary map $\partial_k : C_k(X) \to C_{k-1}(X)$ takes the form of a matrix

$$A_k : \mathbb{Z}^{d_k} \longrightarrow \mathbb{Z}^{d_{k-1}}$$

where the basis is defined by the ~~elementary cubes~~ simplicies.

$$\begin{tikzcd}
 & \mathbb{Z}^{d_k} \arrow[r, "B_k"] & \mathbb{Z}^{d_{k-1}} \arrow[d, "Q_{k-1}^{-1}"] \\
C_{k+1} \arrow[r, "\partial_{k+1}"] \arrow[d, "R_{k+1}"] & C_k \arrow[r, "\partial_k"] \arrow[u, "R_k"] & C_{k-1} \\
\mathbb{Z}^{d_{k+1}} \arrow[r, "B_{k+1}"] & \mathbb{Z}^{d_k} \arrow[u, "Q_k^{-1}"] &
\end{tikzcd}$$

$$H_k(X) \cong \frac{\ker B_k}{R_k Q_k^{-1} B_{k+1}(\mathbb{Z}^{d_{k+1}})}$$

2. Geometric Complexes

Rayleigh-Benard Convection

Raw intensity image captured using 12-bit digital camera and filtering

Binary image obtained by thresholding at the median value of intensity

Data is in not in the form of a chain complex

Data sets are large

Cubical complexes

An *elementary interval* is a closed interval $I \subset \mathbb{R}$ of the form

$$I = [l, l+1] \quad \text{or} \quad I = [l, l]$$

for some $l \in \mathbb{Z}$.

An *elementary cube* is a finite product of elementary intervals

$$Q = I_1 \times I_2 \times \cdots \times I_d \subset \mathbb{R}^d,$$

where each I_i is an elementary interval.

$$\mathcal{K}_k^d := \{k\text{-dimensional cubes in } \mathbb{R}^d\}$$

A cubical set is a finite union of elementary cubes

3. Geometric Preprocessors

Co-Reduction, M. Mrozek, B. Batko,

Dis. Comp. Geom. '09

Idea: Reduce the geometric complex as much as possible before building the chain complex and computing Smith Normal Form.

Fact: If $A \subset X$ is acylic, then

$$H_k(X) \cong \begin{cases} H_k(X,A) \oplus \mathbb{Z} & \text{if } k = 0 \\ H_k(X,A) & \text{otherwise} \end{cases}$$

$$C_*(X,A) := \frac{C_*(X)}{C_*(A)}$$

Goal: Remove a "maximal" acyclic subset A of X.

