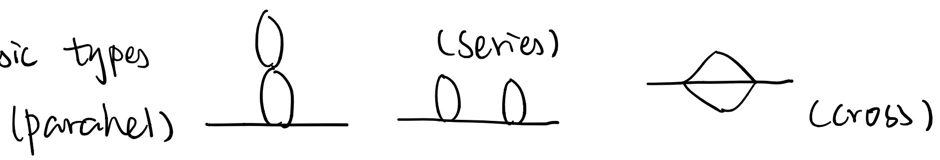
VII3. Extended Persistence

Recall: K = simplicial complex. $f: K \to \mathbb{R}$ is an increasing function, that is, $\tau \leq \sigma \Rightarrow f(\tau) \leq f(\sigma)$. If $a_1 < a_2 < \cdots < a_n$ are the distinct function values of f, we get a filtration $\phi = K_0 \subset K_1 \subseteq \cdots \subseteq K_n = K$ where $K_i = f^{-1}(-\infty, a_i]$. The inclusion induces $\cdots \to H_p(K_i) \to H_p(K_{i+1}) \to \cdots$ for $i = 0, \ldots, n$, for all $p \in \mathbb{Z}$. We computed the persistence diagram of the persistence vector space $\{H_p(K_i)\}$ by reducing the boundary matrix ∂. (The lowest 1's in the reduced matrix R corresponds to the points in the persistence diagram.)

Motivation of extended persistence

Pairwise relations between intramolecular contacts can affect the function and property of a molecule. (e.g. folding rate)

There are three basic types



which we want to distinguish. They are all homotopic, thus cannot be distinguished by homology. The extended persistence homology recovers the construction of the molecules.

Moreover, we get useful duality and symmetry theorems from it.

Basic notions

Let M be a d-manifold, $f: M \to \mathbb{R}$ be a Morse function. (e.g. the height function), $a_1 < a_2 < \cdots < a_n$ be the homological critical values. Insert b_i s.t. $b_0 < a_1 < b_1 < a_2 < \cdots < a_n < b_n$. Then $M_{b_i} := f^{-1}(-\infty, b_i]$ is homotopic to M_{a_i}. Let $M^{b_i} := f^{-1}[b_i, \infty)$. We get inclusions

$\phi = M_{b_0} \subseteq M_{b_1} \subseteq \cdots \subseteq M_{b_n} = M = (M, \phi) = (M, M^{b_n}) \subseteq (M, M^{b_{n-1}}) \subseteq \cdots \subseteq (M, M^{b_0} = M)$.

This induces a persistence vector space

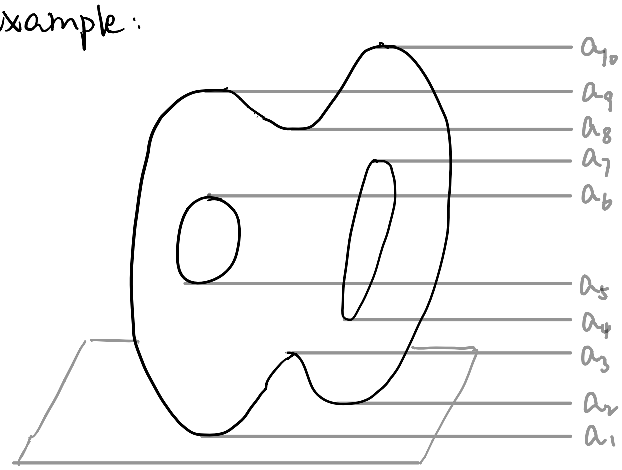
$$0 = H_p(M_{b_0}) \to \cdots \to H_p(M_{b_n}) = H_p(M) = H_p(M, M^{b_n}) \to \cdots \to H_p(M, M^{b_0}) = 0 ,$$

whose persistence diagram contains only finite points since everything getting born eventually dies (ended with 0).

