Topological Data Analysis

2022–2023

Lecture 5

Persistence Modules

17 November 2022

Let us fix an arbitrary field $\mathbb{F}$.

A <u>persistence module</u> over $\mathbb{F}$ is a pair (V,π) where $V=\{V_t\}_{t\in\mathbb{R}}$ is a collection of $\mathbb{F}$-vector spaces indexed by the real numbers and π is a collection of $\mathbb{F}$-linear maps

$$\pi_{s,t} : V_s \longrightarrow V_t \quad \text{for } s\leq t$$

such that $\pi_{s,t}\circ \pi_{r,s} = \pi_{r,t}$ if $r\leq s\leq t$ and $\pi_{t,t}=\mathrm{id}$ for all t.

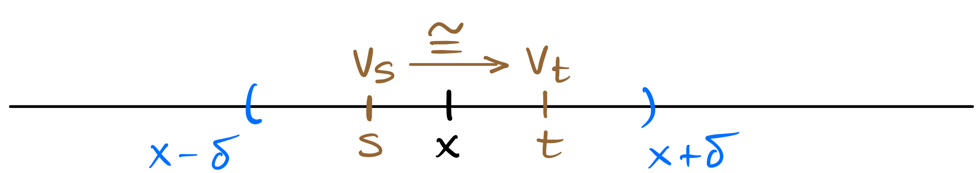
In other words, (V,π) is a <u>functor</u> from $\mathbb{R}$ viewed as an ordered set to the category of $\mathbb{F}$-vector spaces.

A persistence module is of <u>finite type</u> or <u>tame</u> if:

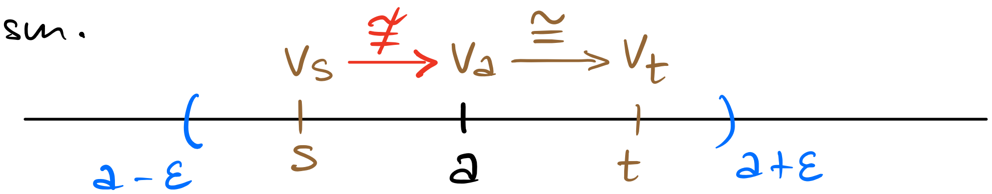
a) $\dim_{\mathbb{F}} V_t$ is finite for all t.

b) There is a finite set $A=\{a_0,\dots,a_n\}\subset\mathbb{R}$ such that:

i) For every $x\in\mathbb{R}\setminus A$ there is a $\delta>0$ such that $\pi_{s,t}$ is an isomorphism for $x-\delta<s\leq t<x+\delta$.



ii) For every $a \in A$ there is an $\varepsilon > 0$ such that if $a \le t < a + \varepsilon$ then $\pi_{a,t}$ is an isomorphism while if $a - \varepsilon < s < a$ then $\pi_{s,a}$ is <u>not</u> an isomorphism.



iii) $V_t = \{0\}$ if $t < a_0$, assuming that $a_0 < a_1 < \dots < a_n$.

The set $A = \{a_0, \dots, a_n\}$ is called the <u>spectrum</u> of (V, π).

Note that $\pi_{t,t}$ is an isomorphism for all $t \in \mathbb{R}$ by (i) or (ii), and therefore $\pi_{t,t} \circ \pi_{t,t} = \pi_{t,t}$ implies that $\pi_{t,t} = id$ for all t.

It also follows from the definition that $\pi_{s,t}$ is an isomorphism whenever $a_n \le s \le t$.

To prove this, let $C = \{ t \ge a_n \mid \pi_{a_n,t} \text{ is an isomorphism} \}$. Then C is an open subset of $[a_n, \infty)$ since if $t \in C$ and $t > a_n$ then $t \notin A$ and therefore there is a $\delta > 0$ such that $(t-\delta, t+\delta) \subset C$, and if $t = a_n$ then there is an $\varepsilon > 0$ such that $[a, a+\varepsilon) \subset C$. Similarly C is closed since $[a_n, \infty) \setminus C$ is open. Hence $C = [a_n, \infty)$ because $[a_n, \infty)$ is connected and $C \neq \emptyset$ as $a_n \in C$. ✓

Example:

Let X be a point cloud and let $R_t(X)$ be the Vietoris–Rips complex for each $t \in \mathbb{R}$, where $R_t(X) = \emptyset$ if $t < 0$.

