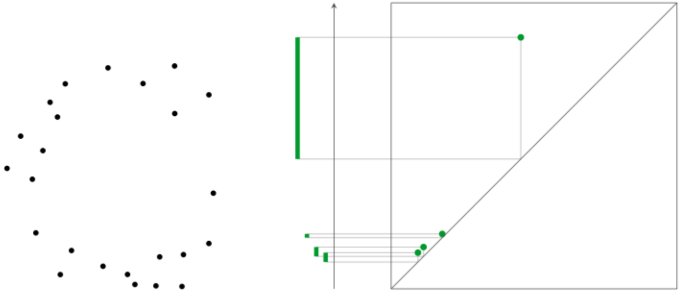
An Introduction to Multi-D Persistent Homology

Michael Lesnick
Princeton University → SUNY Albany

IMA Tutorial on Multiparameter Persistence, Computation, and Applications,
August 13, 2018

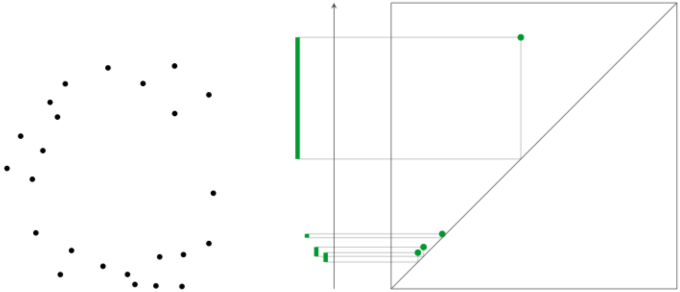
Persistent Homology

- Provides invariants of data called **barcodes**
- Barcode is a collection of intervals $[b, d)$ in $\mathbb{R}$



Persistent Homology

- Provides invariants of data called **barcodes**
- Barcode is a collection of intervals $[b, d)$ in $\mathbb{R}$



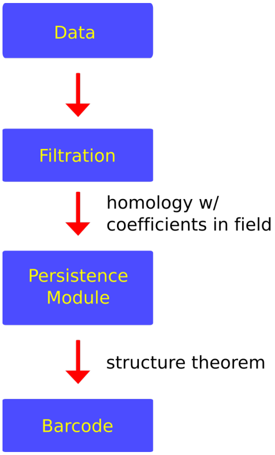
Multi-D (a.k.a. Multiparameter) Persistent Homology:

- Arises naturally in applications
- Yields richer but more complex invariants of data
- Ordinary persistence theory / methodology doesn't extend naively

My belief: Multi-parameter persistence has the potentially to greatly add to the power of one-parameter persistence in applications:

- good recent progress
- core ideas and technology still under development
- lots of opportunities to contribute.

Pipeline for Constructing Barcodes



Data→ Filtration (Vietoris-Rips)

Let P be a metric space with **integer-valued metric**.

For $t \geq 0$, let G_t be the t-**neighborhood graph** of P:

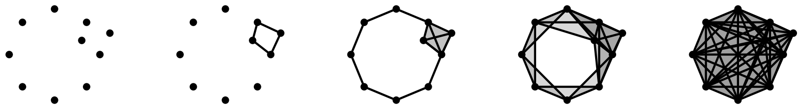
- G_t has vertex set P,
- $[q, r] \in G_t$ iff $d_P(q, r) \leq 2t$.

Let $\mathrm{Rips}(P)_t$ be the **clique complex** of G_t:

- For each 3-clique in G_t, add in a triangle,
- for each 4-clique in G_t, add in a tetrahedron,
- etc.

$\mathrm{Rips}(P)_s \subseteq \mathrm{Rips}(P)_t$ whenever $s \leq t$, so we have a filtration

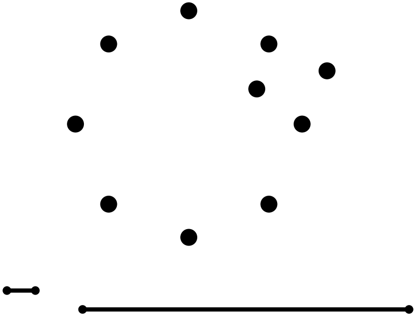
$$\mathrm{Rips}(P) := \{\mathrm{Rips}(P)_t\}_{t \in \mathbb{N}}.$$



Rips Stability Theorem

Theorem [Edelsbrunner et al. 2007, Chazal et al. 2009] For all finite metric spaces P, Q and $i \geq 0$,