To visulize the relative homology, we have the following theorem:
Consider the filtration $\phi = M^{b_n} \subseteq M^{b_{n-1}} \subseteq \cdots \subseteq M^{b_0} = M$, with the filtration $(M, M^{b_n}) \subseteq (M, M^{b_{n-1}}) \subseteq \cdots \subseteq (M, M^{b_0}) = (M, M)$. call a class $\gamma \in H_p(M^{b_i})$ <u>essential</u> if $\mathrm{pers}(\gamma) = \infty$.

Thm. (1) A dimension p homology class of M^b dies at the same time that a $p+1$ relative homology class of (M, M^b) dies.
(2) An inessential dimension p homology class of M^b gets born at the same time that a dimension $p+1$ relative homology class of (M, M^b) gets born.
(3) An essential dimension p homology class of M^b gets born at the same time that a dimension p relative homology class dies.

Instead of giving a proof, we give an example:
M = 2-torus. f = height function on M.
In the ordinary persistence, a_1, a_2 give birth to classes in H_0, with the latter class killed by a_3. a_4, a_5, a_6, a_7, a_8 give birth to classes in H_1, with

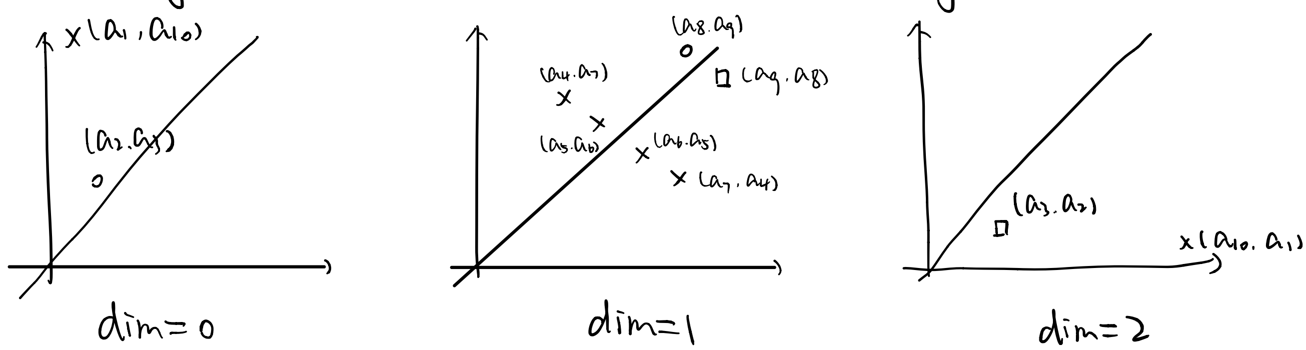


the class created by a_8 dies at a_9. a_{10} gives birth to a 2-cycle. All other classes are essential, i.e. live until the end of the persistence diagram.

Coming down, consider $H_p(M^{b_i})$ and $H_p(M, M^{b_i})$. By the theorem, a_{10} gives birth to an essential homology class in $H_0(M^{a_9})$, so a_{10} kills the 0-cycle created by a_1. a_9 gives birth to an inessential 0-cycle in $H_0(M^{a_9})$, so a_9 gives birth to a 1-relative cycle that killed by a_8. Similarly, by thm (3), a_7, a_6, a_5, a_4 kill the classes in H_1 created by a_4, a_5, a_6, a_7 in the ascending path. a_3 gives birth to a 2-cycle, which is killed by a_2.

Finally, a_1 kills the 2-cycle created by a_{10}.

So we get the extended persistence diagram:



Duality and symmetry theorems

Def. Call the homology classes born and die going up ordinary; Call those born going up and dies coming down extended; Call those born and die coming down relative. Denote the subdiagrams $\mathrm{Ord}(f)$, $\mathrm{Ext}(f)$ and $\mathrm{Rel}(f)$ respectively.

