

Воронежский
Государственный
Университет

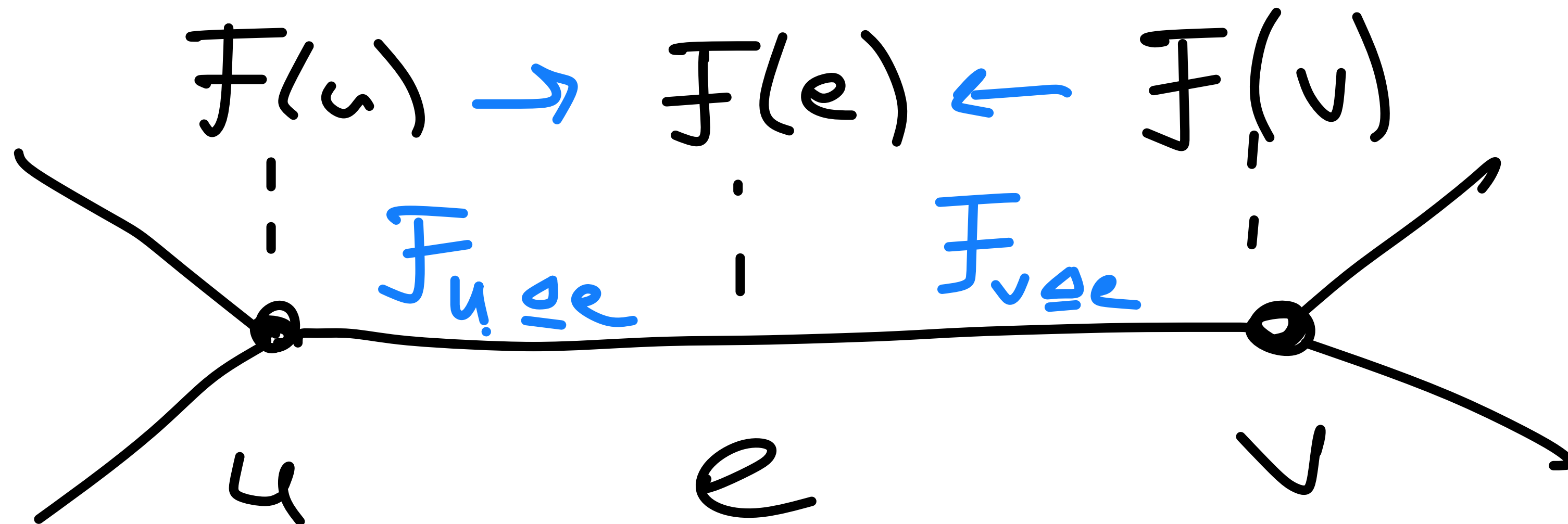
Когомологии пучков и графовые нейросети

«Топология, комбинаторика и анализ данных» – Воронеж,
15-17 ноября 2023

Максим Бекетов, стажер-исследователь и аспирант
Международная Лаборатория Алгебраической Топологии и её Приложений (МЛ АТиП)
Факультета Компьютерных Наук (ФКН) ВШЭ

Пучки

- **Опр 1:** *Граф* G – мн-во вершин V и ребер E с отношением *инцидентности* $v \trianglelefteq e$. (Мы разрешаем мульти-ребра)
- **Опр. 2:** *Пучок* (над графом) \mathcal{F} – это набор векторных пространств $\mathcal{F}(u)$ «над» каждой вершиной u и $\mathcal{F}(e)$ «над» каждым ребром e графа, с линейным преобразованием $\mathcal{F}_{v \trianglelefteq e} : \mathcal{F}(v) \rightarrow \mathcal{F}(e)$



Пучки

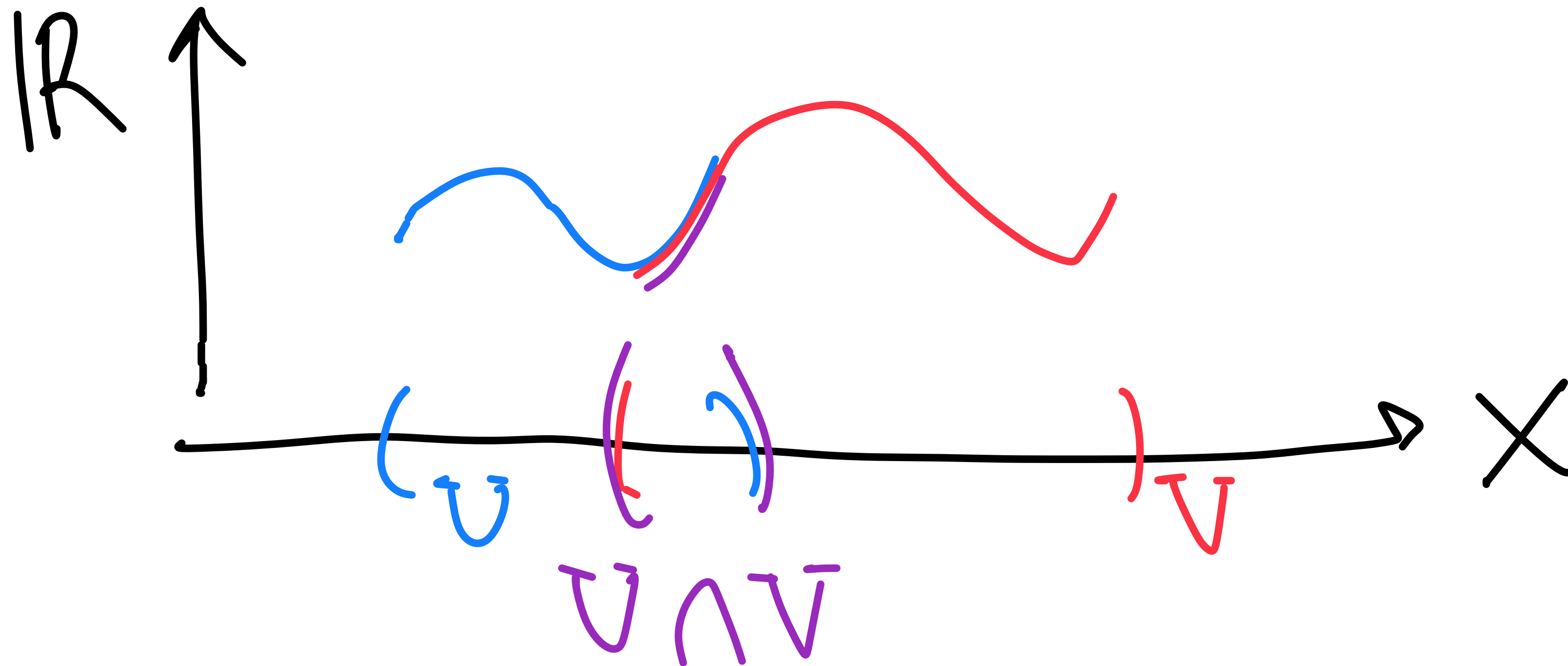
- **Опр 1:** *Граф* G – мн-во вершин V и ребер E с отношением *инцидентности* $v \trianglelefteq e$. (Мы разрешаем мульти-ребра)
- **Опр. 2:** *Пучок* (над графом) \mathcal{F} – это набор векторных пространств $\mathcal{F}(u)$ «над» каждой вершиной u и $\mathcal{F}(e)$ «над» каждым ребром e графа, с линейным преобразованием $\mathcal{F}_{v \trianglelefteq e} : \mathcal{F}(v) \rightarrow \mathcal{F}(e)$
- \mathcal{F} от фр. Faisceau, англ. Sheaf = пучок, связка, сноп (сена)
- Вект. пр-ва $\mathcal{F}(e)$ и $\mathcal{F}(e)$ наз. *слоями* (англ. stalk = стебель, фр. fibre = волокно). У нас в них всегда будет *скалярное произведение*
- $\mathcal{F}_{v \trianglelefteq e}$ – *отображение ограничения*

Сечения

- **Опр 3:** На W – подмн.-во V вершин G с пучком \mathcal{F} , сечение \mathcal{F} над W – это выбор $\forall v \in W$ определенного вектора $x_v \in \mathcal{F}(v)$ такого, что $\forall v, v' \triangleleft e$ имеем $\mathcal{F}_{v \triangleleft e}(x_v) = \mathcal{F}_{v' \triangleleft e}(x_{v'})$
- Если наши векторные пространства – это пр-ва гладких функций, то это условие означает, что они согласуются («сшиваются») на ребрах

Сечения

- Вообще учки (сколько-то гладких функций) можно рассматривать над топологическими пр-вами:



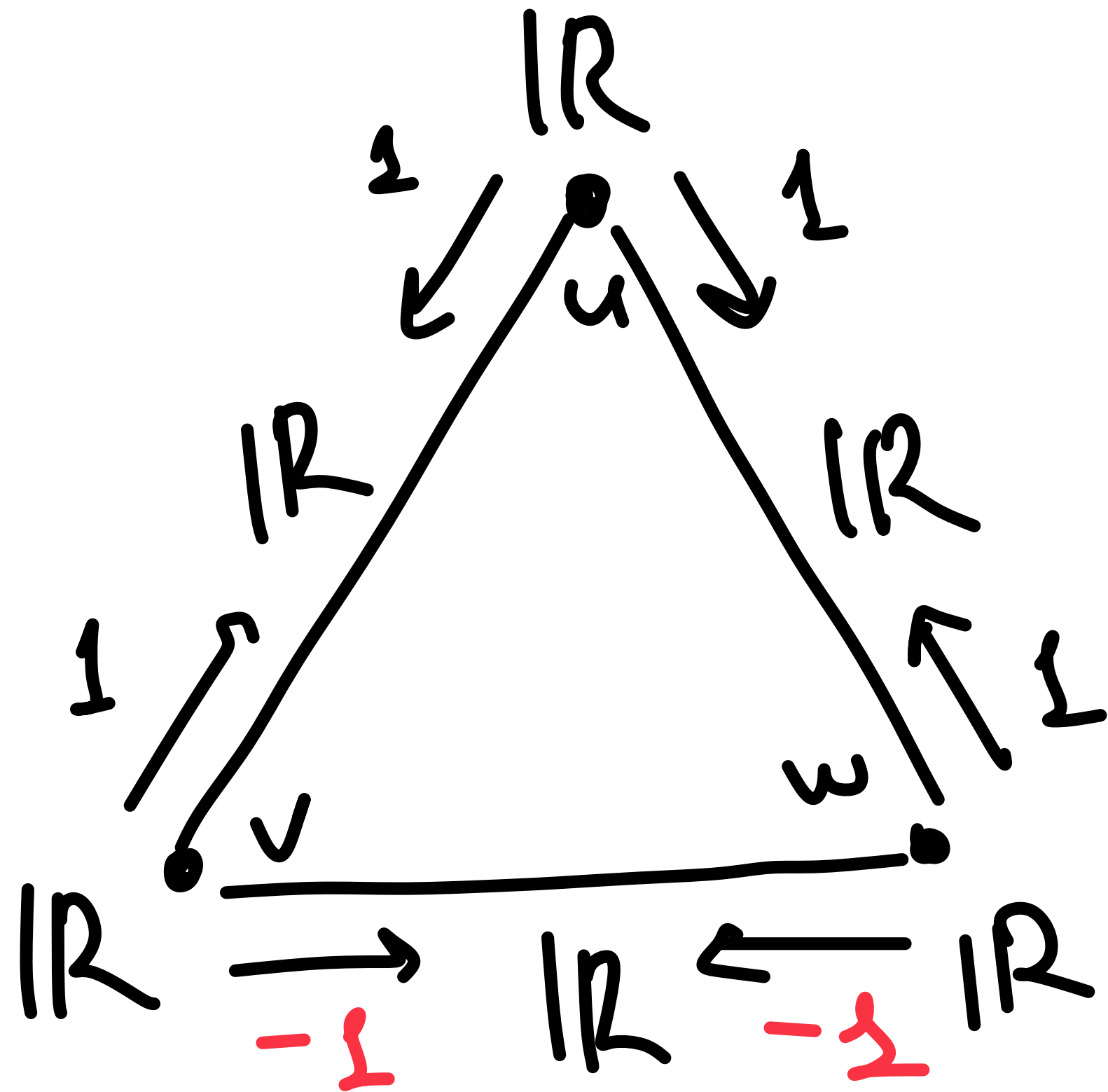
- Но у нас все попроще, у нас над графами, потому нет разницы между *предпучком* и *пучком*

Сечения

- Заметьте, что сечения (над любым $W \subset V$) можно *складывать* (т.к. отображения ограничения линейны)
- а также *скалярно умножать* (в каждом слое есть скалярное произведение)
- Значит, сечения образуют *векторное пространство* – обозначим его $H^0(W, \mathcal{F})$

Сечения

- Пример:** Пусть \mathcal{V} – вект. пр-во, тогда *постоянный пучок* $\underline{\mathcal{V}}$ над графом G – это такой, что $\underline{\mathcal{V}}(v) = \mathcal{V}$ для всех v (и так же для всех ребер, e), и отображение ограничения $\mathcal{F}_{v \leq e}$ – тождественным отображ.-ем



Сечения постоянного пучка – постоянные функции

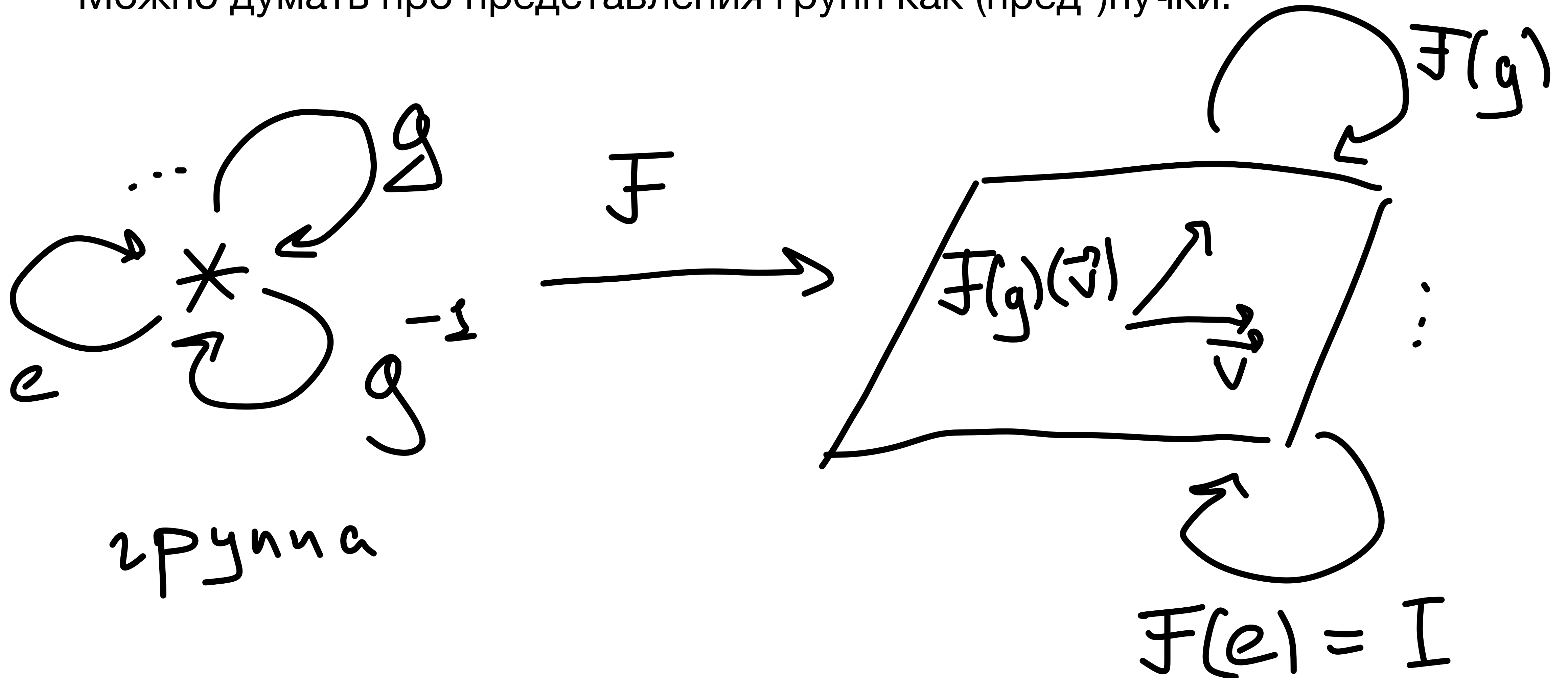
<- в случае $\mathcal{V} = \mathbb{R}$ (важный частный случай)

Есть по крайней мере одно *глобальное сечение* – всюду 0

Но есть ли другие? Наличие *нетривиальных (глобальных) сечений* – важное свойство

Интерлюдия: представления, колчаны

- Можно думать про представления групп как (пред-)пучки:



Коцепи

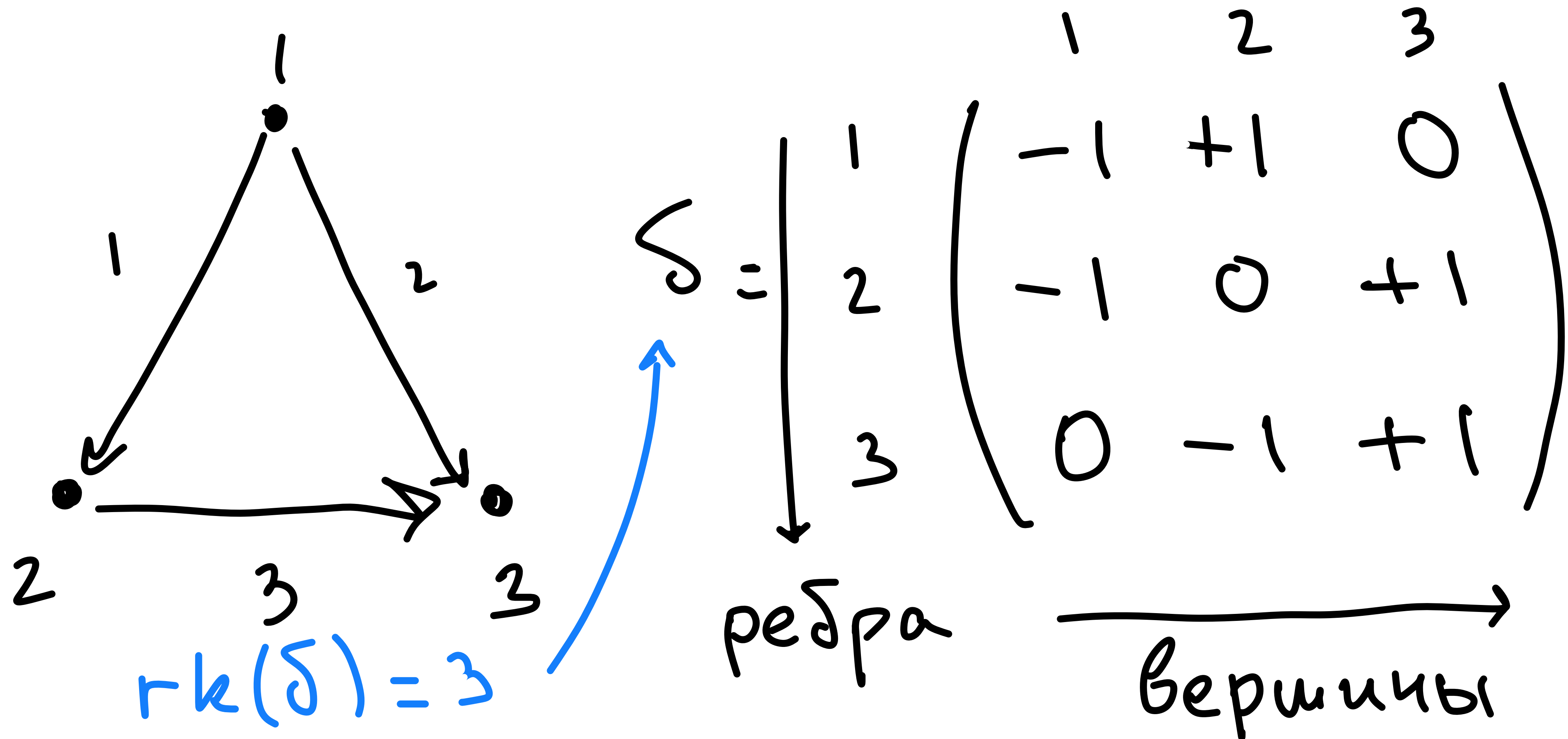
- **Опр. 4:** Пусть \mathcal{F} – пучок над G . 0-коцепи – это функции на вершинах (выбор $x_v \in \mathcal{F}(v)$ для всех v), 1-коцепи – функции на ребрах:

$$C^0(G; \mathcal{F}) = \bigoplus_{v \in V} \mathcal{F}(v), \quad C^1(G; \mathcal{F}) = \bigoplus_{e \in E} \mathcal{F}(e)$$

