

# ***p*-Адические числа и сложные системы**

21 января 2016, МИАН

Иерархические методы для сложных систем

Иерархия — деревья, билдинги, всплески (вейвлеты),  
ультраметрический анализ, *p*-адические числа.

Многомерные иерархические методы в анализе данных:

Кластеризация — деревья,

Мультикластеризация — сети, билдинги Брюа-Титса и аналоги.

Генетический код — 2-адическая плоскость.

Другие приложения:

Спиновые стёкла, динамика белка, упаковка ДНК.

Глубокое обучение — иерархические модели в машинном обучении.

## *p*-Адические числа

Поле *p*-адических чисел  $\mathbb{Q}_p$  — пополнение поля рациональных чисел по норме

$$|x|_p = p^{-j}, \quad x = p^j \frac{a}{b}, \quad |0|_p = 0.$$

*p*-Адические числа — ряды вида

$$x = \sum_{l=a}^{\infty} x_l p^l, \quad x_l \in \{0, 1, \dots, p-1\}.$$

З.И.Боревич, И.Р.Шафаревич, Теория чисел, 3-е изд., М.: Наука, 1985.

В.С.Владимиров, И. В.Волович, Е.И.Зеленов, *p*-Адический анализ и математическая физика, Наука, М., 1994.

**Ультраметрическое пространство** – метрическое пространство  $X$  с сильным неравенством треугольника: для ультраметрики  $d(\cdot, \cdot)$

$$d(x, y) \leq \max(d(x, z), d(z, y)), \quad \forall x, y, z \in X.$$

Дерево шаров в (локально компактном) ультраметрическом пространстве — шары (ненулевого диаметра либо изолированные точки) являются вершинами, рёбра — пары шаров вида (шар, максимальный подшар).

Можно обсуждать приложения ультраметрических пространств для описания иерархии в различных задачах.

С. В. Козырев, Ультраметрика в теории сложных систем, ТМФ, 185:2 (2015), 346–360

Шары в  $\mathbb{Q}_p$  имеют вид

$$p^j (n + \mathbb{Z}_p), \quad n = \sum_{l=a}^{-1} n_l p^l, \quad n_l \in \{0, 1, \dots, p-1\},$$

$j$  целое,

$n$  целое отрицательное,

$\mathbb{Z}_p$  есть кольцо целых  $p$ -адических чисел (шар  $|x|_p \leq 1$ ),

$n$  — параметр на  $\mathbb{Q}_p/\mathbb{Z}_p$  — нумерует единичные шары.

$p^j n$  — нумерует шары ненулевого диаметра.

Дерево шаров в  $\mathbb{Q}_p$ : в каждом шаре  $p$  максимальных подшаров в соответствии с разложением

$$\mathbb{Z}_p = \bigcup_{m=0}^{p-1} (m + p\mathbb{Z}_p).$$

Кластеризация – метод иерархической классификации данных.

Данные размечаются иерархической системой (деревом) кластеров. В типичном случае ( $k$ -средние, кластеризация по ближайшему соседу) — кластеры строятся по метрике на данных.

МультиклUSTERизация — несколько систем кластеров на тех же данных. Несколько метрик, различные метрики порождают различные деревья кластеров.

Сети кластеров с циклами — циклы возникают, когда кластеры относительно разных метрик совпадают как множества.

A. Strehl, J. Ghosh, C. Cardie, *Cluster ensembles — a knowledge reuse framework for combining multiple partitions*. Journal of Machine Learning Research, 2002. 3. P.583–617.

$p$ -Адический случай: сети кластеров связаны с аффинными билдингами Брюа–Титса.

Кластеризация по ближайшему соседу:

$(M, \rho)$  – метрическое пространство.

$a = x_0, x_1, \dots, x_{n-1}, x_n = b$  –  $\varepsilon$ -цепь в пространстве  $(M, \rho)$ .

Цепное расстояние между  $a$  и  $b$ :

$d(a, b) = \inf(\varepsilon: \text{существует } \varepsilon\text{-цепь между } a, b).$

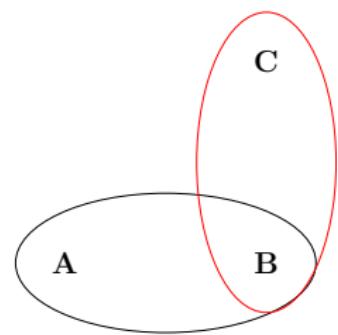
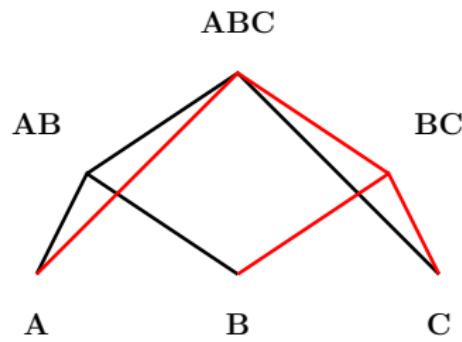
все свойства ультраметрики кроме невырожденности  
(несовпадающие точки могут иметь нулевое цепное  
расстояние).