Then
$$V_t = H_*(R_t(X)) = \bigoplus_{k=0}^{\infty} H_k(R_t(X))$$

is a persistence module over the coefficient field $\mathbb{F}$ with the maps $\pi_{s,t} : V_s \to V_t$ induced by the inclusions $i_{s,t} : R_s(X) \hookrightarrow R_t(X)$ if $s \leq t$.

The relation $\pi_{s,t} \circ \pi_{r,s} = \pi_{r,t}$ if $r \leq s \leq t$ follows from the functoriality of homology, since $i_{s,t} \circ i_{r,s} = i_{r,t}$ if $r \leq s \leq t$.

$$\pi_{r,t} = (i_{r,t})_* = (i_{s,t} \circ i_{r,s})_* = (i_{s,t})_* \circ (i_{r,s})_* = \pi_{s,t} \circ \pi_{r,s}. \checkmark$$

This persistence module is of finite type. Its spectrum is the set of parameter values $\{\varepsilon_0, \dots, \varepsilon_n\}$ where the homology of $R_\varepsilon(X)$ changes.

A <u>morphism</u> $f:(V,\pi)\to(V',\pi')$ of persistence modules over $\mathbb{F}$ is a collection of $\mathbb{F}$-linear maps $f_t : V_t \longrightarrow V'_t$ such that

$$f_t\circ\pi_{s,t} = \pi'_{s,t}\circ f_s \quad \text{whenever } s\le t.$$

$$\begin{tikzcd} V_s \arrow[r, "f_s"] \arrow[d, "\pi_{s,t}"'] & V'_s \arrow[d, "\pi'_{s,t}"] \\ V_t \arrow[r, "f_t"'] & V'_t \end{tikzcd}$$

commutes.

In other words, f is a <u>natural transformation</u> of functors.

Suppose given morphisms $(V,\pi)\xrightarrow{f}(V',\pi')\xrightarrow{g}(V'',\pi'')$.
Then the composite $g\circ f$, which is defined as $(g\circ f)_t = g_t\circ f_t$ for all t, is also a morphism of persistence modules.

$$\begin{tikzcd} V_s \arrow[r, "f_s"] \arrow[d, "\pi_{s,t}"'] & V'_s \arrow[r, "g_s"] \arrow[d, "\pi'_{s,t}"'] & V''_s \arrow[d, "\pi''_{s,t}"] \\ V_t \arrow[r, "f_t"] & V'_t \arrow[r, "g_t"] & V''_t \end{tikzcd}$$

$(g\circ f)_t\circ\pi_{s,t} = g_t\circ f_t\circ\pi_{s,t} =$

$= g_t\circ\pi'_{s,t}\circ f_s = \pi''_{s,t}\circ g_s\circ f_s =$

$= \pi''_{s,t}\circ(g\circ f)_s$ ✓

A morphism $f:(V,\pi)\to(V',\pi')$ of persistence modules is an <u>isomorphism</u> if there is a morphism $g:(V',\pi')\to(V,\pi)$ such that $g\circ f=\mathrm{id}_V$ and $f\circ g=\mathrm{id}_{V'}$.

It follows that f is an isomorphism if and only if f_t is an isomorphism of vector spaces for all t.

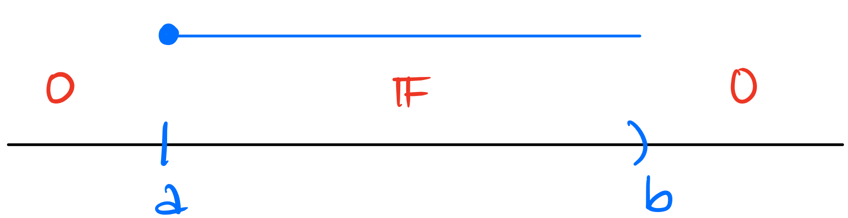
This is left as an exercise.