- Сечения образуют подпространство C^0 – ядро некоторого линейного отображения $\delta : C^0 \rightarrow C^1$ – отображения кограницы
- Ориентируем ребра: $[v : e] = +1$, если $v \leftarrow e$ (v – «сток»), и $[v : e] = -1$, если $v \rightarrow e$ (v – «исток»)
- Тогда $\delta(x_v) = \sum_{v \triangleleft e} [v : e] \mathcal{F}_{v \triangleleft e}(x_v)$ – зависит от пучка, будем обозн. $\delta^{\mathcal{F}}$

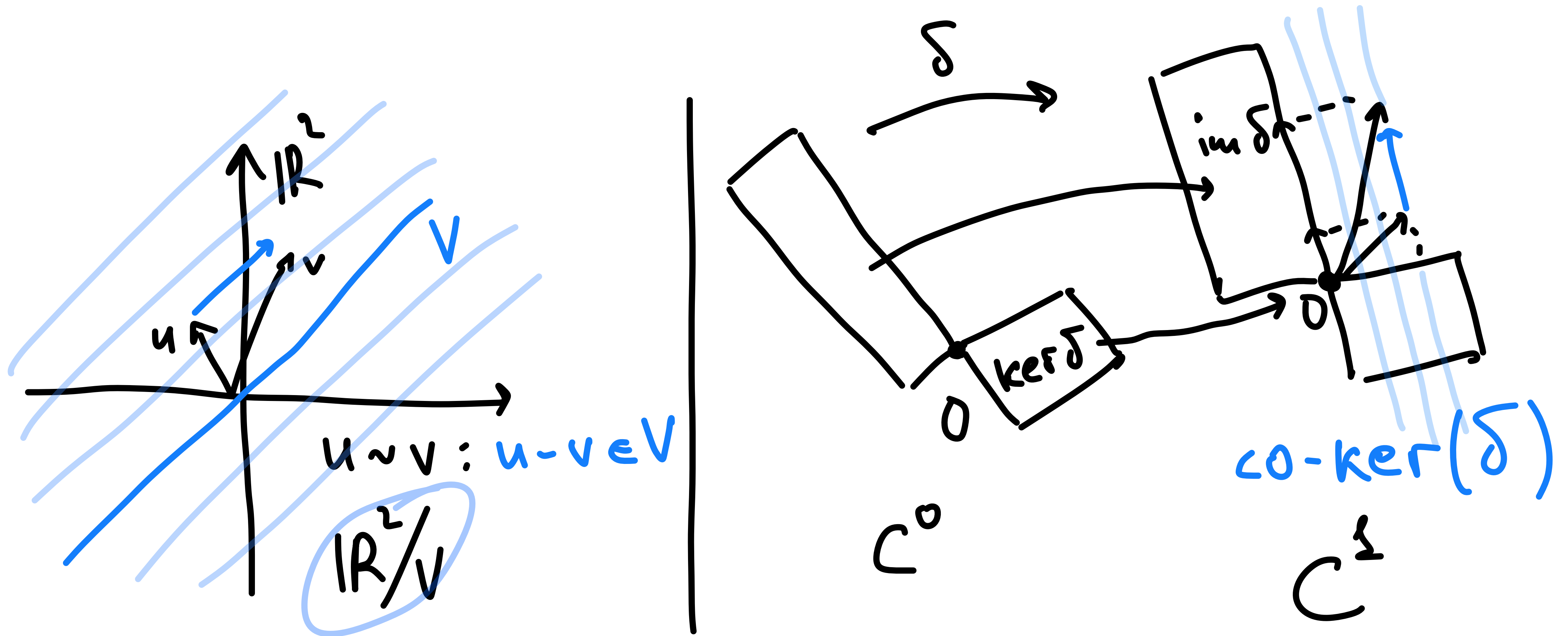
Коцепи

- Пример: Если $\mathcal{F} = \underline{\mathbb{R}}$, то δ – просто матрица инцидентности (со знаком)



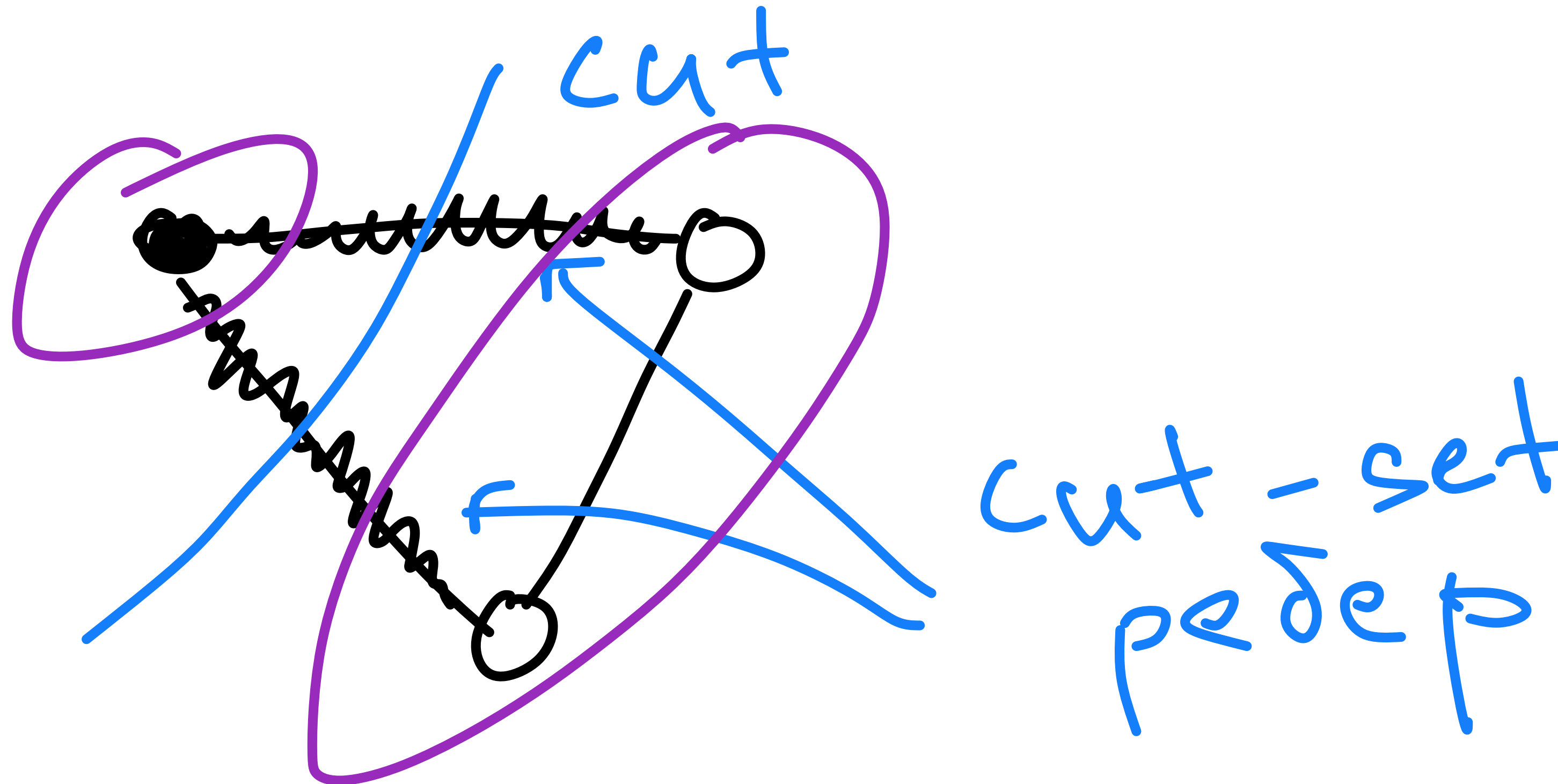
Когомологии

- 0-когом. $H^0(G; \mathcal{F}) = \ker \delta$, 1-когом. $H^1(G; \mathcal{F}) = C^1(G; \mathcal{F}) / (\text{im } \delta)$



Когомологии

- 1-КОГОМ. $H^1(G; \mathcal{F}) = C^1(G; \mathcal{F}) / (\text{im } \delta)$
- $\text{im } \delta \simeq \text{cut-space}(G) = \text{Span}(\{I(\text{cut-set})\}) = \text{im } \delta^*$



Лапласиан

- **Опр. 5:** Лапласиан (оператор Лапласа) пучка $\Delta_{\mathcal{F}} = \delta^* \delta$
- По построению полож. (полу-) определен: $\langle x, \delta^* \delta x \rangle = \langle \delta x, \delta x \rangle \geq 0$
- И ядро его: $\ker \delta^* \delta = \ker \delta = H^0(G; \mathcal{F})$
- Для пучка \mathcal{F} лапласиан Δ – симметричная блочная матрица,