The Co-Reduce Algorithm

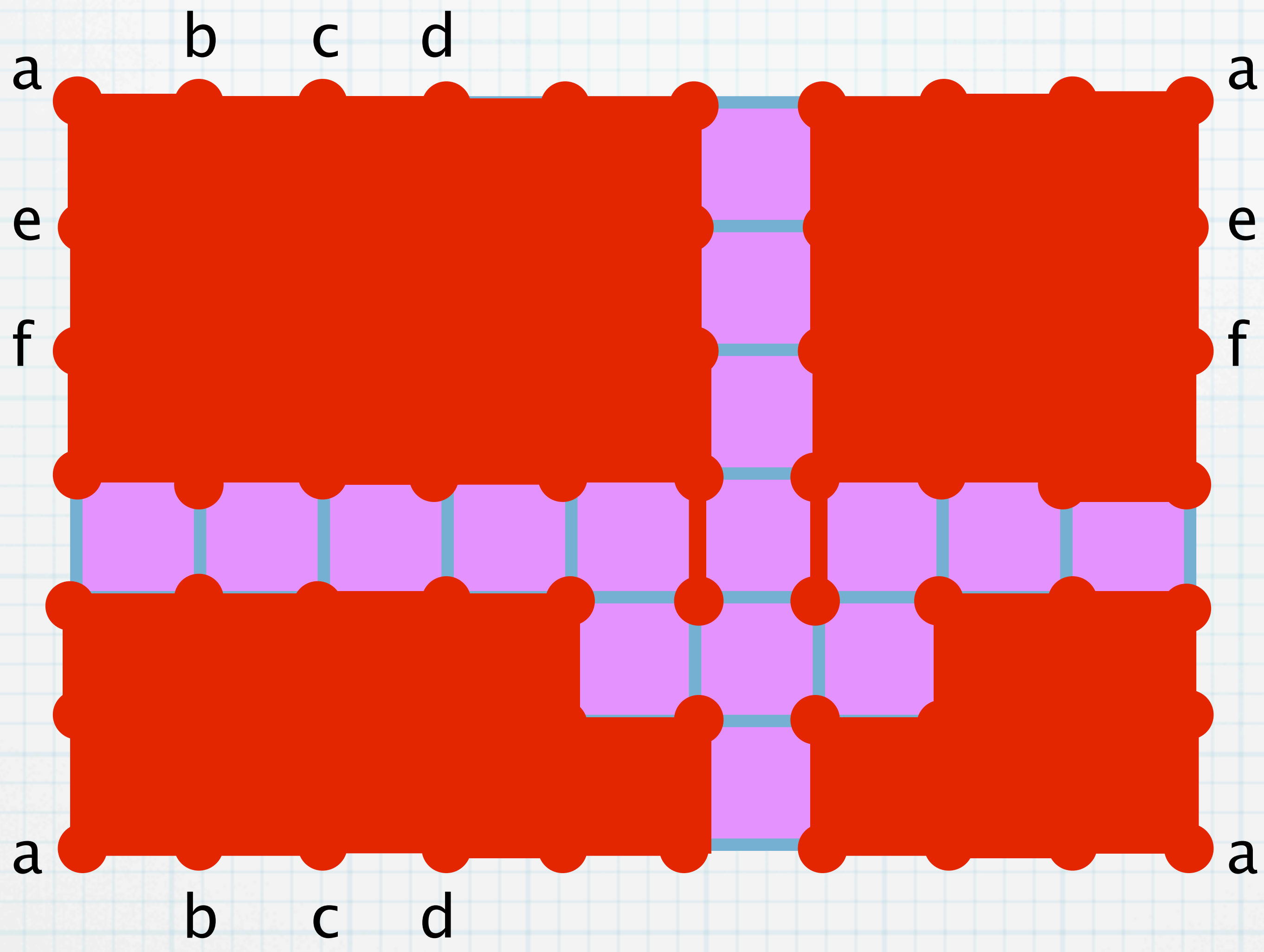
Queue Q := NULL
$Q \leftarrow \sigma_0$
while $(|Q| \neq 0)$
    $Q \rightarrow \sigma$
    if $(\partial_{\mathcal{X}}(\sigma) = \{\beta\})$
        $\mathcal{X} = \mathcal{X} \setminus \{\sigma\}$
        $Q \leftarrow \delta_{\mathcal{X}}(\beta)$
        $\mathcal{X} = \mathcal{X} \setminus \{\beta\}$
    else if $(\partial_{\mathcal{X}}(\sigma) = \emptyset)$
        $Q \leftarrow \delta_{\mathcal{X}}(\beta)$
    end if
end while
return $\mathcal{X}$

Geometric boundary

Geometric co-boundary

X is a Torus (70 vertices, 123 edges, 54 squares)

Goal: Construct a large acyclic set A



C(X,A) generated by 19 edges, 16 squares

Run Times co-reduce

Random cubical complexes in $\mathbb{R}^4$

Size	$\beta_0, \beta_1, \beta_2$	Time
62,636	24, 2988, 20	0.53
116,284	36, 5552, 38	1.06
168,750	48, 8060, 54	1.56
222,480	54, 10493, 71	2.09

3.6 GHz PC
Pentium 4 processor
2 GB RAM

Large and high complexes

Dim.	Size	Betti	Time
3	8,392,997	1, 1057	4.38
4	350,000,000	1, 7, 352	935
12	135,351	1, 0, 0, 1, 4, 5, 0	0.77
12	332,309	1, 0, 0, 0, 0, 1, 1	1.92

Summary: The goal

Input: Large Geometrical Complex

Output: Small/Simple Chain Complex

3. Maps

A. Chain Maps (Algebra)

B. Approximating Continuous Maps

C. Constructing Chain Maps

D. Geometric Reductions

A. Chain Maps (Algebra)

Let $\mathcal{C} = \{C_k, \partial_k\}$ and $\mathcal{C}' = \{C'_k, \partial'_k\}$ be chain complexes. A sequence of abelian group homomorphisms $\varphi_k : C_k \to C'_k$ is a chain map if, for every $k \in \mathbb{Z}$,

$$\partial'_k \varphi_k = \varphi_{k-1} \partial_k.$$

Proposition: If $\varphi : \mathcal{C} \to \mathcal{C}'$ is a chain map, then

$$\begin{aligned} \varphi_k(Z_k) &\subset Z'_k \\ \varphi_k(B_k) &\subset B'_k \end{aligned}$$

for all $k \in \mathbb{Z}$.