Цепное расстояние определяет ультраметрику на множестве  
классов эквивалентности точек в  $M$ , где  $a, b$  лежат в одном  
классе эквивалентности, если  $d(a, b) = 0$ .

Кластер – шар относительно цепного расстояния.

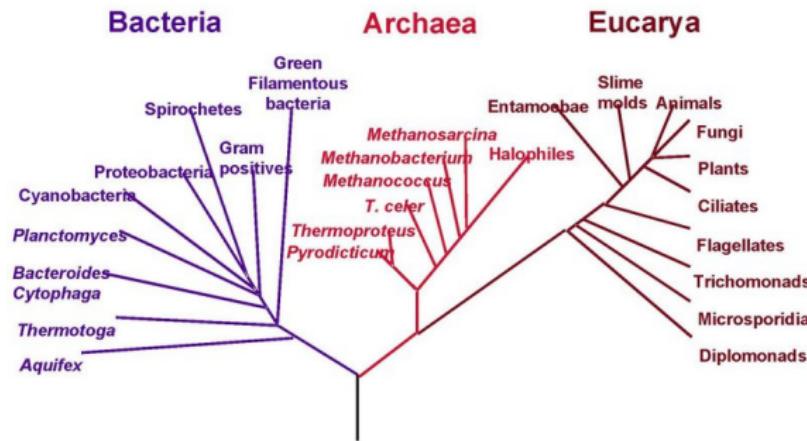
Кластеризация – покрытие пространства  $(M, \rho)$  кластерами.

Мультикластеризация – несколько метрик на  $M$ , несколько  
деревьев кластеров. Деревья склеиваются в сеть.



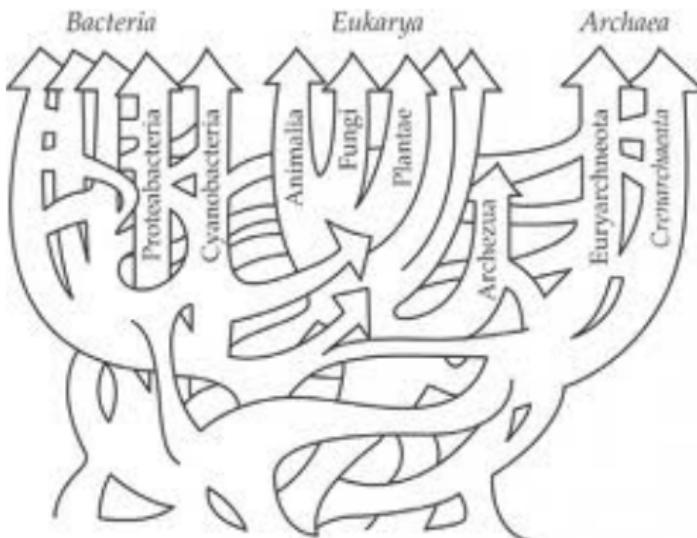
Кластеризация в науках о жизни: дерево жизни  
Карл Линней, *Systema Naturae*, 1735

## Phylogenetic Tree of Life



Рибосомное дерево жизни  
Карл Вёзе, 1977, 1985

# Филогенетическая сеть



Е.В. Кунин, Логика случая. О природе и происхождении биологической эволюции, Центрполиграф, 2014

Циклы — связь с мультикластеризацией.

Аффинный билдинг Брюа–Титса — симплексиальный комплекс, Вершины — классы эквивалентности решёток.

Решётка в  $\mathbb{Q}_p^d$  есть открытый компактный  $\mathbb{Z}_p$ -модуль в  $\mathbb{Q}_p^d$ :

$$\bigoplus_{i=1}^d \mathbb{Z}_p e_i,$$

$\{e_i\}$  есть базис в  $\mathbb{Q}_p^d$ .

Две решётки эквивалентны, если одна получается из другой умножением на число.