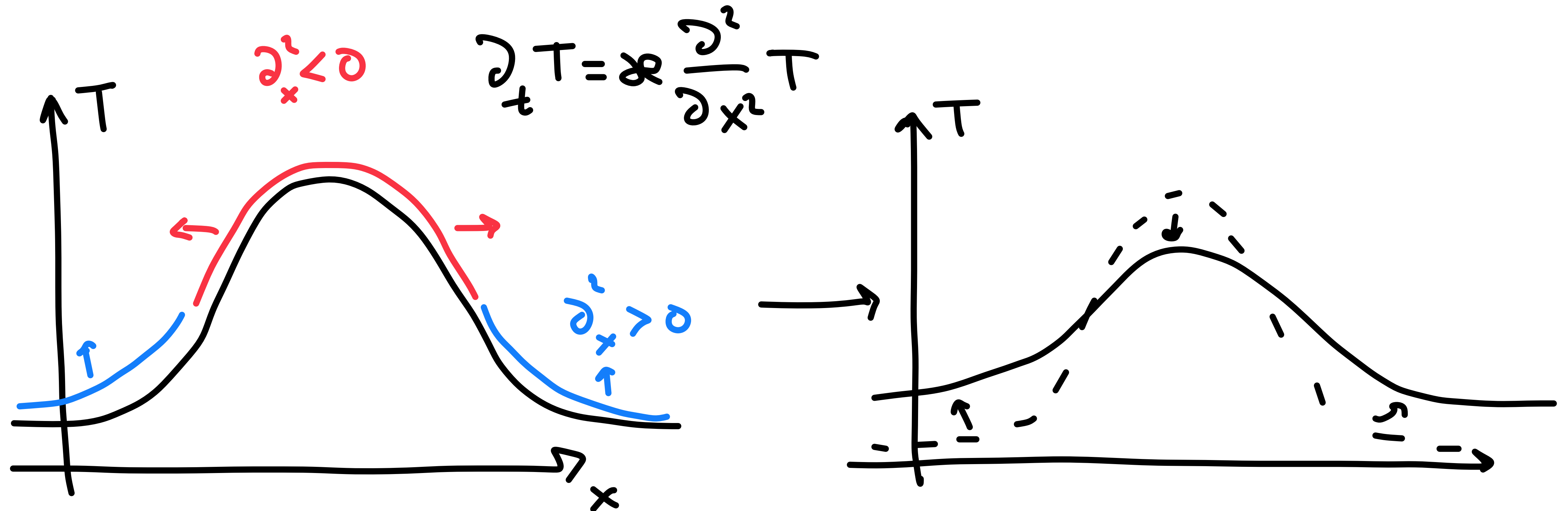
$$\Delta_{(v_1, v_2)} = - \mathcal{F}_{v_2 \triangleleft e}^* \mathcal{F}_{v_1 \triangleleft e},$$

а диагональные элементы:
$$\Delta_{(v_1, v_1)} = \sum_{e: v_1 \triangleleft e} \mathcal{F}_{v_1 \triangleleft e}^* \mathcal{F}_{v_1 \triangleleft e}$$

Диффузия

- Лапласиан задает на пучке диффузию (уравнение теплопроводности):

$\partial_t f = -\Delta f$. Как выглядит обычная диффузия:

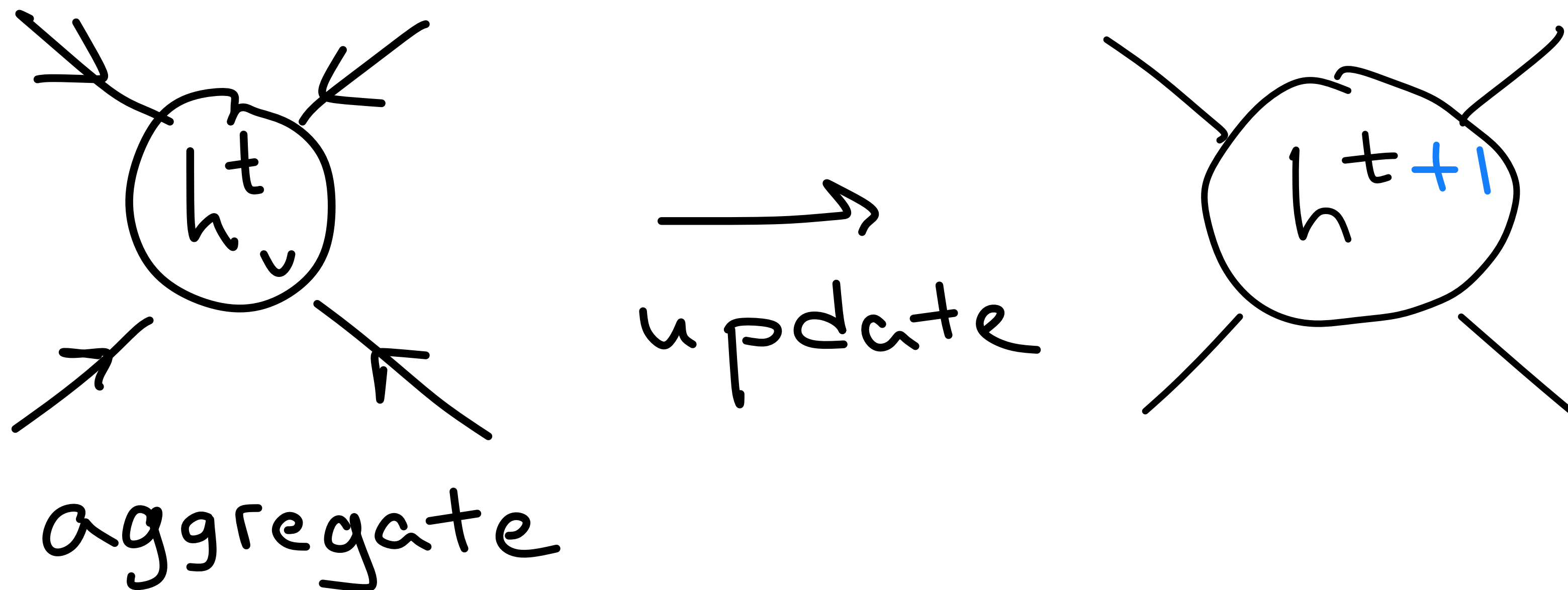


Графовые нейросети

- Почти все GNN представляют собой message passing:

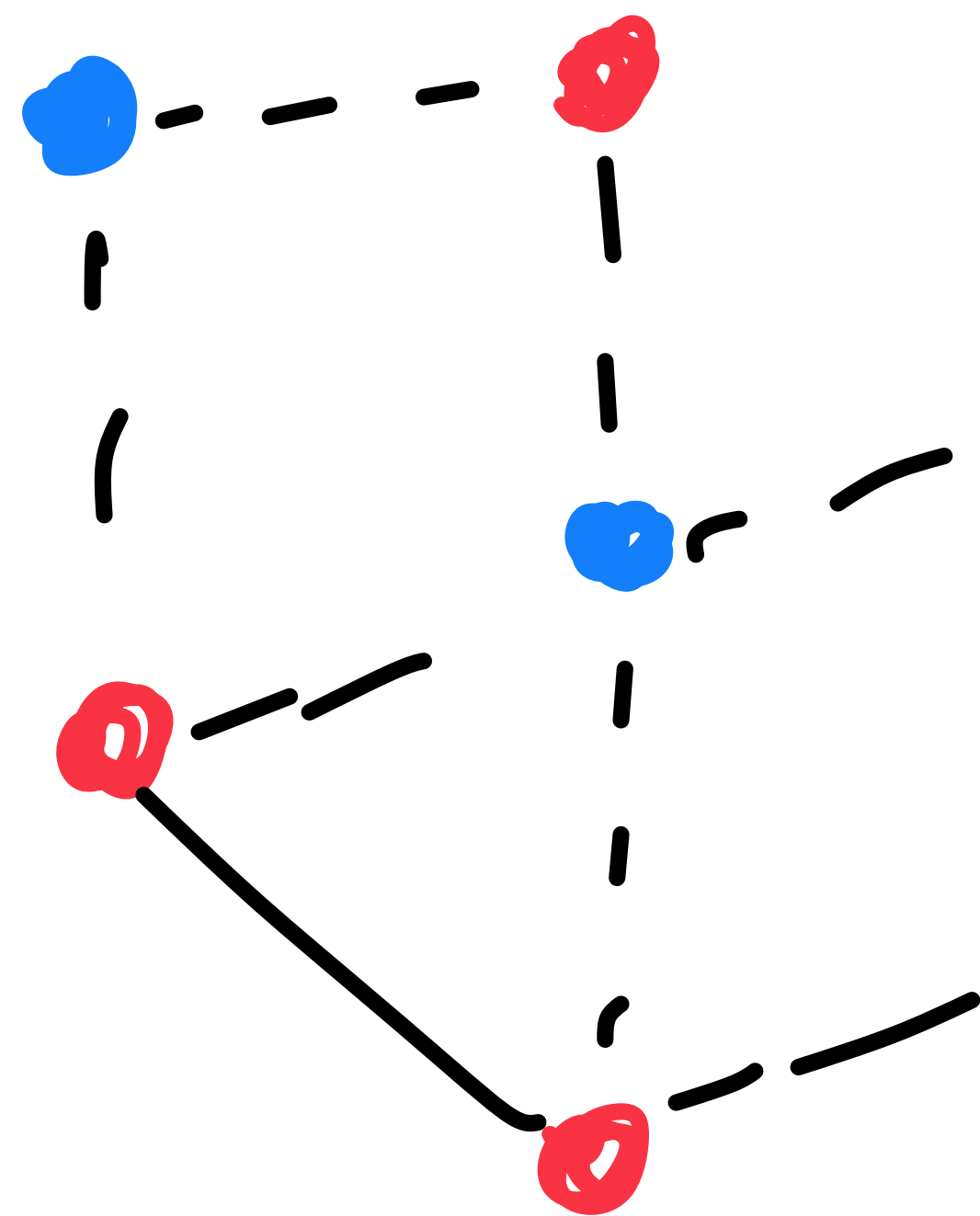
$$m_v^t = \text{Aggregate}(\{h_u^{t-1} \mid u \in \mathcal{N}(v)\}) \quad h_v^t = \text{Update}(h_v^{t-1}, m_v^t)$$

h for hidden (state/feature), m for message, U – какая-то нелинейная функция (нейросеть)