Define the group homomorphisms $\varphi_{k*} : H_k(\mathcal{C}) \to H_k(\mathcal{C}')$ by

$$\varphi_{k*}([z]) := [\varphi_k(z)]$$

B. Approximating Continuous Maps

Consider a continuous map $f : (X, A) \to (Y, B)$

Need to construct a chain map $f_{\#} : C_*(X, A) \to C_*(Y, B)$

First Question:

Consider $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$. Let $x \in X$. What is $f(x)$?

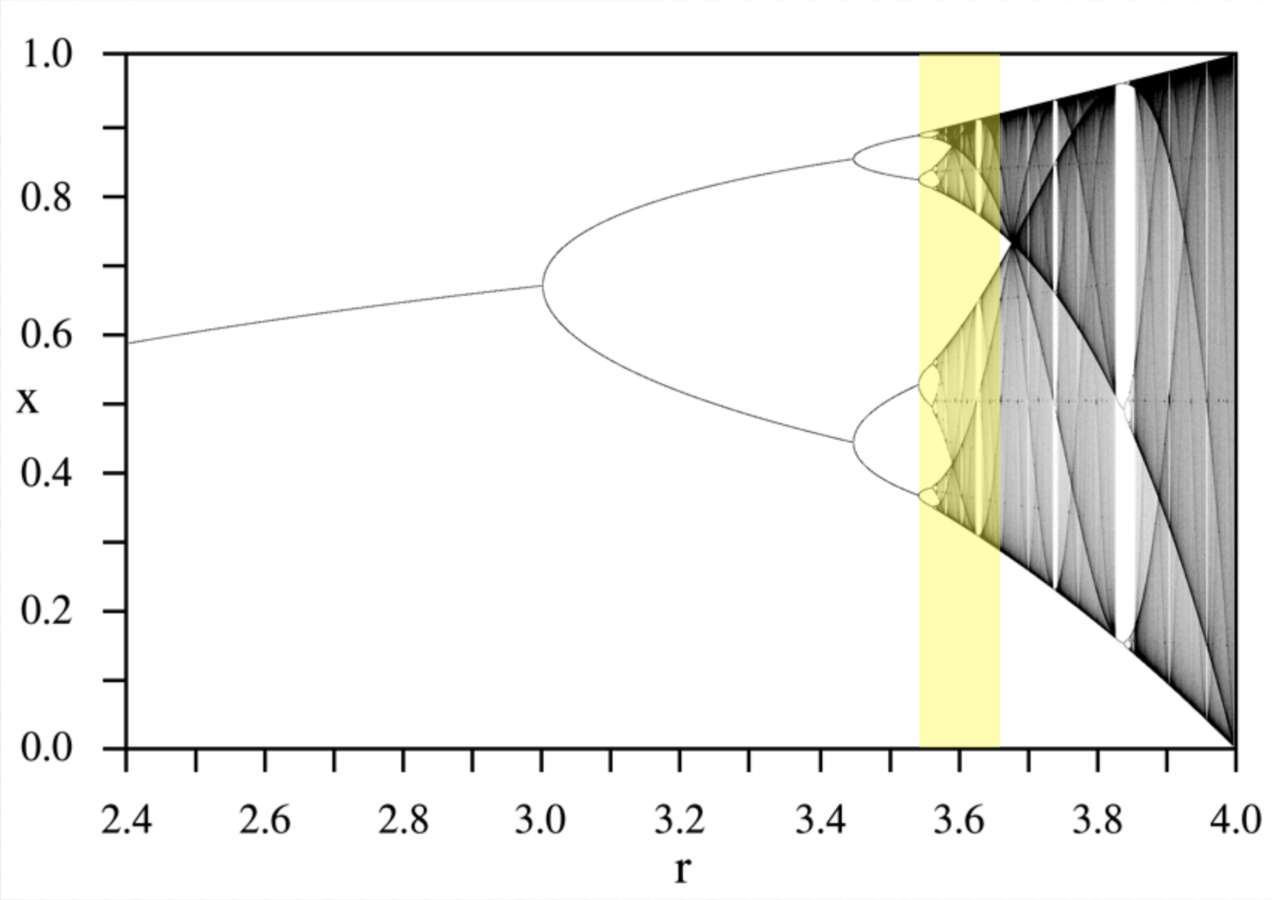
1. Consider a differential equation $\dot{x} = g(x)$
   which generates a flow $\psi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$.
   A natural continuous map is $f(x) := \psi(1, x)$.
   $f(x)$ is obtained by numerically integrating ode.