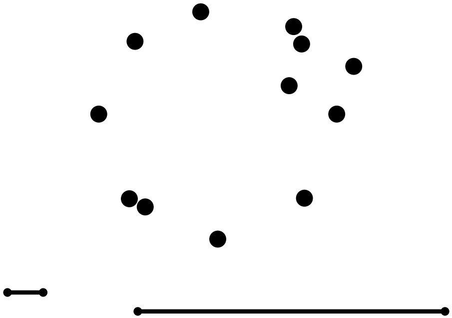
$$d_b(\mathcal{B}_i(\text{Rips}(P)), \mathcal{B}_i(\text{Rips}(Q))) \leq d_{GH}(P, Q).$$



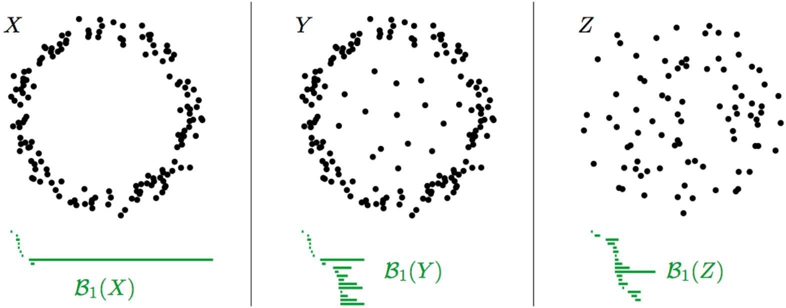
Rips Stability Theorem

Theorem [Edelsbrunner et al. 2007, Chazal et al. 2009] For all finite metric spaces P, Q and $i \geq 0$,

$$d_b(\mathcal{B}_i(\mathrm{Rips}(P)), \mathcal{B}_i(\mathrm{Rips}(Q))) \leq d_{GH}(P, Q).$$



Limitations of 1-Parameter Persistence

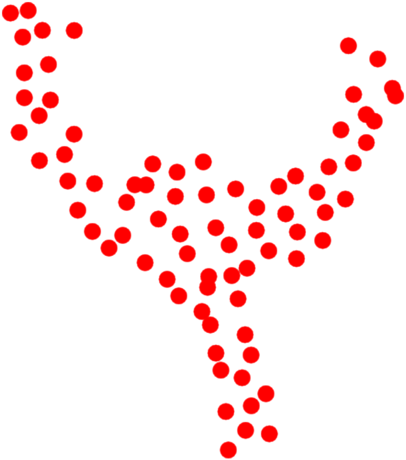


- Persistent homology is **not** stable with respect to outliers,
- Can be insensitive to structure in high density regions of data.

This leads us to 2-parameter persistence [Carlsson / Zomorodian '09].

- 2^{nd} parameter controls how aggressively we remove outliers.

Another Motivation: “Tendrils” in data



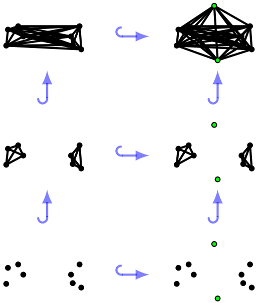
2-Parameter (Multiparameter) Persistent Homology

- Introduced by [Frosini, Mulazzani ’99], [Carlsson, Zomorodian ’09]

Bifiltrations

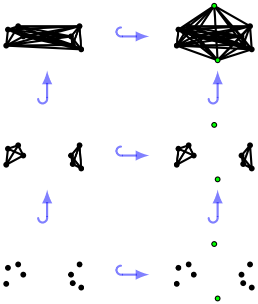
A **bifiltration** is a diagram of simplicial complexes of the form:

$$
\begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \\
\uparrow & & \uparrow & & \uparrow & & \\
F_{1,3} & \hookrightarrow & F_{2,3} & \hookrightarrow & F_{3,3} & \hookrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \\
F_{1,2} & \hookrightarrow & F_{2,2} & \hookrightarrow & F_{3,2} & \hookrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \\
F_{1,1} & \hookrightarrow & F_{2,1} & \hookrightarrow & F_{3,1} & \hookrightarrow & \cdots
\end{array}
$$