In the above diagram, we used symbols o, x, □ to indicate points in Ord, Ext and Rel respectively. It's easy to see that there are symmetries of these subdiagrams in complementary dimensions. This is not a coincidence:

Thm (Persistence Duality Theorem) Let M be a d-manifold. $f: M \to \mathbb{R}$ be a Morse function. Then $\mathrm{Ord}_p(f) = \mathrm{Rel}^T_{d-p}(f)$, $\mathrm{Ext}_p(f) = \mathrm{Ext}^T_{d-p}(f)$, $\mathrm{Rel}_p(f) = \mathrm{Ord}^T_{d-p}(f)$. where the superscript "T" is used to indicate the reflection along the main diagonal.

Equivalently. $\mathrm{Dgm}_p(f) = \mathrm{Dgm}^T_{d-p}(f)$.

Pf: Idea: Use Lefschetz duality $H_p(M_{b_i}) \cong H_p(M, M^{b_i})$.

Thm (Persistence Symmetry Theorem) With the same notations, $\mathrm{Ord}_p(f) = \mathrm{Ord}^R_{d-p-1}(-f)$. $\mathrm{Ext}_p(f) = \mathrm{Ext}^0_{d-p}(-f)$. $\mathrm{Rel}_p(f) = \mathrm{Rel}^R_{d-p+1}(-f)$.

where the superscript "0" is used to indicated the reflection w.r.t. the origin $((a,b) \mapsto (-a,-b))$ and "R" is used to indicate the reflection along the minor diagonal. $((a,b) \mapsto (-b,-a))$

Pf: Idea: Use the previous two theorems.

Computation of Extended Persistence

Recall that we computed the ordinary persistence by reducing the boundary matrix ∂. In the extended case, we also reduce a matrix, but reducing an augemented one.

Recall that $H_p(X, A) = \tilde{H}_p(X \cup CA)$ for $A \subseteq X$, where $CA := A \times I / A \times 1$ is the cone over A. Therefore, the extended persistence homology groups of a simplicial complex K are the ordinary persistence homology groups of CK, with the last half being the reduced homology. Since CK is contractible, we also go from the trivial group to trivial group. Actually, the boundary matrix of CK can be obtained from that of K. To make these precise, we recall the lower & upper star filtrations:

Let K = triangulation of a d-manifold M. Assign different values (heights) to each vertex, say $f(v_1) < f(v_2) < \cdots < f(v_n)$.

Let $f: |K| \to \mathbb{R}$ be the piecewise linear extension.

$K_i :=$ the full subcomplex of the first i vertices

$K^i :=$ - - - - - - - - - - last $n-i$ - - -.

Let v_i be a vertex, define the lower & upper star of v_i:

$St_- v_i := \{\sigma \in St\, v_i \mid x \in \sigma \Rightarrow f(x) \le f(v_i)\}$.

$St^+ v_i := \{\sigma \in St\, v_i \mid x \in \sigma \Rightarrow f(x) \ge f(v_i)\}$.

Then $K_i = K_{i-1} \cup St_- v_i = \dot{\bigcup}_{j=1}^{i} St_- v_j$. $K^i = K^{i+1} \cup St^+ v_{i+1} = \dot{\bigcup}_{j=1}^{i} St^+ v_j$.

The extended persistence homology groups are

$$\phi = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K = (K, K^{n+1}) \subseteq (K, K^n) \subseteq \cdots \subseteq (K, K^0) = (K, K)$$

Note that each $K_i - K_{i-1}$ can now consist of more than one simplices. Within each $K_i - K_{i-1}$, order the simplices arbitrarily, as long as in non-decreasing dimension.

Let A be the boundary matrix of the filtration in this order, $K_0 \subseteq K_1 \subseteq \cdots \subseteq K$, i.e. $A[i,j] = 1 \Leftrightarrow \sigma_i$ is a codimension-1 face of σ_j.