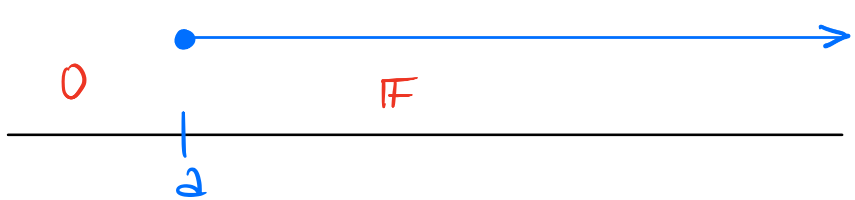
Interval modules

For $I=[a,b)$ or $I=[a,\infty)$, define a persistence module $\mathbb{F}(I)$ as

$$F(I)_t=\begin{cases}\mathbb{F} & \text{if } t\in I\\ 0 & \text{otherwise}\end{cases}$$

with $\pi_{s,t}=\mathrm{id}$ if $s,t\in I$ or $\pi_{s,t}=0$ otherwise.





These are persistence modules of finite type. The spectrum of $\mathbb{F}(I)$ is $\{a, b\}$ if $I = [a, b)$ or $\{a\}$ if $I = [a, \infty)$.

They are called <u>interval modules</u>.

Direct sum

If (V, π) and (V', π') are persistence modules, their <u>direct sum</u> is the persistence module $(V \oplus V', \pi \oplus \pi')$ with

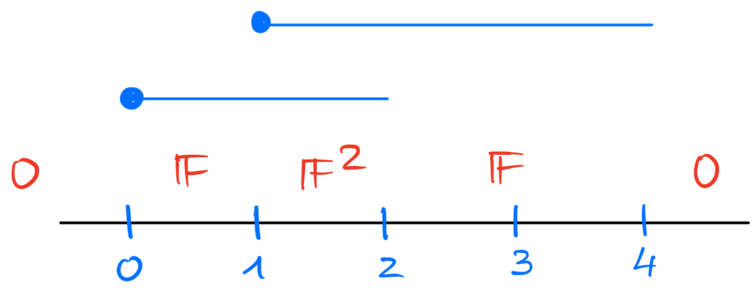
$$\begin{cases} (V \oplus V')_t = V_t \oplus V'_t \\ (\pi \oplus \pi')_{s,t} = \pi_{s,t} \oplus \pi'_{s,t} \end{cases}$$

That is, $(\pi \oplus \pi')_{s,t} (v, v') = (\pi_{s,t}(v), \pi'_{s,t}(v'))$.

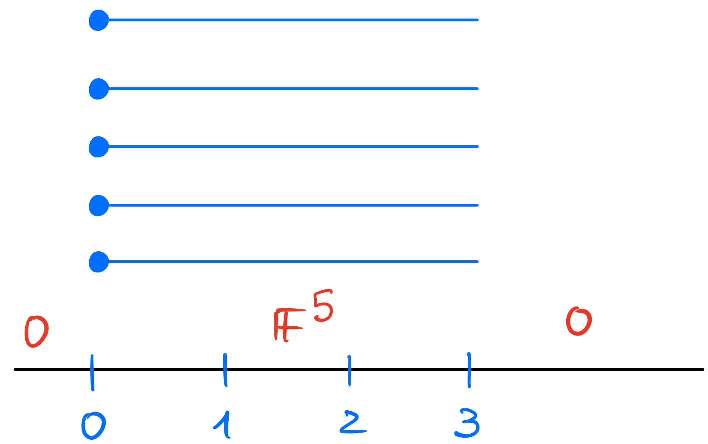
If V and V' are of finite type, then $V \oplus V'$ is also of finite type.

If A is the spectrum of V and A' is the spectrum of V' then the spectrum of $V \oplus V'$ is $A \cup A'$.

For every $m \geq 1$ we denote $\mathbb{F}(I)^m = \mathbb{F}(I) \oplus \overset{m}{\overbrace{\cdots}} \oplus \mathbb{F}(I)$.