A **bifiltration** is a diagram of simplicial complexes of the form:

$$\begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \\
\downarrow & & \downarrow & & \downarrow & & \\
F_{1,3} & \hookrightarrow & F_{2,3} & \hookrightarrow & F_{3,3} & \hookrightarrow & \cdots \\
\downarrow & & \downarrow & & \downarrow & & \\
F_{1,2} & \hookrightarrow & F_{2,2} & \hookrightarrow & F_{3,2} & \hookrightarrow & \cdots \\
\downarrow & & \downarrow & & \downarrow & & \\
F_{1,1} & \hookrightarrow & F_{2,1} & \hookrightarrow & F_{3,1} & \hookrightarrow & \cdots
\end{array}$$



Bipersistence Modules

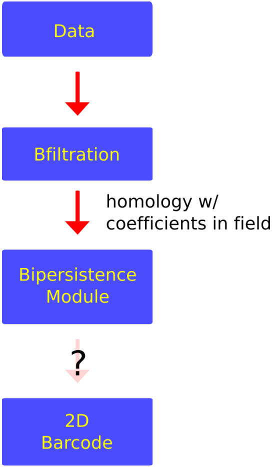
A **bipersistence module** M is a commutative diagram of k-vector spaces of the form:

$$
\begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \\
\uparrow & & \uparrow & & \uparrow & & \\
M_{1,3} & \longrightarrow & M_{2,3} & \longrightarrow & M_{3,3} & \longrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \\
M_{1,2} & \longrightarrow & M_{2,2} & \longrightarrow & M_{3,2} & \longrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \\
M_{1,1} & \longrightarrow & M_{2,1} & \longrightarrow & M_{3,1} & \longrightarrow & \cdots
\end{array}
$$

For $a, b \in \mathbb{N}^2$, write $a \leq b$ if $a_i \leq b_i$ for $i = 1, 2$.

Note: If $a \leq b$, we have a unique map $M_a \to M_b$.

Pipeline for 2-Parameter Persistence



Data → Bifiltration

Three density-sensitive constructions based on Rips complexes.

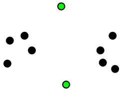
1. density-Rips bifiltration [Carlsson and Zomorodian '09].

Two **parameter-free** constructions:

2. degree-Rips bifiltration [L., Wright '15]
3. subdivision-Rips bifiltration [Sheehy, '12]

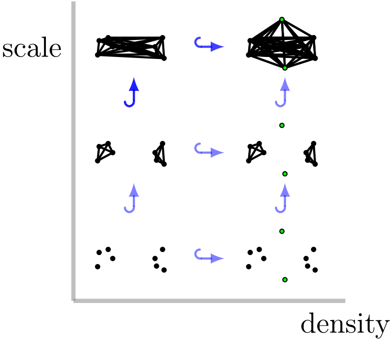
For P a metric space and $r > 0$, define a **density function** $\gamma_r : P \to \mathbb{R}$ by

$$\gamma_r(x) = \#\text{ other points within distance } r \text{ of } x.$$



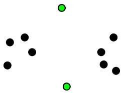
Define the **density-Rips** bifiltration

$$\mathrm{Rips}(\gamma_r)_{(a,b)} := \mathrm{Rips}(\{q \in P \mid \gamma_r(q) \geq a\})_b.$$



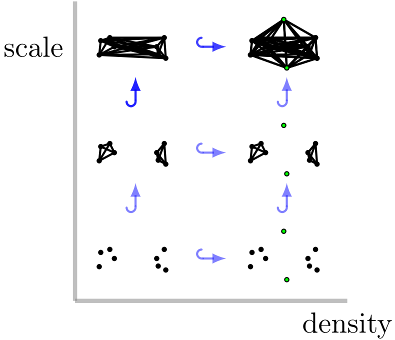
For P a metric space and $r > 0$, define a **density function** $\gamma_r : P \to \mathbb{R}$ by

$$\gamma_r(x) = \#\text{ other points within distance } r \text{ of } x.$$



Define the **degree-Rips** bifiltration

$$:= \mathrm{Rips}(\{q \in P \mid \gamma_b(q) \geq a\})_b.$$

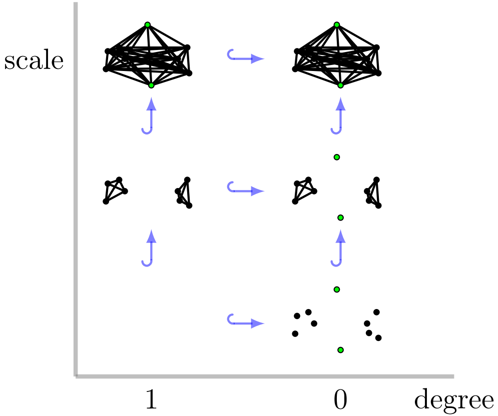


Degree-Rips Bifiltration

For P a metric space, $t, s \geq 0$, let

- $G_{s,t}$ = graph obtained by removing all vertices of degree $< s$ from the t-neighborhood graph of P.
- $\mathrm{DRips}_{s,t}(P)$ be the clique complex on $G_{s,t}$