B be the boundary matrix of the filtration

$K^{n+1} \subseteq K^n \subseteq \cdots \subseteq K$, with the simplices ordered differently: The simplices of K^n go first, then $K^{n-1} - K^n$, still with non-decreasing dimension, then $K^{n-2} - K^{n-1}$, ..., and so on. Then B is also upper-triangular.

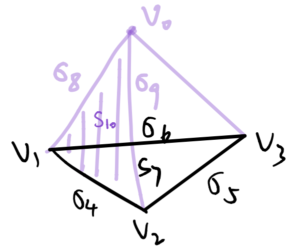
Let P be the permutation matrix connecting the two orderings.

Let $\partial = \begin{bmatrix} A & P \\ 0 & B \end{bmatrix}$ be the augmented upper-triangular matrix.

We claim that ∂ is the boundary matrix of CK.

Actually, add a dummy vertex v_0, it is easy to see that $cK = K \cup \{(v_0\sigma) \mid \sigma \in K\}$ where $(v_0\sigma)$ is the $p+1$-simplex with vertex v_0 and bottom σ. The part $\begin{bmatrix} A \\ 0 \end{bmatrix}$ of ∂ represents the boundary of the bottom K of cK, and the columns of $\begin{bmatrix} P \\ B \end{bmatrix}$ represents $\{(v_0\sigma)\}$ in cK, with the σ's ordered as in B.

$$\begin{array}{c c} & \begin{array}{cccccc} St_-v_1 & \cdots & St_-v_n & St^+v_n & \cdots & St^+v_1 \end{array} \\ \partial = \begin{array}{c} St_-v_1 \\ \vdots \\ St_-v_n \\ St^+v_n \\ \vdots \\ St^+v_1 \end{array} & \begin{bmatrix} A & P \\ 0 & B \end{bmatrix} \end{array}$$



For a simplex $\tau \in cK - K$, $\tau = (v_0\sigma)$ for some σ. Then $\partial\tau = \sigma + (v_0\,\partial\sigma)$. We want to show that the i-th column of $\begin{bmatrix} P \\ B \end{bmatrix}$ gives the boundary of τ, where τ is the i-th simplex in the order of B. By the order of simplices in ∂, the part $(v_0\partial\sigma)$ in $\partial\tau$ is contained in B and the part σ is contained in P. Suppose $\partial\sigma = \sum_j \alpha_j$. then $B[j,i]=1$ by definition of B. This is just saying that B gives the $(v_0\partial\sigma)$ part of $\partial\tau$. Now let k be the order of τ in the order of A. then $P[i,k]=1$ by definition of P. This is just saying that P gives the σ part of $\partial\tau$. (You may think of S_{10} in the figure as τ. $\partial S_{10} = \partial(v_0\sigma_4) = \sigma_4 + \sigma_8 + \sigma_9 = \sigma_4 + v_0(\partial\sigma_4)$).

As a result, by reducing ∂, we get the extended persistence diagram of K. The ordinary, extended and relative subdiagrams correspond to the lowest 1's at the upper-left, the upper-right and the lower-right of the reduced ∂, respectively.

VII 4. Spectral Sequence

In this Section. we only consider the ordinary persistence. Spectral sequence is related to estimating the homology group. There is a natural spectral sequence once a filtrated chain complex is given. We'll state the spectral sequence theorem, which gives the relation between the spectral sequence and the persistence diagram. Then we explain how to compute the spectral sequence.

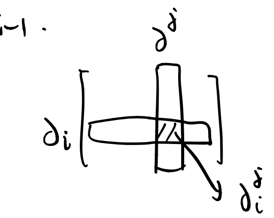
Notations

Start with a filtration of simplicial complexes $\phi = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$

$k_i := \text{card}\, K_i$, # of simplices in K_i. $\partial = [\partial^1\ \partial^2 \cdots \partial^n]$ is the boundary matrix, where ∂^j stores the boundaries of the $k_{j-1}+1$ to k_j simplices. ∂_i denote the $k_{i-1}+1$ to k_i rows of ∂.