$\mathbb{F}[0,2) \oplus \mathbb{F}[1,4)$



$\mathbb{F}[0,3)^5$

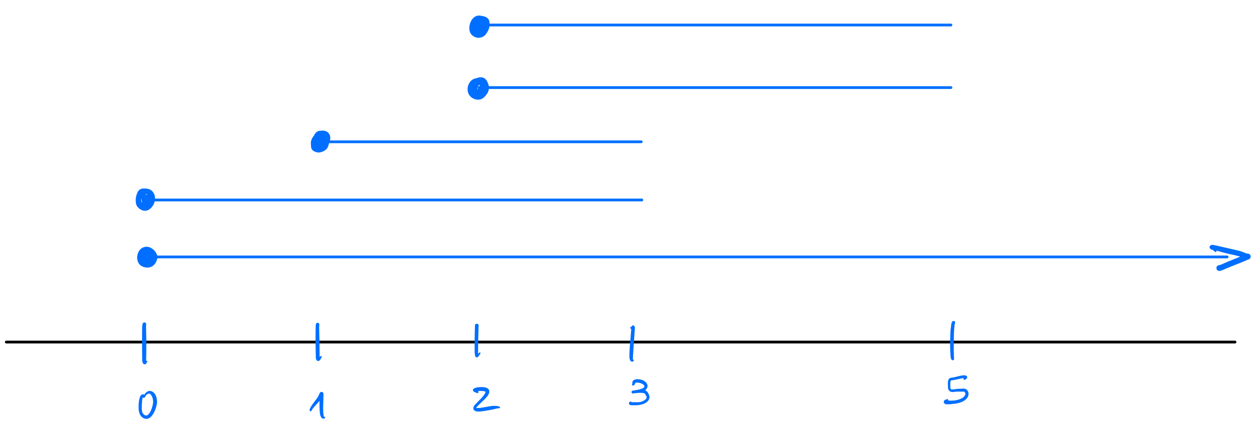
Normal Form Theorem:

For every persistence module V of finite type there is a finite collection of intervals $\{I_1, \dots, I_N\}$ with $I_i = [a_i, b_i)$ or $I_i = [a_i, \infty)$ for every i such that $I_i \neq I_j$ if $i \neq j$ and there is an isomorphism of persistence modules

$$V \cong \mathbb{F}(I_1)^{m_1} \oplus \dots \oplus \mathbb{F}(I_N)^{m_N}$$

with $m_i > 0$ for all i. Moreover, the set $\{I_1, \dots, I_N\}$ and the integers $m_1, \dots, m_N$ are unique.

As a consequence of this fact, every persistence module of finite type yields a unique <u>barcode</u> (up to permutation of bars):

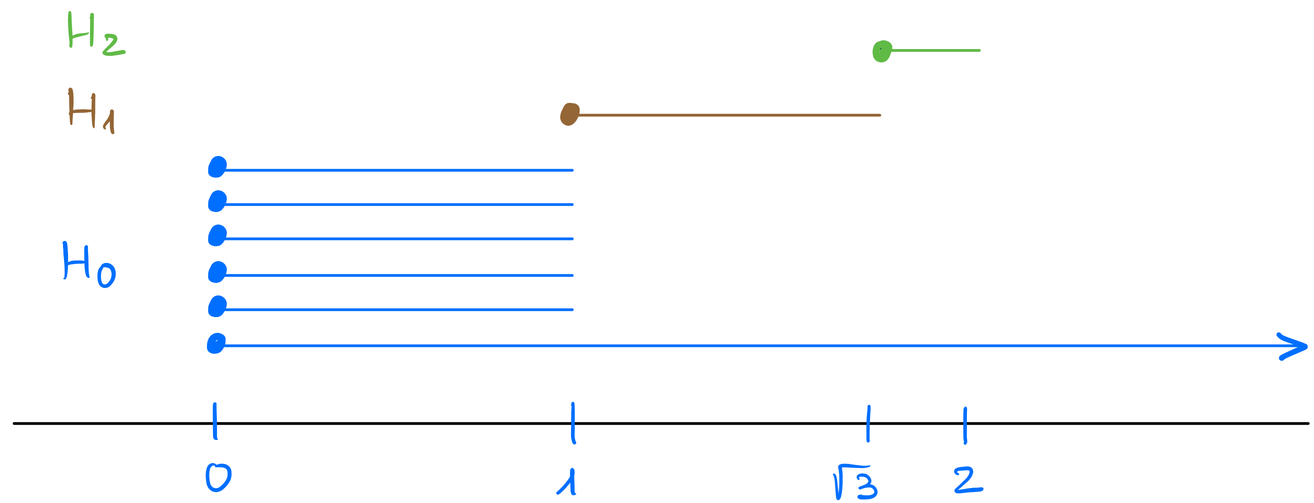


$$\mathbb{F}[0,\infty) \oplus \mathbb{F}[0,3) \oplus \mathbb{F}[1,3) \oplus \mathbb{F}[2,5)^2$$