This defines a bifiltration $\mathrm{DRips}(P)$



Subdivision-Rips Bifiltration

For T a simplicial complex, $\mathrm{Bary}(T)$ is naturally filtered:

- Vertices of $\mathrm{Bary}(T)$ correspond to simplices of T.
- Let T^k be the max. subcomplex of $\mathrm{Bary}(T)$ whose vertices correspond to simplices in T of dimension $\geq k$:

$$\cdots \hookrightarrow T^k \hookrightarrow T^{k-1} \hookrightarrow \cdots \hookrightarrow T^0 = \mathrm{Bary}(T).$$

- Applying this to each space in a **filtration** yields a **bifiltration**.

When filtration is $\mathrm{Rips}(P)$, we'll denote this bifiltration as $\mathrm{SRips}(P)$.

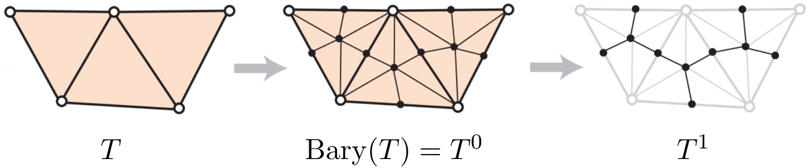
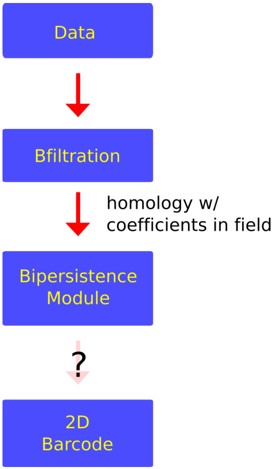


Fig: [Sheehy '12]

Fact: degree-Rips and subdivision-Rips bifiltrations do have (provably) good stability properties w.r.t. outliers [L., Blumberg, in preparation].

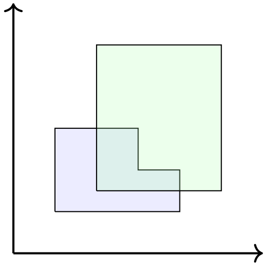
- can formulate using Gromov-Prokhorov metric and **interleavings**.

Pipeline for 2D Persistence



Barcodes of a Bipersistence Module?

Can we define the barcode of bipersistence module as a collection of nice regions in $\mathbb{R}^2$?



Not in any good way.

Non-Existence of a Good Barcode

Def: A collection $\mathcal{B}(M)$ of subsets of $\mathbb{R}^2$ is a **good barcode** for a bipersistence module M if for all $a \leq b \in \mathbb{N}^2$,

$$\text{rank}(M_a \to M_b) = \#\text{ of sets in } \mathcal{B}(M) \text{ containing both } a \text{ and } b.$$

Non-Existence of a Good Barcode

Def: A collection $\mathcal{B}(M)$ of subsets of $\mathbb{R}^2$ is a **good barcode** for a bipersistence module M if for all $a \leq b \in \mathbb{N}^2$,

$$\mathrm{rank}(M_a \to M_b) = \#\text{ of sets in } \mathcal{B}(M) \text{ containing both } a \text{ and } b.$$

Fact: $\exists$ a bipersistence module M with no good barcode.

$$\begin{array}{ccccccc}
0 & \longrightarrow & k & \xrightarrow{=} & k & \longrightarrow & 0 \\
\uparrow & & \twoheaduparrow & & \twoheaduparrow & & \uparrow \\
k & \xhookrightarrow{f} & k^2 & \xrightarrow{=} & k^2 & \twoheadrightarrow & k \\
\uparrow & & \uparrow g & & \uparrow = & & \uparrow = \\
0 & \longrightarrow & k & \xhookrightarrow{g} & k^2 & \twoheadrightarrow & k \\
\uparrow & & \uparrow & & \uparrow h & & \uparrow \\
0 & \longrightarrow & 0 & \longrightarrow & k & \longrightarrow & 0
\end{array}$$

$$f = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; g = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; h = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