Cost of Integration vs. Numerical Errors

2. Consider a parameterized family of maps

$$f : X \times \Lambda \to X$$

Want to study dynamics over a set $Q \subset \Lambda$ of parameter values



Example: The Logistic Map

$$f(x, r) = r \cdot x \cdot (1 - x)$$

Explicit nonlinearities may give misleading information

$f(x, Q)$ is the minimal object of interest.

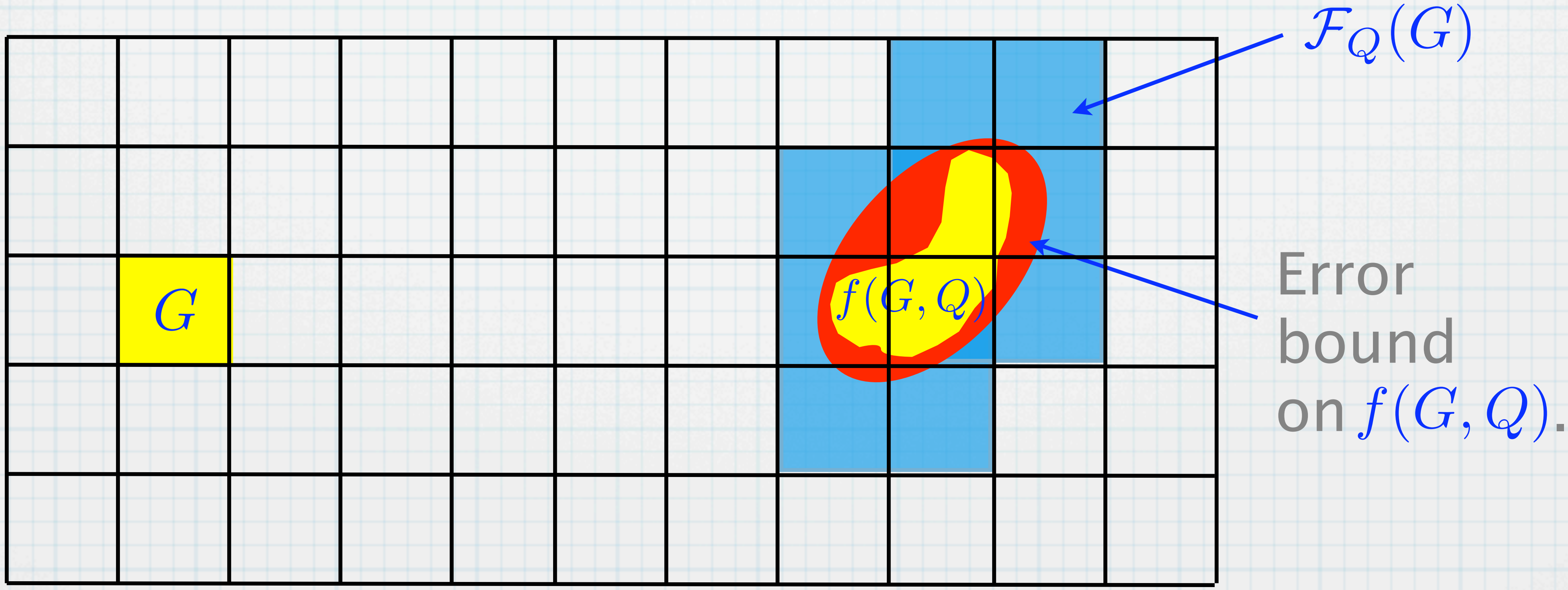
Constructing an acyclic multivalued approximation

Consider $f : \mathbb{R}^2 \to \mathbb{R}^2$. Accounting for errors

$$x \mapsto [f_1(x) - \epsilon, f_1(x) + \epsilon] \times [f_2(x) - \epsilon, f_2(x) + \epsilon]$$

Choose a cubical grid $\mathcal{X}$ that covers X.