∂_i^j denote the intersection of ∂^j and ∂_i, i.e. the codimension 1 faces of simplices in $K_j - K_{j-1}$ that lies in $K_i - K_{i-1}$.

$\partial_i^j = 0$ whenever $i > j$, since ∂ is upper-triangular.



$C_p := C_p(K)$, the group of p-chains in K.

$C_p^j := C_p(K_j)/C_p(K_{j-1}) = C_p(K_j - K_{j-1})$, the groups of p-chains in $K_j - K_{j-1}$.

For $c \in C_p^j$. Let $\partial_i^j c$ be the sum of terms in ∂c that lies in $K_i - K_{i-1}$. We have $\partial c = \partial_j^j c + \partial_{j-1}^j c + \cdots + \partial_i^j c$. There is a natural 1-1 correspondence between the blocks ∂_i^j in ∂ and the restricted boundary maps ∂_i^j.

Construction

In our construction of the spectral sequence $\{E^r_{p,q}, d^r_{p,q}\}$, one should bear in mind that p denotes the column block index. $p+q$ denotes the dimension and r is what we call "phase".

E^0-term. $E^0_{p,q} := C^p_{p+q}$, the $(p+q)$-chains in $K_p - K_{p-1}$. Varying $q \in \mathbb{Z}$, these groups are generated by the columns of ∂^p, that is simplices of $K_p - K_{p-1}$. $E^0_{p,q} \cong C_{p+q}(K_p, K_{p-1})$.

Let $d^0_{p,q}: E^0_{p,q} \to E^0_{p,q-1}$, given by the boundary map omitting the faces in K_{p-1}. I.e. $d^0_{p,q}\sigma = \partial^p_p \sigma$. Clearly $d^0_{p,q-1} \circ d^0_{p,q} = 0$. So $\cdots \to E^0_{p,q+1} \to E^0_{p,q} \to E^0_{p,q-1} \to \cdots$ is a chain complex for $p = 1, \ldots, n$.

$$\begin{array}{cccccc}
 & \vdots & \vdots & \vdots & \\
\cdots & E^0_{1,1} & E^0_{2,1} & E^0_{3,1} & \cdots \\
 & \downarrow & \downarrow & \downarrow & \\
\cdots & E^0_{1,0} & E^0_{2,0} & E^0_{3,0} & \cdots \\
 & \downarrow & \downarrow & \downarrow & \\
\cdots & E^0_{1,-1} & E^0_{2,-1} & E^0_{3,-1} & \cdots \\
 & \downarrow & \downarrow & \downarrow & \\
\cdots & 0 & E^0_{2,-2} & E^0_{3,-2} & \cdots \\
 & \downarrow & \downarrow & \downarrow & \\
\cdots & 0 & 0 & E^0_{3,-3} & \cdots \\
 & \vdots & & \vdots &
\end{array}$$

E^1-term. Let $E^1_{p,q} := H_{p+q}(E^0_p) = \ker(d^0_{p,q}) / \operatorname{im}(d^0_{p,q+1}) \cong H_{p+q}(K_p, K_{p-1})$.

An element of $E^1_{p,q}$ is an equivalence class of a chain $c \in C^p_{p+q}$ s.t. $\partial^p_p c = 0$, i.e. c has zero boundary in ∂_p.

Let $d^1_{p,q}: E^1_{p,q} \to E^1_{p-1,q}$ be the boundary map restricted to ∂_{p-1}. I.e. $d^1_{p,q}(c) = \partial^p_{p-1} c$. Note that $c \in E^1_{p,q}$ means $\partial^p_p c = 0$, but c possibly has nonzero boundary in ∂_{p-1}, so $d^1_{p,q}(c)$ maybe nonzero.