<u>Example</u>: Let X be the set of vertices of a regular hexagon of radius 1, and let V be the Vietoris–Rips persistence module of X. Thus $V_t = H_*(R_t(X))$ for all $t \in \mathbb{R}$. The spectrum of V is the set

$$A = \{0, 1, \sqrt{3}, 2\}.$$

The <u>persistence</u> <u>barcode</u> of X is the following:



$$\mathbb{F}[0,\infty) \oplus \mathbb{F}[0,1)^5 \oplus \mathbb{F}[1,\sqrt{3}) \oplus \mathbb{F}[\sqrt{3},2)$$

In general, if (V,π) is a persistence module of finite type, a nonzero vector $v \in V_t$ is <u>born</u> at t if $v \notin \operatorname{Im} \pi_{s,t}$ for any $s<t$.

A nonzero vector $v \in V_s$ <u>dies</u> or <u>vanishes</u> at $t>s$ if $\pi_{s,t}(v)=0$ and $\pi_{s,r}(v) \neq 0$ for $s \leq r < t$.

If v is born at $t=b$ and dies at $t=d$, then $d-b$ is its <u>life</u> or <u>persistence</u>. If $\pi_{b,t}(v) \neq 0$ for all $t>b$ then v is <u>permanent</u>.

Shift action

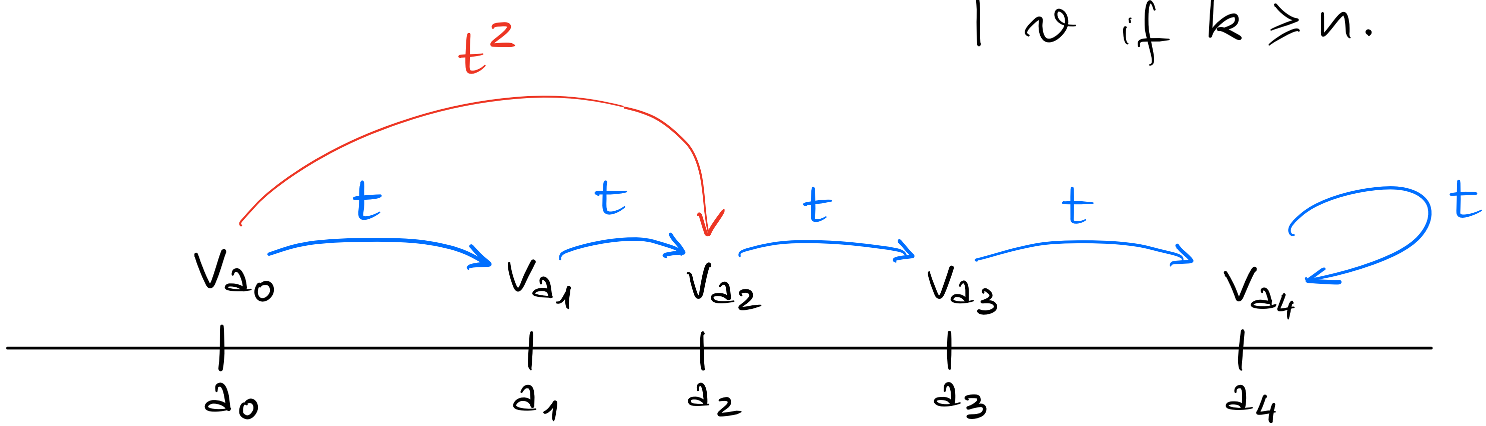
Let (V, π) be a persistence module of finite type with spectrum $A = \{a_0, \dots, a_n\}$.

Consider a graded vector space $V_* = \bigoplus_{k=0}^{\infty} V(k)$, where

$$V(k) = \begin{cases} V_{a_k} & \text{for } 0 \le k \le n \\ V_{a_n} & \text{if } k > n. \end{cases}$$

We turn V_* into a graded module over the graded polynomial ring $\mathbb{F}[t]$ by defining, for $v \in V(k)$,

$$t \cdot v = \begin{cases} \pi_{a_k, a_{k+1}}(v) & \text{if } 0 \le k < n \\ v & \text{if } k \ge n. \end{cases}$$