Construct a combinatorial multivalued map $\mathcal{F}_Q : \mathcal{X} \rightrightarrows \mathcal{X}$.



C. Constructing Chain Maps and Homology

$f : X \to Y$ continuous

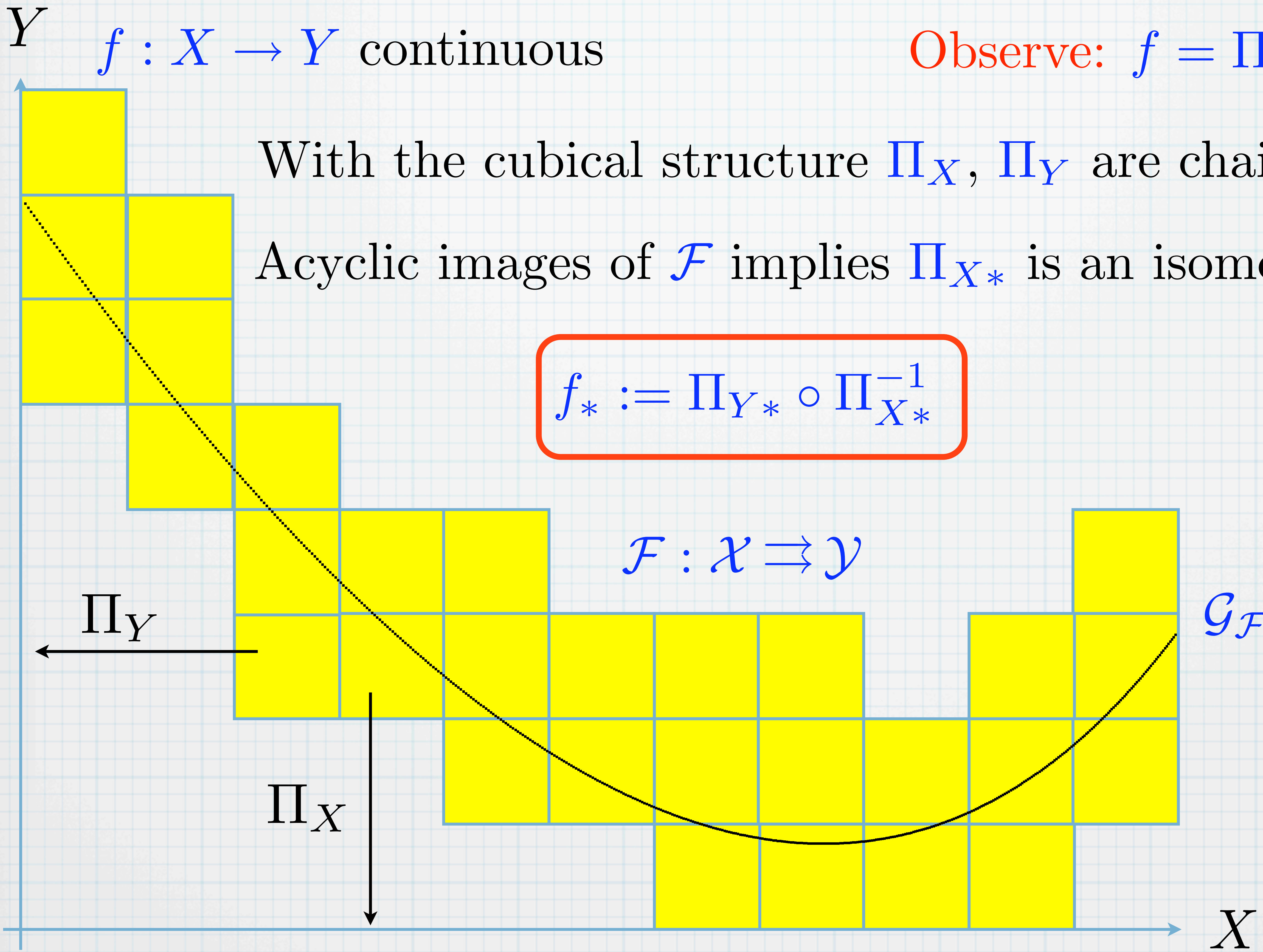
Observe: $f = \Pi_Y \circ \Pi_X^{-1}$

With the cubical structure Π_X, Π_Y are chain maps.