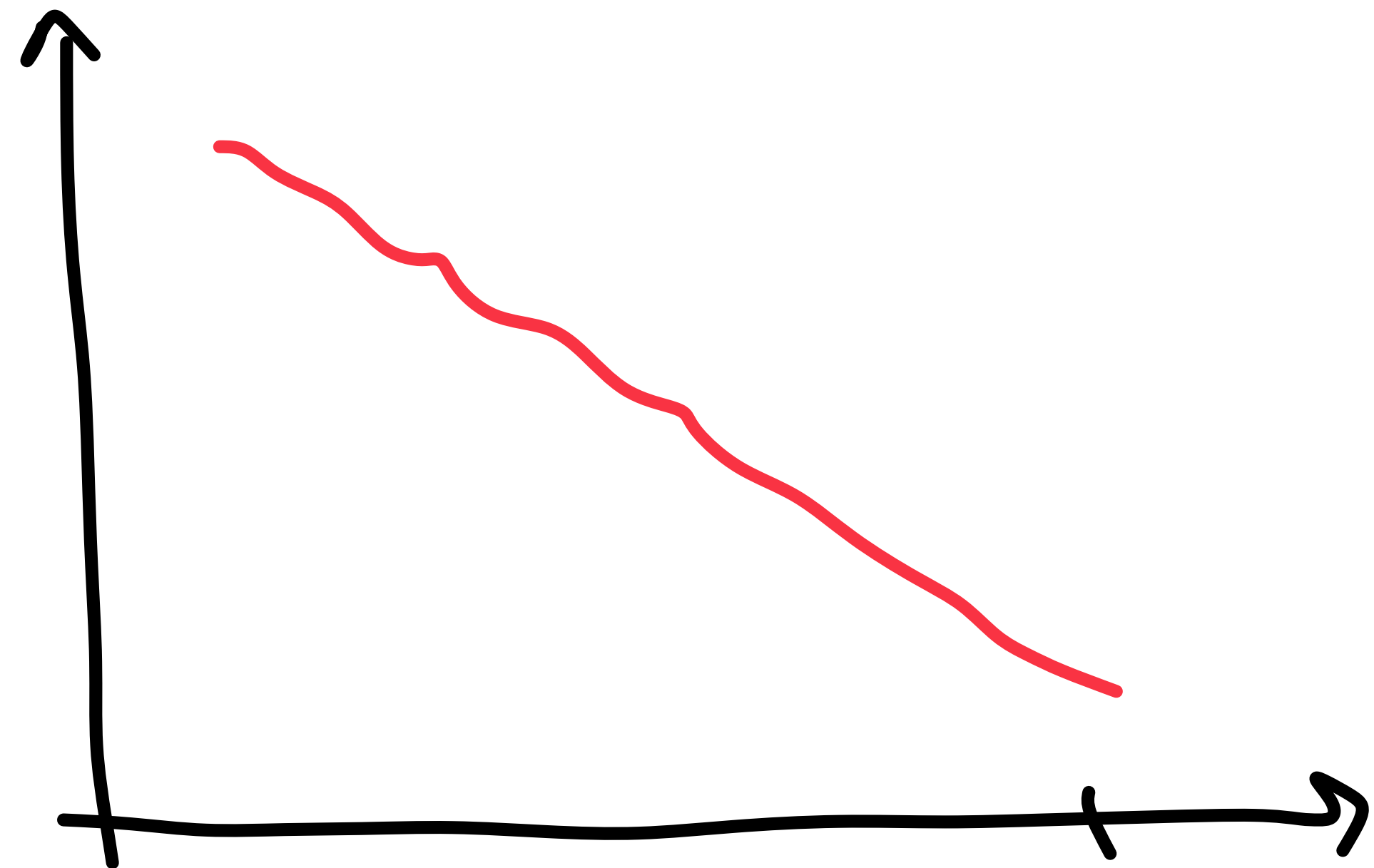


Графовые нейросети: Heterophily

- Обычные GNN плохо работают на разнородных (heterophilic) графах



Accuracy

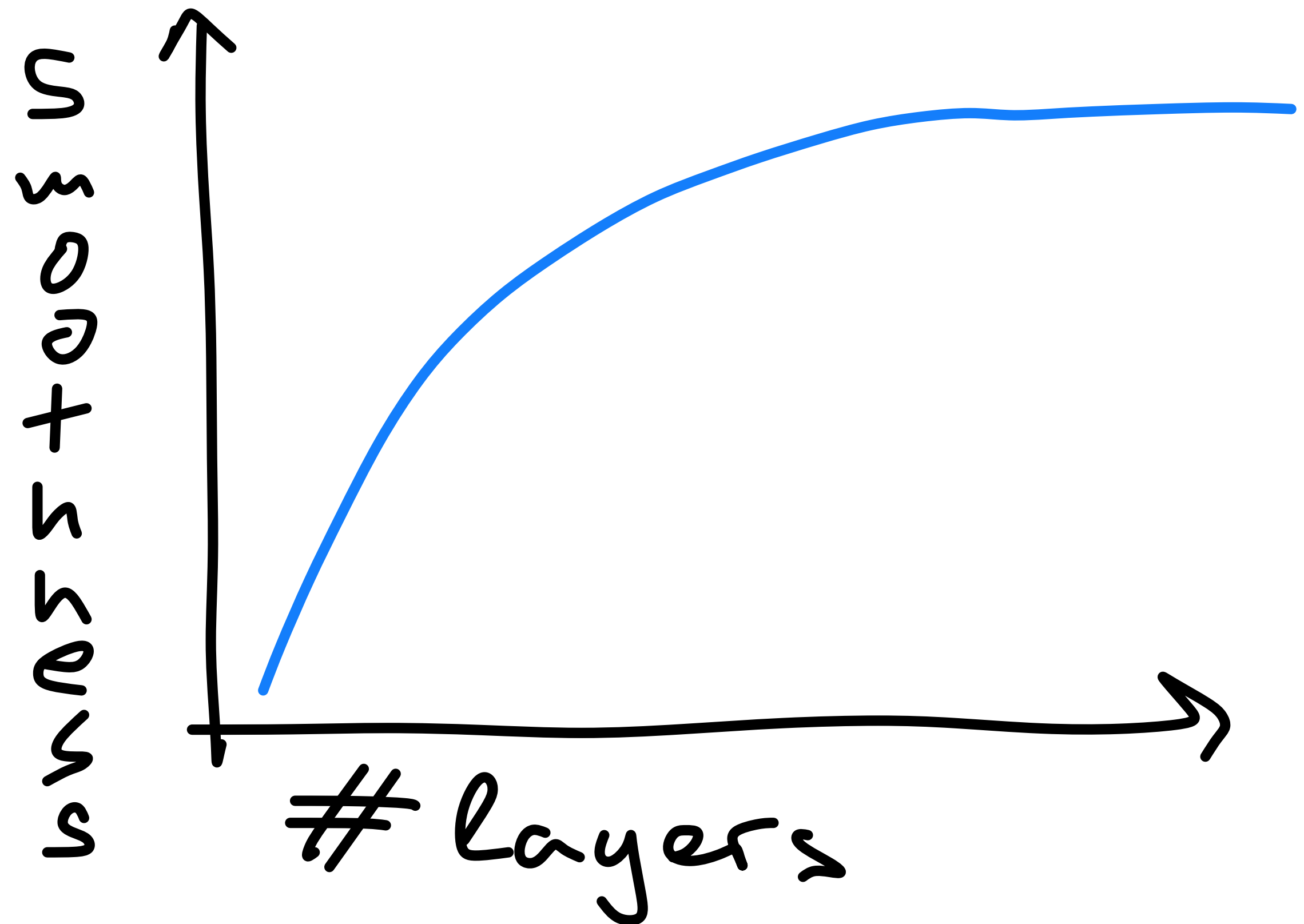
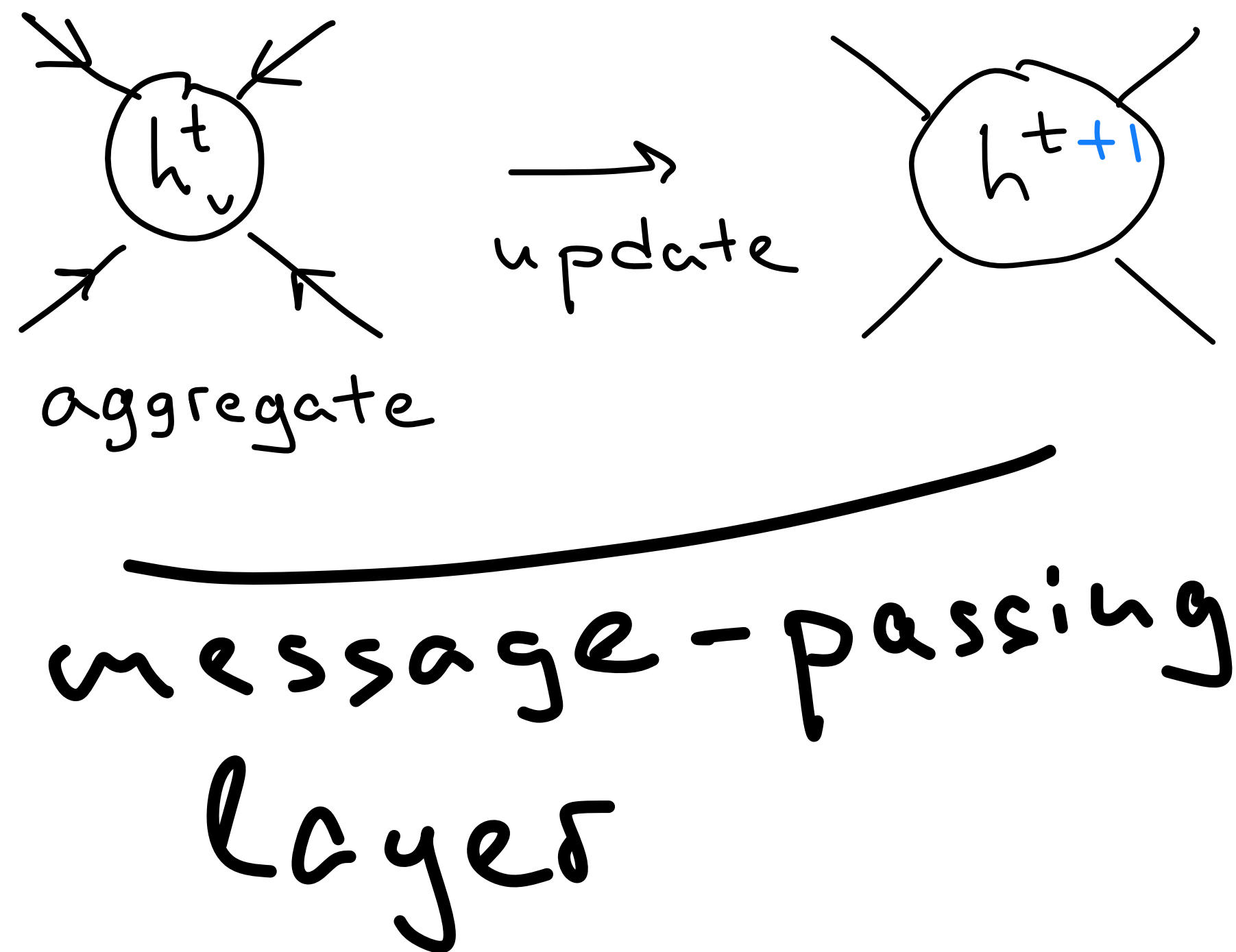