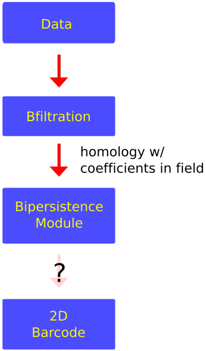
Krull-Schmidt Theorem: For M a finitely generated bipersistence module, $\exists$ collection of indecomposables $M_1, \ldots M_k$, unique up to iso., such that:

$$M \cong \oplus_{i=1}^{k} M_i.$$

The set of possible iso. types of the M_i is **quite complicated**.

Also, indecomposabes are **not** stable: They can merge and split upon small perturbations.

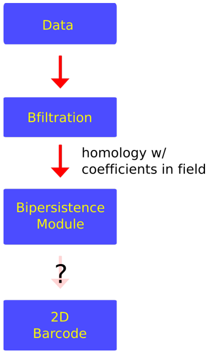
How to work with 2-parameter persistence without a good definition of a barcode?



Ways forward:

1. Work directly with the modules, in a barcode free-way, w/ **metrics** on persistence modules.

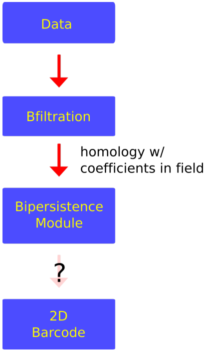
How to work with 2-parameter persistence without a good definition of a barcode?



Ways forward:

1. Work directly with the modules, in a barcode free-way, w/ **metrics** on persistence modules.
2. Consider invariants that serve as a surrogate for the barcode.

How to work with 2-parameter persistence without a good definition of a barcode?



Ways forward:

1. Work directly with the modules, in a barcode free-way, w/ **metrics** on persistence modules.
2. Consider invariants that serve as a surrogate for the barcode.
3. Focus on cases where indecomposables are nice: Interlevel set persistence.

question : How do we concretely represent a bipersistence module?

one answer : a minimal presentation.

- Specifies the isomorphism type of a bipersistence module,
- A matrix with coefficients in k, together with a a $\mathbb{Z}^2$-label for each row and each column.
- Matrix is not unique,
- An efficient way to represent a module for further computation/analysis,
- Computable in practice!

Free Modules, Resolutions, Presentations, and Betti Numbers

Preliminaries

A **morphism** $f : M \to N$ of bipersistence modules is a collection of linear maps

$$\{f_z : M_z \to N_z\}_{z \in \mathbb{Z}^2}$$

such that the following commutes for all $y \leq z \in \mathbb{Z}^2$:

$$\begin{array}{ccc} M_y & \xrightarrow{M_{y,z}} & M_z \\ {\scriptstyle f_y}\downarrow & & \downarrow{\scriptstyle f_z} \\ N_y & \xrightarrow[N_{y,z}]{} & N_z. \end{array}$$

1-D Case:

$$\begin{array}{lccccccc} M: & M_0 & \longrightarrow & M_1 & \longrightarrow & M_2 & \longrightarrow & \cdots \\ & \downarrow{\scriptstyle f_0} & & \downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle f_2} & & \\ N: & N_0 & \longrightarrow & N_1 & \longrightarrow & N_2 & \longrightarrow & \cdots \end{array}$$

With this definition of morphism, persistence modules form a category.

Fact : This is equivalent to the category of **bigraded** $k[x, y]$-modules.

With this definition of morphism, persistence modules form a category.

Fact : This is equivalent to the category of **bigraded** $k[x, y]$-modules.

- Recall: A $k[x, y]$-module M has the structure of a vector space.
- We say M is **bigraded** if, as a vector space, it is equipped with a decomposition:

  $$M \cong \oplus_{z \in \mathbb{Z}^2} M_z$$

  such that

  $$x(M_{z_1,z_2}) \subset M_{z_1+1,z_2}, \qquad y(M_{z_1,z_2}) \subset M_{z_1,z_2+1}.$$

- Morphisms of bigraded $k[x, y]$-modules are module homomorphisms $f : M \to N$ with $f(M_z) \subset N_z$.