Acyclic images of $\mathcal{F}$ implies Π_{X*} is an isomorphism.

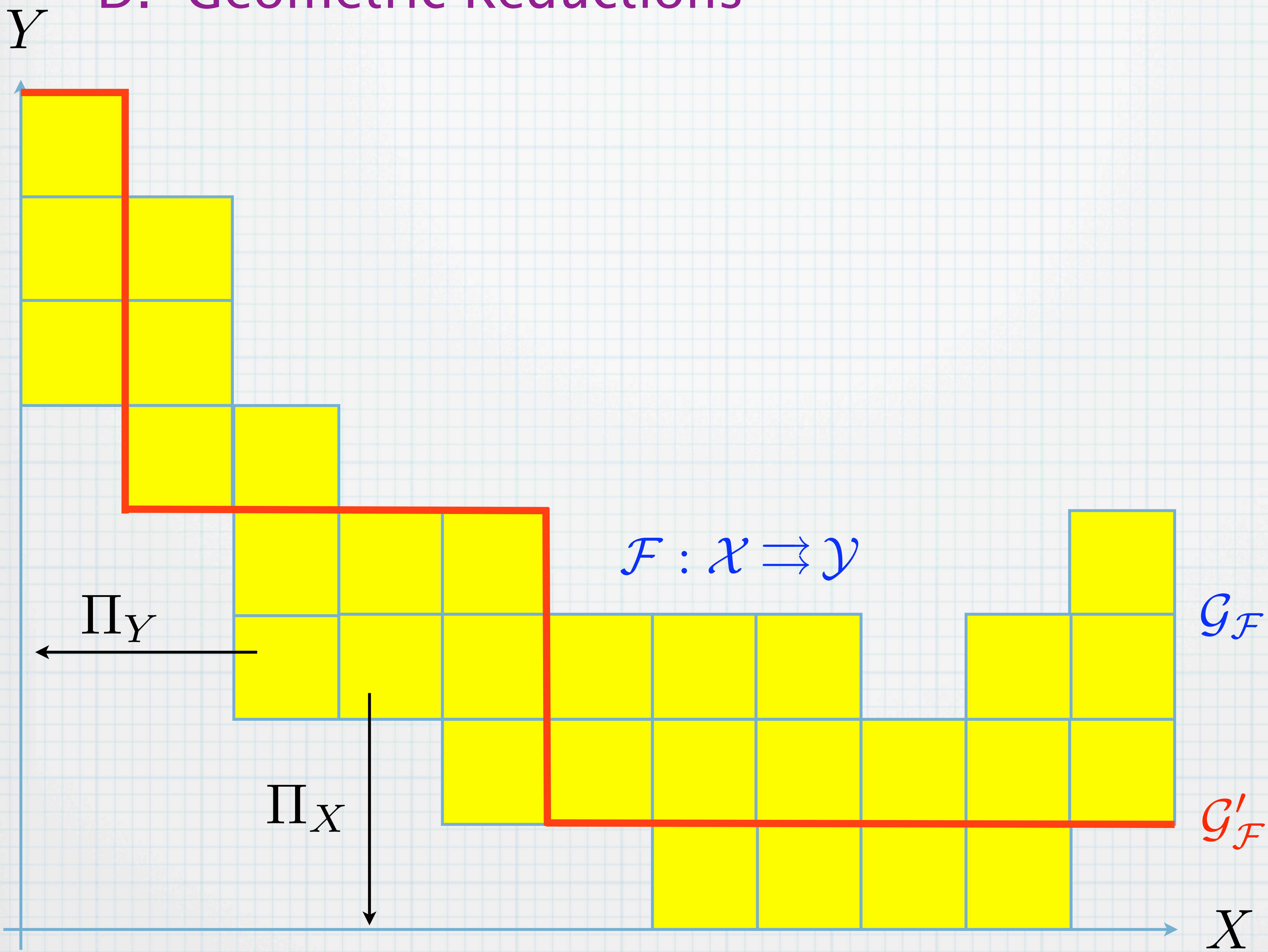
$$\boxed{f_* := \Pi_{Y*} \circ \Pi_{X*}^{-1}}$$