We get a chain complex $\cdots \to E^1_{p+1,q} \to E^1_{p,q} \to E^1_{p-1,q} \to \cdots$

$$
\begin{array}{ccccccc}
 & \vdots & & \vdots & & \vdots & \\
\cdots & E^1_{1,1} & \leftarrow & E^1_{2,1} & \leftarrow & E^1_{3,1} & \cdots \\
\cdots & E^1_{1,0} & \leftarrow & E^1_{2,0} & \leftarrow & E^1_{3,0} & \cdots \\
\cdots & E^1_{1,-1} & \leftarrow & E^1_{2,-1} & \leftarrow & E^1_{3,-1} & \cdots \\
\cdots & 0 & \leftarrow & E^1_{2,-2} & \leftarrow & E^1_{3,-2} & \cdots \\
\cdots & 0 & \leftarrow & 0 & \leftarrow & E^1_{3,-3} & \cdots \\
 & \vdots & & & & \vdots &
\end{array}
$$

E^2-term Let $E^2_{p.q} := \ker d^1_{p.q} / \operatorname{im} d^1_{p+1.q} = H_{p+q}(E^1_{p.q})$.

An element of $E^2_{p.q}$ is represented by the sum of two chains $c \in C^p_{p+q}$ and $c' \in C^{p-1}_{p+q}$ s.t. $\partial^p_p c = 0$ and $\partial^p_{p-1} c + \partial^{p-1}_{p-1} c' = 0$.

$([c] \in \ker d^1_{p.q} \Leftrightarrow \partial^p_{p-1} c \in \operatorname{im} d^0_{p-1.q+1} \Leftrightarrow \exists\, c' \in C^{p-1}_{p+q}$ s.t. $\partial^p_{p-1} c = \partial^{p-1}_{p-1} c')$

Moreover, $E^2_{p.q} \oplus E^1_{p-1.q+1} \cong H_{p+q}(K_p . K_{p-2})$. ($E^2_{p.q}$ gives the $K_p - K_{p-1}$ part and $E^1_{p-1.q+1}$ gives the $K_{p-1} - K_{p-2}$ part)

Let $d^2_{p.q} : E^2_{p.q} \to E^2_{p-2.q+1}$ be the $(p+q)$-th boundary map restricted to ∂_{p-2}. I.e. $d^2_{p.q}(c) = \partial^p_{p-2}(c)$ for $\partial^p_p c + \partial^p_{p-1} c = 0$.

The boundary of c in $K_p - K_{p-2}$ is empty, but is possibly nonempty in K_{p-2}. Similarly we get a chain complex

$$\cdots \to E^2_{p+2.q-1} \to E^2_{p.q} \to E^2_{p-2.q+1} \to \cdots$$

$$
\begin{tikzcd}
 & \vdots & \vdots & \vdots & & \\
\cdots & E^2_{1,1} & E^2_{2,1} & E^2_{3,1} & E^2_{4,1} & \cdots \\
\cdots & E^2_{1,0} & E^2_{2,0} & E^2_{3,0} \arrow[ull] & E^2_{4,0} \arrow[ull] & \cdots \\
\cdots & E^2_{1,-1} & E^2_{2,-1} & E^2_{3,-1} \arrow[ull] & E^2_{4,-1} \arrow[ull] & \cdots \\
\cdots & 0 & E^2_{2,-2} & E^2_{3,-2} \arrow[ull] & E^2_{4,-2} \arrow[ull] & \cdots \\
\cdots & 0 & 0 & E^2_{3,-3} \arrow[ull] & E^2_{4,-3} \arrow[ull] & \cdots \\
 & \vdots & & & &
\end{tikzcd}
$$

In general, let $E^r_{p,q} := H_{p+q}(E^{r-1}_{p,q})$. $d^r_{p,q}: E^r_{p,q} \to E^r_{p-r,q+r-1}$ be the boundary matrix restricted to ∂_{p-r}.

By results in homological algebra, $E^r =$ some E^∞ when r large enough. and $E^\infty_{p,q} \cong H_{p+q}(K_p)/H_{p+q}(K_{p-1})$.