Две решётки  $L_1$  и  $L_2$  смежны (соединены ребром) если представители из классов эквивалентности  $L_1$  и  $L_2$  удовлетворяют

$$pL_1 \subset L_2 \subset L_1.$$

$k - 1$ –Симплексы — классы эквивалентности  $k$  смежных решёток, то есть последовательности

$$pL_k \subset L_1 \subset L_2 \subset \cdots \subset L_k, \quad 1 \leq k \leq d.$$

F.Bruhat, J.Tits, Groupes reductifs sur un corps local, I. Donnees radicielles valuees, Publ. Math. IHES 41: 5–251 (1972)

# Метрики в многомерных $p$ -адических пространствах.

Метрики  $s$  и нормы  $N$  в  $\mathbb{Q}_p^d$

$$s_{q_1, \dots, q_d}(x, y) = N_{q_1, \dots, q_d}(x - y),$$

$$N_{q_1, \dots, q_d}(z) = \max_{i=1, \dots, d}(q_i |z_i|_p), \quad q_i > 0,$$

$$N_{q_1, \dots, q_d}^A(z) = N_{q_1, \dots, q_d}(Az),$$

$$s_{q_1, \dots, q_d}^A(x, y) = N_{q_1, \dots, q_d}^A(x - y).$$

Для нормы  $N_{q_1, \dots, q_d}$  при

$$p^{-1} < q_1 < \dots < q_d \leq 1$$

набор шаров, лежащих между  $p\mathbb{Z}_p^d$  и  $\mathbb{Z}_p^d$ :

$$B_j = \mathbb{Z}_p \times \dots \times \mathbb{Z}_p \times p\mathbb{Z}_p \times \dots \times p\mathbb{Z}_p$$

с  $j$  компонентами  $\mathbb{Z}_p$  и  $d - j$  компонентами  $p\mathbb{Z}_p$ ,  $j = 0, \dots, d$ .

Симплексиальный комплекс шаров  $\mathcal{C}$  по отношению к семейству метрик  $s_{q_1, \dots, q_d}^A$ ,  $A \in \mathrm{GL}_d(\mathbb{Q}_p)$ .

Пусть  $s$  есть метрика из такого семейства и  $I$  есть  $s$ -шар, содержащий нуль ( $s$ -шар есть шар относительно  $s$ ).

Тогда растяжение  $pI$  есть также  $s$ -шар.

Содержащие нуль  $s$ -шары  $I$  и  $J$  смежны, если  $pI \subset J \subset I$ .

$k$ -Симплексы — наборы  $k$  смежных  $s$ -шаров

$$pI_k \subset I_1 \subset I_2 \subset \dots \subset I_k.$$

Если параметры  $q_i$  нормы — общего положения, то максимальная последовательность  $s$ -шаров между  $pI$  и  $I$  содержит  $d + 1$  шар и определяет  $d$ -симплекс.

(Параметры общего положения, если никакие два не могут быть сделаны равными умножением на степени  $p$ ).

Общие симплексы в комплексе шаров относительно метрики  $\{s_{q_1, \dots, q_d}^A\}$  — сдвиги описанных выше (как наборы подмножеств в  $\mathbb{Q}_p^d$ ).

Симплексиальный комплекс  $\mathcal{C}$  шаров относительно семейства метрик  $\{s_{q_1, \dots, q_d}^A\}$  — объединение комплексов шаров для каждой из метрик  $s_{q_1, \dots, q_d}^A$ .  $s$ -Шар и  $s'$ -шар, совпадающие как множества, совпадают как вершины (аналогично для  $s$  и  $s'$ -симплексов как наборов шаров).

Шар в  $\mathbb{Q}_p^d$  относительно любой из описанных норм, содержащий нуль, есть решётка.

$\mathcal{C}$  — симплексиальный комплекс шаров (для норм общего положения),

$\mathcal{C}_0$  — подкомплекс шаров с центром в нуле.

Фактор  $\mathcal{C}_0/\Gamma$  по группе растяжений на  $p^k$ ,  $k \in \mathbb{Z}$  также есть симплексиальный комплекс.

$\mathcal{C}_0/\Gamma$  изоморфен аффинному билдингу Брюа–Титса как симплексиальный комплекс.

**Размерность:** Ранг аффинного билдинга равен числу промежуточных решёток между решётками  $L$  и  $pL$ . Как перенести понятие размерности на произвольные сети кластеров?

Множества  $p\mathbb{Z}_p^d$  и  $\mathbb{Z}_p^d$  являются шарами для всех норм  $N_{q_1, \dots, q_d}$  при  $p^{-1} < q_i \leq 1$ . При этом для разных норм этого семейства набор промежуточных шаров между  $p\mathbb{Z}_p^d$  и  $\mathbb{Z}_p^d$  различен, и промежуточные шары не будут шарами для всех норм из описанного семейства одновременно.