$\mathcal{F} : \mathcal{X} \rightrightarrows \mathcal{Y}$



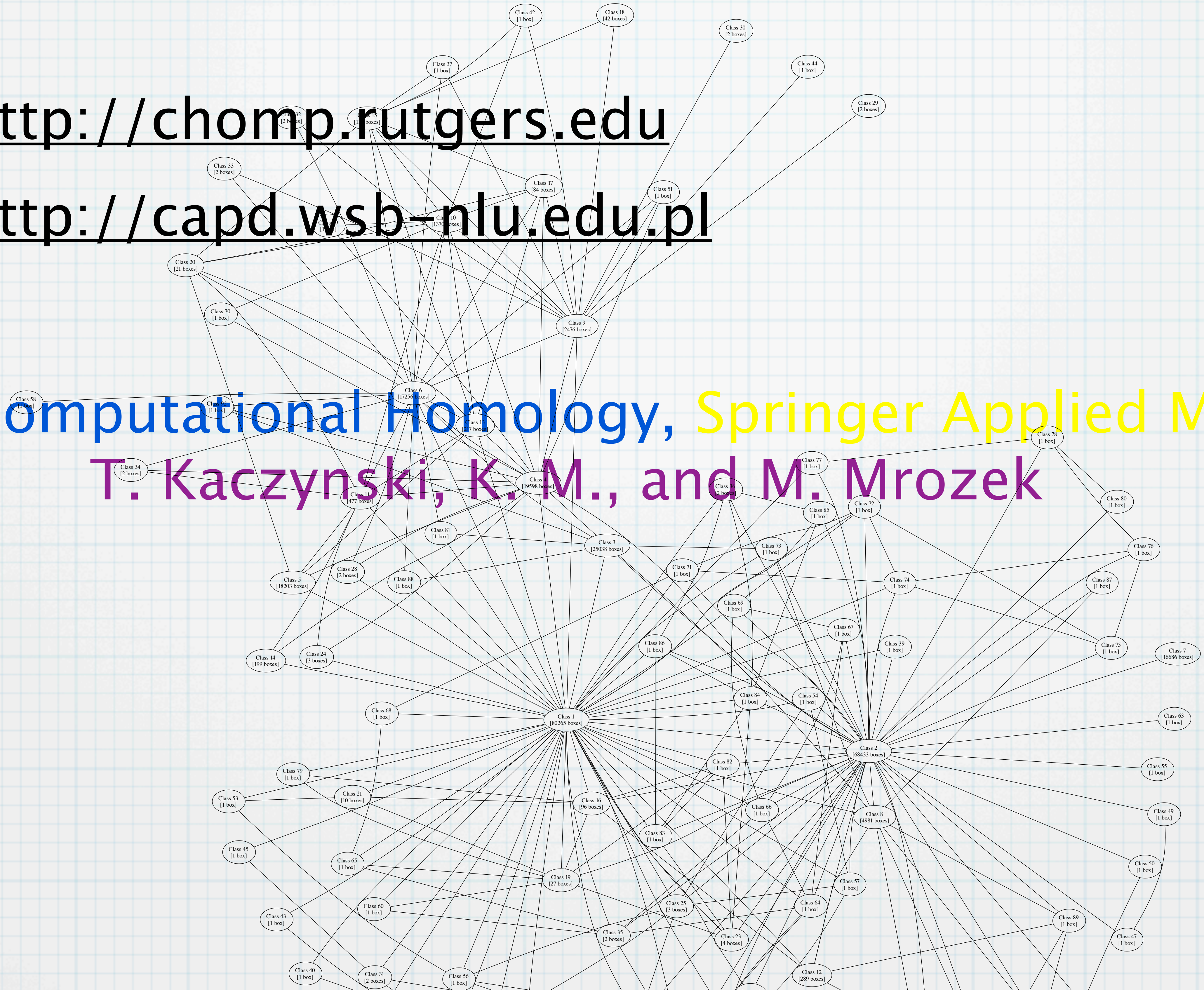
K. M., M. Mrozek, P. Pilarczyk FoCM (2005)

D. Geometric Reductions



http://chomp.rutgers.edu

http://capd.wsb-nlu.edu.pl



Computational Homology, Springer Applied Math

T. Kaczynski, K. M., and M. Mrozek

DARPA

National Science Foundation
WHERE DISCOVERIES BEGIN

U.S. DEPARTMENT OF ENERGY