$$\text{Heterophily} = \frac{\# \text{ Blue-Red Edges}}{\text{Total \# Edges}}$$

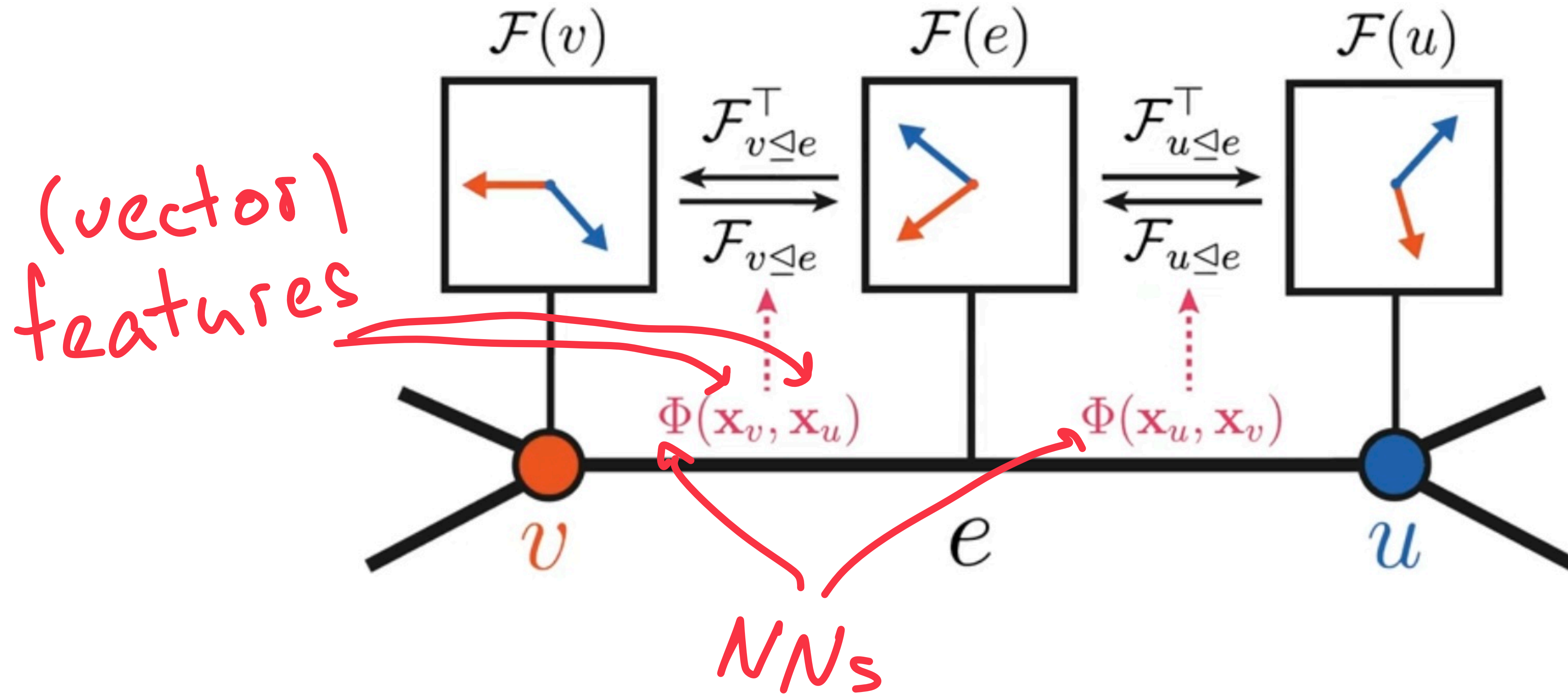
Heterophily

Графовые нейросети: Oversmoothing

- И, несмотря на нелинейную природу нейросетей, с глубиной они все равно «сглаживают» любой сигнал (тепло распределяется равномерно)



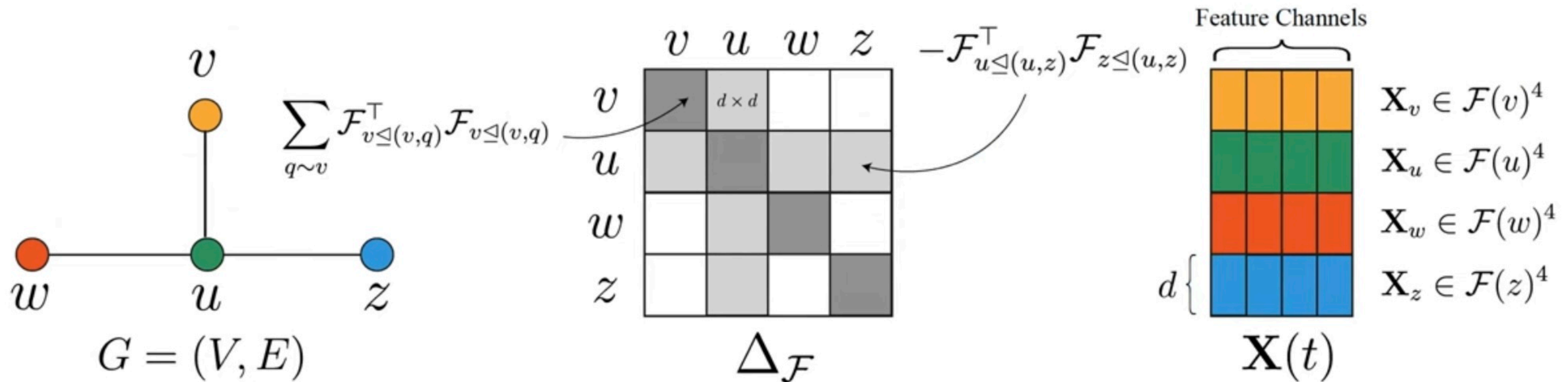
Решение*: Sheaf GNNs (SNNs)



Neural Sheaf Diffusion: A Topological Perspective on Heterophily and Oversmoothing in GNNs, C. Bodnar, F. Di Giovanni, B. Chamberlain, P. Liò, M. Bronstein *NeurIPS 2022*

Neural Sheaf Diffusion

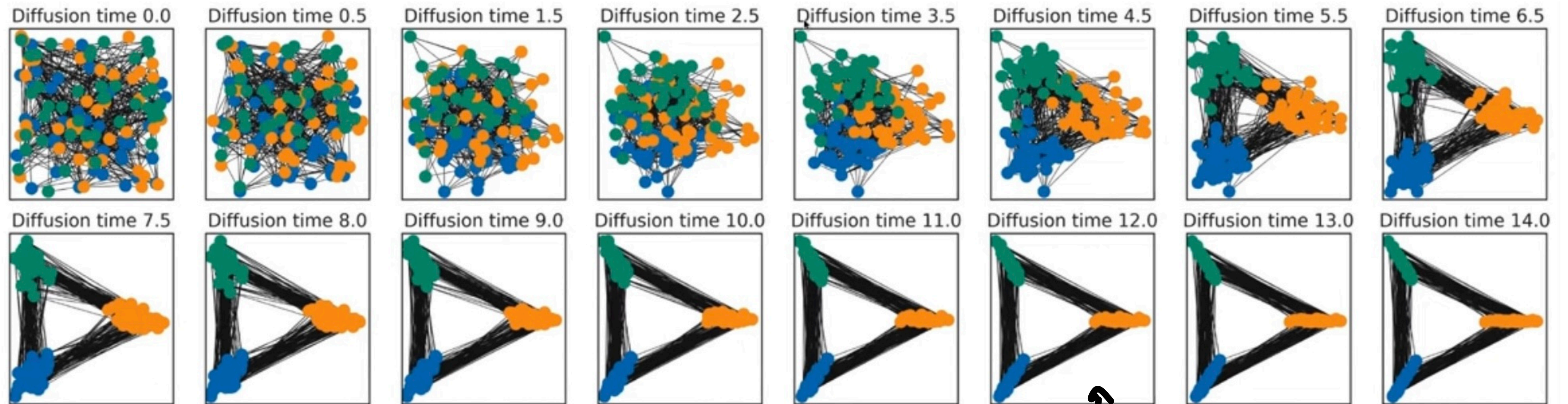
- Пучковая (нейронная) диффузия: $\mathbf{X}(0) = \mathbf{X}, \quad \partial_t \mathbf{X}(t) = -\Delta_{\mathcal{F}} \mathbf{X}(t)$



Neural Sheaf Diffusion: A Topological Perspective on Heterophily and Oversmoothing in GNNs, C. Bodnar, F. Di Giovanni, B. Chamberlain, P. Liò, M. Bronstein *NeurIPS 2022*

Neural Sheaf Diffusion – успехи

- И ее успехи например в задаче community detection:



$$\mathcal{F} \approx \mathbb{R}^2 \uparrow$$

И такое получается при почти любых начальных условиях $\mathbf{X}(t = 0)$!

Neural Sheaf Diffusion: A Topological Perspective on Heterophily and Oversmoothing in GNNs, C. Bodnar, F. Di Giovanni, B. Chamberlain, P. Liò, M. Bronstein *NeurIPS 2022*