Since the ring $\mathbb{F}[t]$ is a principal ideal domain, we have the following result. For a graded module M, we denote by ΣM the graded module defined as $(\Sigma M)_k = M_{k-1}$ if $k \geq 1$, and $(\Sigma M)_0 = \{0\}$.

$$M: \quad M_0 \xrightarrow{t} M_1 \xrightarrow{t} M_2 \xrightarrow{t} M_3 \longrightarrow \cdots$$

$$\Sigma M: \quad 0 \xrightarrow{t} M_0 \xrightarrow{t} M_1 \xrightarrow{t} M_2 \longrightarrow \cdots$$

Structure Theorem:

Let M be a finitely generated graded module over $\mathbb{F}[t]$, where $\mathbb{F}$ is a field. Then

$$M \cong \bigoplus_{i=1}^{m} \Sigma^{p_i} \mathbb{F}[t] \oplus \left(\bigoplus_{j=1}^{n} \Sigma^{q_j} \mathbb{F}[t] / (t^{r_j}) \right)$$

for some collection of integers $p_i \geq 0$, $q_j \geq 0$, $r_j \geq 1$. Moreover, this decomposition is unique up to a permutation of summands.

Zomorodian – Carlsson (2005)

The Structure Theorem implies the Normal Form Theorem for persistence modules of finite type using the shift action.

For a persistence module (V, π) of finite type with spectrum $A = \{a_0, \dots, a_n\}$, let V_* be the corresponding graded $\mathbb{F}[t]$-module. Then V_* is finitely generated as an $\mathbb{F}[t]$-module and hence

$$V_* \cong \bigoplus_{i=1}^{m} \Sigma^{p_i} \mathbb{F}[t] \oplus \left(\bigoplus_{j=1}^{n} \Sigma^{q_j} \mathbb{F}[t] / (t^{r_j}) \right).$$