Direct sums of persistence modules are constructed pointwise:

$$M: \quad M_0 \xrightarrow{f_0} M_1 \xrightarrow{f_1} M_2 \xrightarrow{f_2} \cdots$$

$$N: \quad N_0 \xrightarrow{g_0} N_1 \xrightarrow{g_1} N_2 \xrightarrow{g_1} \cdots$$

$$M \oplus N: \quad M_0 \oplus N_0 \xrightarrow{f_0 \oplus g_0} M_1 \oplus N_1 \xrightarrow{f_1 \oplus g_1} M_2 \oplus N_2 \xrightarrow{f_2 \oplus g_2} \cdots$$

Free bipersistence modules:

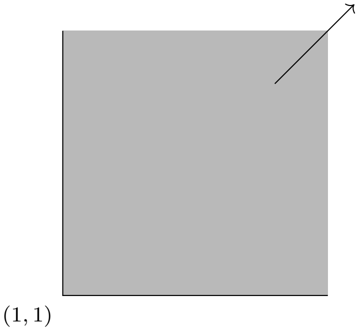
- Needed to define presentations, resolutions, Betti numbers,
- Algebra of these looks a lot like ordinary linear algebra!

For $c \in \mathbb{Z}^2$, define the **quadrant** bipersistence module $\mathcal{Q}^c$ by

$$\mathcal{Q}^c_a = \begin{cases} K & \text{if } a \geq c, \\ 0 & \text{otherwise.} \end{cases} \qquad \mathcal{Q}^c_{a,b} = \begin{cases} \mathrm{Id}_K & \text{if } a \geq c, \\ 0 & \text{otherwise.} \end{cases}$$

$Q^{(1,1)}$:

$$\begin{array}{ccccccccc}
\vdots & & \vdots & & \vdots & & \vdots & & \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & k & \longrightarrow & k & \longrightarrow & k & \longrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & k & \longrightarrow & k & \longrightarrow & k & \longrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & k & \longrightarrow & k & \longrightarrow & k & \longrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \cdots
\end{array}$$

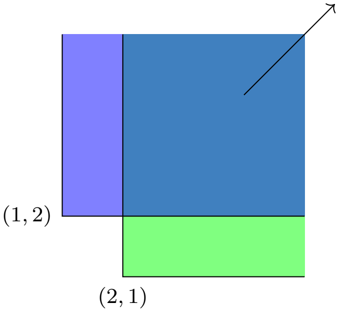


Exercise: For which $y, z \in \mathbb{Z}^2$ does there exists a non-zero morphism $\mathcal{Q}^y \to \mathcal{Q}^z$?

A **free** module is one isomorphic to a direct sum of quadrant modules.

$Q^{(2,1)} \oplus Q^{(1,2)}$:

$$
\begin{array}{ccccccccc}
\vdots & & \vdots & & \vdots & & \vdots & & \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & k & \longrightarrow & k^2 & \longrightarrow & k^2 & \longrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & k & \xrightarrow{\binom{1}{0}} & k^2 & \longrightarrow & k^2 & \longrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow \binom{0}{1} & & \uparrow & & \\
0 & \longrightarrow & 0 & \longrightarrow & k & \longrightarrow & k & \longrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \cdots
\end{array}
$$



(Maps between equal vector spaces are identities).

Bases of Free modules

Let M be any bipersistence module, G be a set of vectors in M.

We say G **generates** M if every vector in M is a linear combination of elements of G.

Bases of Free modules

Let M be any bipersistence module, G be a set of vectors in M.

We say G **generates** M if every vector in M is a linear combination of elements of G.

A **basis** for a free module F is a minimal generating set.

$$
\begin{array}{ccccccccc}
\vdots & & \vdots & & \vdots & & \vdots & & \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \to & k & \to & k^2 & \to & k^2 & \to & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \to & k & \xrightarrow{\binom{1}{0}} & k^2 & \to & k^2 & \to & \cdots \\
\uparrow & & \uparrow & & \uparrow\binom{0}{1} & & \uparrow & & \\
0 & \to & 0 & \longrightarrow & k & \longrightarrow & k^2 & \to & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \to & 0 & \longrightarrow & 0 & \longrightarrow & 0 & \to & \cdots
\end{array}
$$

Example: For $Q^{(2,1)} \oplus Q^{(1,2)}$,

- $\{1^{(2,1)}, 1^{(1,2)}\}$ is a basis.
- $\{1^{(2,1)}, 1^{(1,2)}, (1,1)^{(2,2)}\}$ is a generating set, not a basis.
- $\{1^{(2,1)}\}$ is not a generating set.