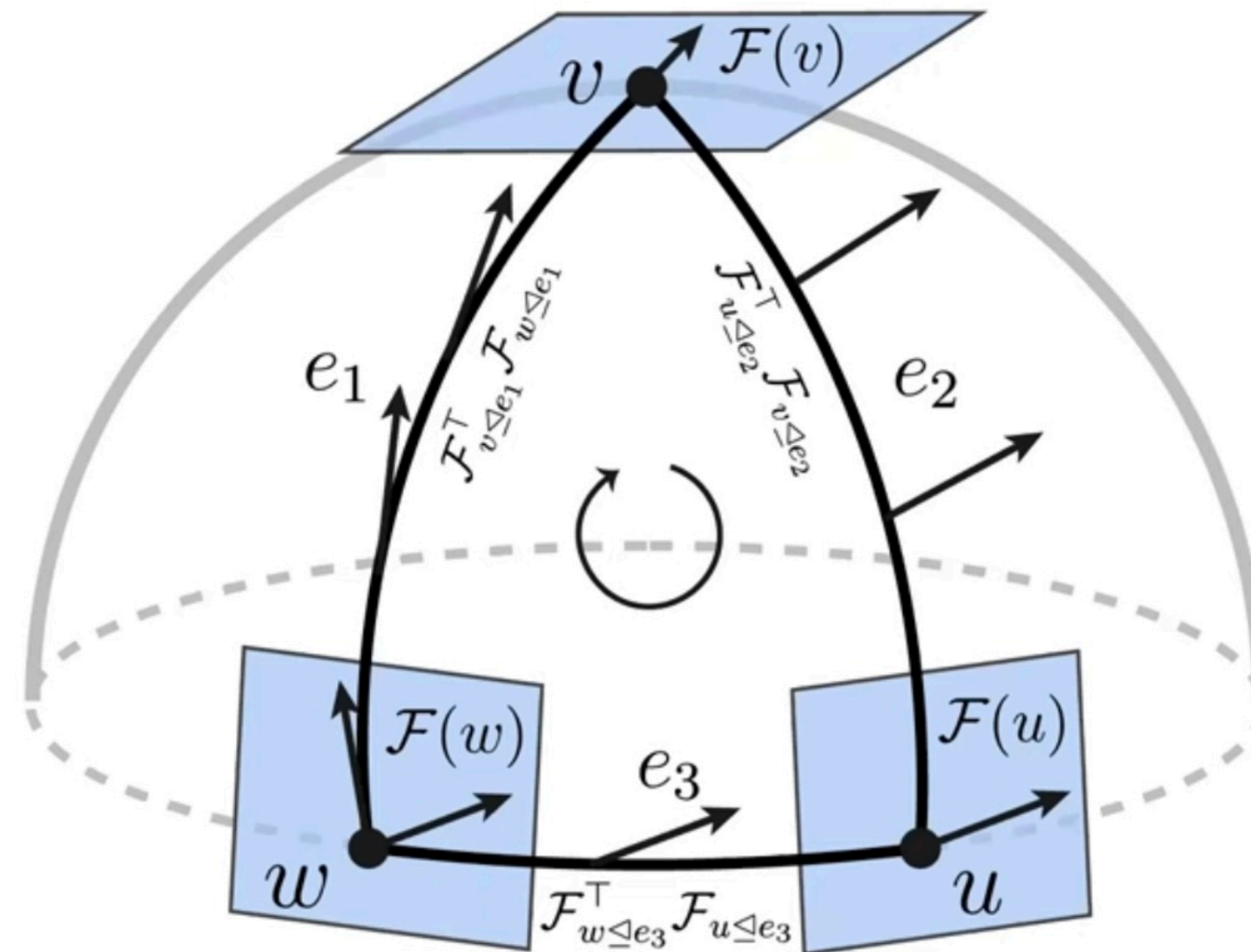
Neural Sheaf Diffusion – успехи

	Texas	Wisconsin	Film	Squirrel	Chameleon	Cornell	Citeseer	Pubmed	Cora
Hom level	0.11	0.21	0.22	0.22	0.23	0.30	0.74	0.80	0.81
#Nodes	183	251	7,600	5,201	2,277	183	3,327	18,717	2,708
#Edges	295	466	26,752	198,493	31,421	280	4,676	44,327	5,278
#Classes	5	5	5	5	5	5	7	3	6
Diag-NSD	85.67 ±6.95	88.63 ±2.75	37.79 ±1.01	54.78 ±1.81	68.68 ±1.73	86.49 ±7.35	77.14 ±1.85	89.42±0.43	87.14±1.06
O(d)-NSD	85.95 ±5.51	89.41 ±4.74	37.81 ±1.15	56.34 ±1.32	68.04 ±1.58	84.86 ±4.71	76.70±1.57	89.49 ±0.40	86.90±1.13
Gen-NSD	82.97±5.13	89.21 ±3.84	37.80 ±1.22	53.17±1.31	67.93±1.58	85.68 ±6.51	76.32±1.65	89.33±0.35	87.30±1.15
GGCN	84.86 ±4.55	86.86±3.29	37.54±1.56	55.17 ±1.58	71.14 ±1.84	85.68 ±6.63	77.14 ±1.45	89.15±0.37	87.95 ±1.05
H2GCN	84.86 ±7.23	87.65±4.98	35.70±1.00	36.48±1.86	60.11±2.15	82.70±5.28	77.11±1.57	89.49 ±0.38	87.87 ±1.20
GPRGNN	78.38±4.36	82.94±4.21	34.63±1.22	31.61±1.24	46.58±1.71	80.27±8.11	77.13±1.67	87.54±0.38	87.95 ±1.18
FAGCN	82.43±6.89	82.94±7.95	34.87±1.25	42.59±0.79	55.22±3.19	79.19±9.79	N/A	N/A	N/A
MixHop	77.84±7.73	75.88±4.90	32.22±2.34	43.80±1.48	60.50±2.53	73.51±6.34	76.26±1.33	85.31±0.61	87.61±0.85
GCNII	77.57±3.83	80.39±3.40	37.44±1.30	38.47±1.58	63.86±3.04	77.86±3.79	77.33 ±1.48	90.15 ±0.43	88.37 ±1.25
Geom-GCN	66.76±2.72	64.51±3.66	31.59±1.15	38.15±0.92	60.00±2.81	60.54±3.67	78.02 ±1.15	89.95 ±0.47	85.35±1.57
PairNorm	60.27±4.34	48.43±6.14	27.40±1.24	50.44±2.04	62.74±2.82	58.92±3.15	73.59±1.47	87.53±0.44	85.79±1.01
GraphSAGE	82.43±6.14	81.18±5.56	34.23±0.99	41.61±0.74	58.73±1.68	75.95±5.01	76.04±1.30	88.45±0.50	86.90±1.04
GCN	55.14±5.16	51.76±3.06	27.32±1.10	53.43±2.01	64.82±2.24	60.54±5.30	76.50±1.36	88.42±0.50	86.98±1.27
GAT	52.16±6.63	49.41±4.09	27.44±0.89	40.72±1.55	60.26±2.50	61.89±5.05	76.55±1.23	87.30±1.10	86.33±0.48
MLP	80.81±4.75	85.29±3.31	36.53±0.70	28.77±1.56	46.21±2.99	81.89±6.40	74.02±1.90	75.69±2.00	87.16±0.37

Neural Sheaf Diffusion: A Topological Perspective on Heterophily and Oversmoothing in GNNs, C. Bodnar, F. Di Giovanni, B. Chamberlain, P. Liò, M. Bronstein *NeurIPS 2022*

СВЯЗНОСТЬ

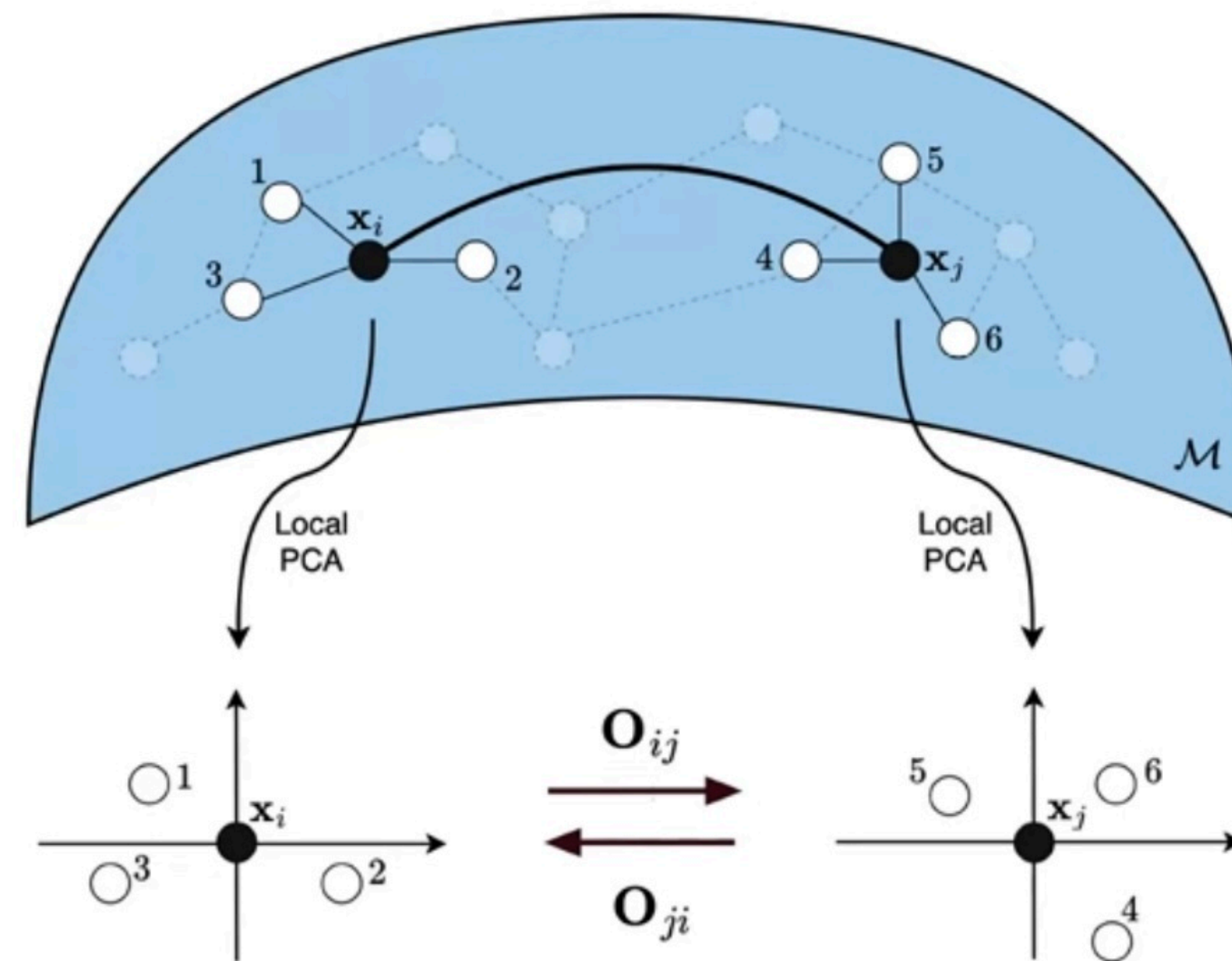
- Можно думать про это как про параллельный перенос и диффузию векторного поля по многообразию



Neural Sheaf Diffusion: A Topological Perspective on Heterophily and Oversmoothing in GNNs, C. Bodnar, F. Di Giovanni, B. Chamberlain, P. Liò, M. Bronstein *NeurIPS 2022*