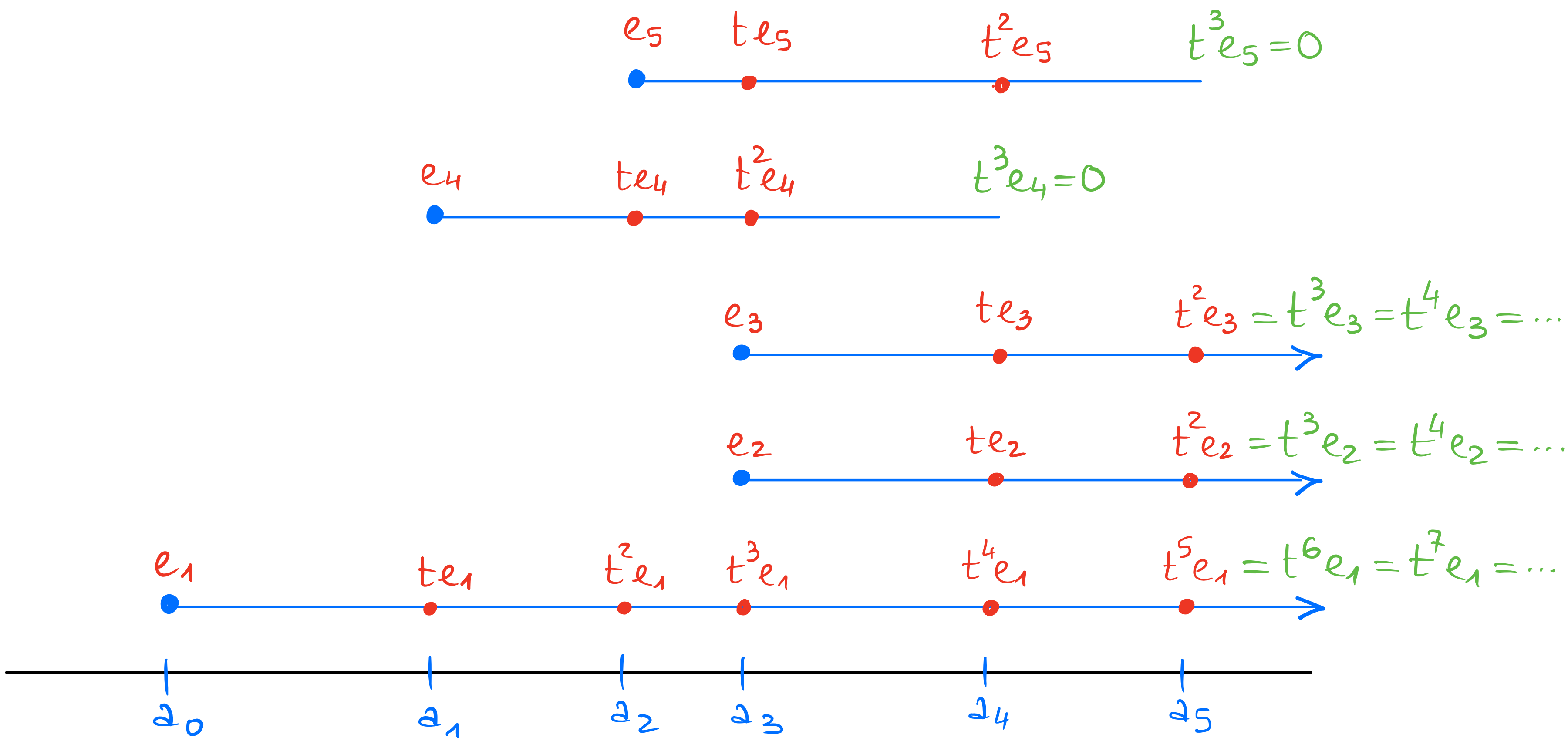
This implies that

$$V = \bigoplus_{i=1}^{m} \mathbb{F}[a_{p_i}, \infty) \oplus \left(\bigoplus_{j=1}^{n} \mathbb{F}[a_{q_j}, a_{q_j + r_j}) \right).$$

Example:

$$V_* \cong \mathbb{F}[t] \oplus \Sigma^3 \mathbb{F}[t] \oplus \Sigma^3 \mathbb{F}[t] \oplus \Sigma\, \mathbb{F}[t]/(t^3) \oplus \Sigma^2\, \mathbb{F}[t]/(t^3)$$

$$V = \mathbb{F}[a_0, \infty) \oplus \mathbb{F}[a_3, \infty)^2 \oplus \mathbb{F}[a_1, a_4) \oplus \mathbb{F}[a_2, a_5)$$



e_5 te_5 t^2e_5 $t^3e_5=0$
e_4 te_4 t^2e_4 $t^3e_4=0$
e_3 te_3 $t^2e_3=t^3e_3=t^4e_3=\dots$
e_2 te_2 $t^2e_2=t^3e_2=t^4e_2=\dots$
e_1 te_1 t^2e_1 t^3e_1 t^4e_1 $t^5e_1=t^6e_1=t^7e_1=\dots$
a_0 a_1 a_2 a_3 a_4 a_5

Proof of the Structure Theorem:

Let R be any graded PID, and let M be a graded R-module. Suppose that M is finitely generated. Pick a minimal generating set $\{x_1, \dots, x_n\}$ in M. There is an R-module epimorphism $\varphi: R^n \longrightarrow M$, $\varphi(e_i) = x_i$, where $e_i = (0, \dots, \overset{i}{1}, \dots, 0)$. Since R is a PID, $\operatorname{Ker}\varphi$ is a free R-module of rank $m \leq n$. By the Smith normal form algorithm, there is a basis $v_1, \dots, v_n$ of R^n and there are non-invertible elements $d_1, \dots, d_m$ in R such that $d_i \mid d_j$ if $i \leq j$ and $d_1 v_1, \dots, d_m v_m$ is a basis of $\operatorname{Ker}\varphi$.

It follows that
$$M \cong R^n / \operatorname{Ker}\varphi \cong R^{n-m} \oplus R/(d_1) \oplus \cdots \oplus R/(d_m)$$
where $(d_1) \supseteq (d_2) \supseteq \cdots \supseteq (d_m)$.

Moreover, $v_1, \dots, v_n$ can be chosen so that each v_i is homogeneous (i.e., all its summands have the same degree), and each d_j is also homogeneous. Thus, in the case $R = \mathbb{F}[t]$, $d_j = t^{r_j}$ with $r_j \geq 1$.

Hence $\mathbb{F}[t] v_i \cong \Sigma^{q_i} \mathbb{F}[t]$ if v_i has degree q_i and
$$\mathbb{F}[t] v_j / (d_j v_j) \cong \Sigma^{q_j} \mathbb{F}[t] / (t^{r_j}).$$ ✓