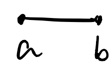
Thm (Spectral Sequence Theorem) The total rank of the images of the differential maps of the spectral sequence when $r \geq 1$ equals # of points in the $(p+q)$-th diagram whose persistence is r: $\sum_{p=1}^{n} \operatorname{rank} A^r_{p,q} = \operatorname{card}\{a \in \operatorname{Dgm}_{p+q}(f) \mid \operatorname{pers}(a) = r\}$. where

$A^r_{p,q} = \operatorname{Im}(d^r_{p+r,q-r+1}: E^r_{p+r,q-r+1} \to E^r_{p,q}) \subseteq E^r_{p,q}$, and q varies with p s.t. $p+q$ remains constant.

In particular, when $r \to \infty$, $\sum_{p=1}^{n} \operatorname{rank} A^r_{p,q}$ gives the homology of K.

[Wrong version? $\sum_{p=1}^{n} \operatorname{rank} E^r_{p,q} = \operatorname{card}\{a \in \operatorname{Dgm}_{p+q}(f) \mid \operatorname{pers}(a) \geq r\}$

Counterexample: $K_1 = \{a, b\}$. $K_2 = \{a, b, (a,b)\}$. $E^1_{2,-1} = \mathbb{Z}_2$,



but no cycle at all in dimension ≥ 1.]

Computation of spectral sequences

We want information about $E^r_{p,q}$, $\ker(d^r_{p,q})$ and $\operatorname{im}(d^r_{p+r,q-r+1})$.

To get it, we do nothing different from reducing the boundary matrix. However, instead of do it column by column, we do it along the diagonal.

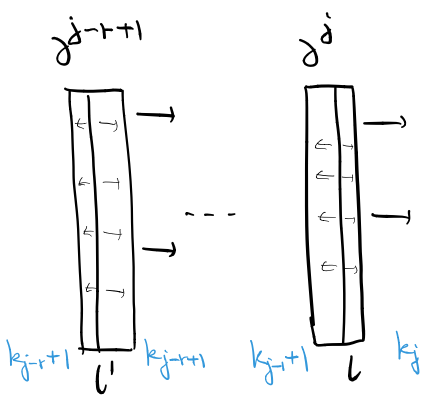
First observe that $E^1_{p,q} \cong H_{p+q}(K_p, K_{p-1})$, an element in $E^1_{p,q}$ is represented by a zero column in the reduced ∂^p_p.

Similarly, an element in $E^2_{p,q}$ is represented by $c \in C^p_{p+q}$

and $c' \in C^{p-1}_{p+q}$ s.t. $\partial^p_p c = 0$, $\partial^p_{p-1} c + \partial^{p-1}_{p-1} c' = 0$. Thus, it suffices to consider the zero columns of ∂^p_p, ∂^{p-1}_{p-1} and ∂^p_{p-1}.

The following algorithm reduces all $\partial^p_p, \partial^p_{p-1}, \dots, \partial^p_{p-r}$ after phase r:

```
for r=1 to n do
    for j=r to n do
        for l=k_{j-1}+1 to k_j do
            while ∃ k_{j-r} < l' < l s.t. k_{j-r} < low(l') = low(l) ≤ k_{j-r+1} do
                add column l' to column l
            endwhile
        endfor
    endfor
endfor
```



As illustrated in the figure, at each phase, r is fixed. For each $j \geq r$, we consider ∂^j and ∂^{j-r+1}, then reduce ∂^j by the columns of ∂^{j-r+1}, by the same method as before. When $r=1$, we are reducing ∂^j by columns of itself.

Recall that $\partial^j_i = 0$ when $i > j$. Thus, after the phase $r=1$, all ∂^j_j are reduced. Since there are all 0's on the left of them; After the phase $r=2$, all ∂^j_{j-1} are reduced; ... and so on, as desired.