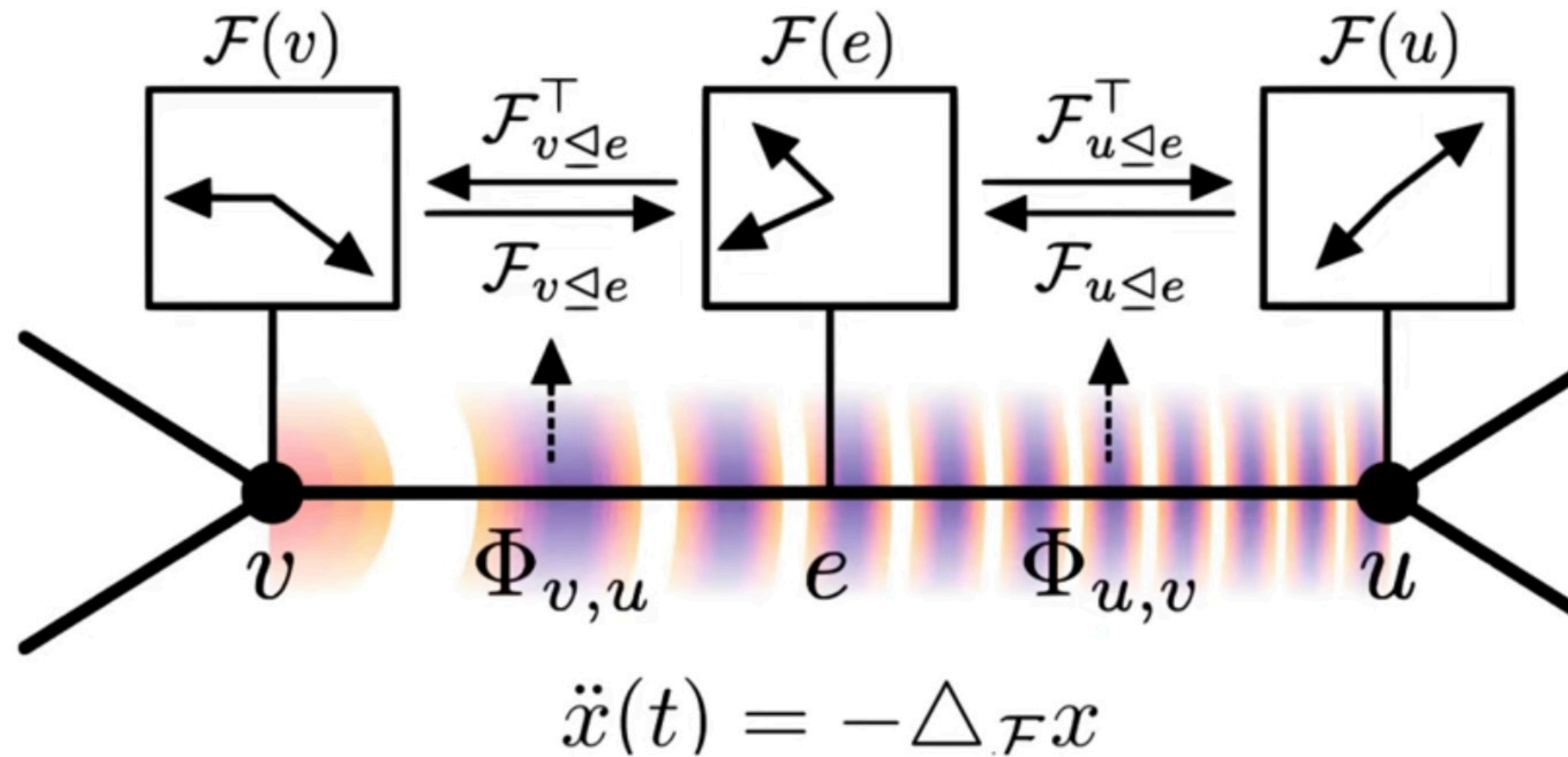
СВЯЗНОСТЬ

- И люди даже пытаются упростить нейронную диффузию, восстанавливая локально это многообр-е и связность на нем:



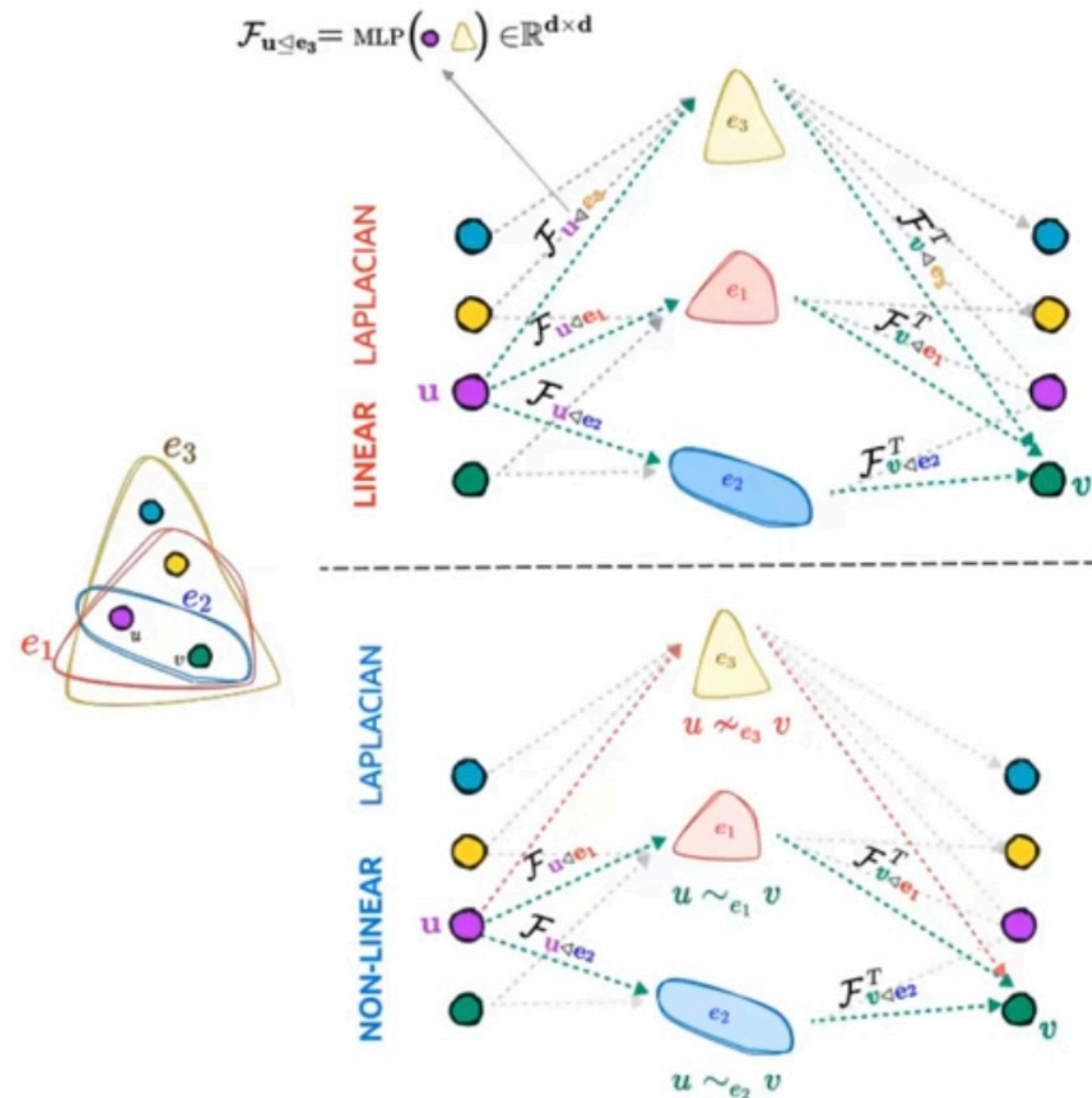
Barbero, Bodnar, Borde, et al., "Sheaf Neural Networks with Connection Laplacians", 2022

And beyond: волновое уравнение



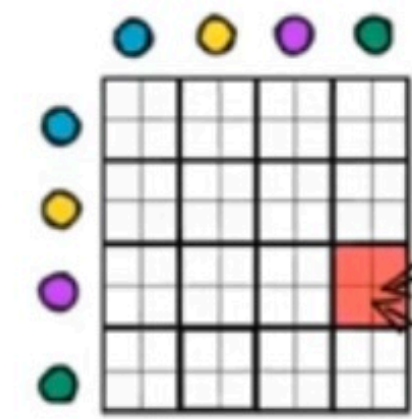
Suk et al., "Surfing on the Neural Sheaf", 2022.

And beyond: гиперграфы



IMPLICIT BIAS:

$$E_{L_2}^{\mathcal{F}}(x) = \frac{1}{2} \sum_e \frac{1}{\delta_e} \sum_{u,v \in e} \|\mathcal{F}_{v \trianglelefteq e} D_v^{-\frac{1}{2}} x_v - \mathcal{F}_{u \trianglelefteq e} D_u^{-\frac{1}{2}} x_u\|_2^2$$



$$(\mathcal{L}_{\mathcal{F}})_{uv} = - \sum_{\substack{e \in \mathbf{E} \\ u,v \in e}} \frac{1}{\delta_e} \mathcal{F}_{u \trianglelefteq e}^T \mathcal{F}_{v \trianglelefteq e} \in \mathbb{R}^{d \times d}$$

$$(\bar{\mathcal{L}}_{\mathcal{F}})_{uv} = - \sum_{\substack{e \in \mathbf{E} \\ u \sim_e v}} \frac{1}{\delta_e} \mathcal{F}_{u \trianglelefteq e}^T \mathcal{F}_{v \trianglelefteq e} \in \mathbb{R}^{d \times d}$$

$u \sim_e v$ if $(u, v) = \operatorname{argmax}_{u', v' \in e} \|\mathcal{F}_{u' \trianglelefteq e} x_{u'} - \mathcal{F}_{v' \trianglelefteq e} x_{v'}\|$

IMPLICIT BIAS:

$$E_{TV}^{\mathcal{F}}(x) = \frac{1}{2} \sum_e \frac{1}{\delta_e} \max_{u,v \in e} \|\mathcal{F}_{v \trianglelefteq e} D_v^{-\frac{1}{2}} x_v - \mathcal{F}_{u \trianglelefteq e} D_u^{-\frac{1}{2}} x_u\|_2^2$$

Duta et al., "Sheaf Hypergraph Networks", 2023.