Это свойство (шары являются шарами относительно семейства метрик) можно использовать как замену свойства "шары отличаются умножением на  $p$ ".

Используя такую точку зрения, можно перенести конструкцию симплексиального комплекса (типа аффинного билдинга) на произвольную сеть кластеров, связанную с мультикластеризацией, и ввести понятие размерности для сети кластеров.

## Общие сети кластеров

Пусть  $X$  есть локально компактное пространство с конечным набором ультраметрик  $\mathbf{S}$ . Пусть также для любых двух метрик  $s, r \in \mathbf{S}$  каждый  $s$ -шар есть конечное объединение  $r$ -шаров.

Дерево  $\mathcal{T}(X, s)$   $s$ -шаров в  $X$ : вершины  $s$ -шары, вершины соединены ребром, если соответствующие шары вложены без промежуточных.

Сеть  $\mathcal{C}(X, \mathbf{S})$  кластеров в  $X$  относительно семейства метрик  $\mathbf{S}$  — объединение деревьев  $\mathcal{T}(X, s)$   $s$ -шаров,  $s \in \mathbf{S}$ , склеенных по вершинам таким, что соответствующие шары совпадают как множества. Частичный порядок по вложению.

Выберем  $\mathbf{S}$ -шар  $I$  ( $s$ -шар относительно всех  $s \in \mathbf{S}$ ). Пусть  $J$  есть наименьший  $\mathbf{S}$ -шар, строго больший  $I$ .

Симплекс — подмножество набора промежуточных  $s$ -шаров между  $I$  и  $J$  (для метрики  $s \in \mathbf{S}$ ).  $k$ -Симплекс содержит  $k + 1$   $s$ -шаров, ребро в  $\mathcal{T}(X, s)$  есть 1-симплекс.

Объединение по метрикам  $s \in \mathbf{S}$  определяет структуру симплексиального комплекса на сети  $\mathcal{C}(X, \mathbf{S})$ .

## Размерность сети кластеров

Обобщить определение размерности (число  $p$ -адических параметров) на общие сети кластеров.

Пусть  $I, J$  есть пара  $\mathbf{S}$ -шаров,  $I \subset J$  и  $J$  есть минимальный  $\mathbf{S}$ -шар, содержащий  $I$ .

Рассмотрим максимальный  $s$ -симплекс,  $s \in \mathbf{S}$ , отвечающий  $I, J$  (максимальная последовательность  $s$ -шаров между  $I$  и  $J$ ).

Ранг такого симплекса назовём  $\mathbf{S}$ -размерностью пары  $I, J$  (для  $s \in \mathbf{S}$ , для которой такой ранг максимальен).

Таким образом, понятие размерности обобщается на любую сеть кластеров.

## Приложения к анализу данных.

Классификационные деревья (деревья кластеров) описывают разнообразие данных, сети кластеров отвечают многообразию способов классификации, размерность сетей кластеров описывает число источников разнообразия.

В некоторых случаях в задачах анализа данных существуют более далеко идущие аналогии с аффинными билдингами. На сети кластеров можно рассмотреть не только метрики, но и аналог структуры модуля (действие сдвигов).

## word2vec —

T.Mikolov, et al. Efficient Estimation of Word Representations in Vector Space, arXiv:1301.3781

Примерная схема:

Рассматривается представление слов из корпуса текстов векторами в гильбертовом пространстве. При этом корреляции между словами связаны со скалярным произведением в гильбертовом пространстве. Тогда семантика (смысл) слов оказывается связана с действием сдвигов в гильбертовом пространстве.

Примеры: после соответствующей нормировки

Мужчина – Женщина = Король – Королева,

Париж – Франция = Берлин – Германия.

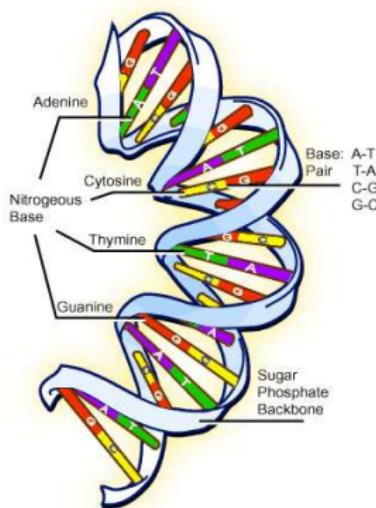
Генетический код — способ кодирования аминокислот тройками нуклеотидов.

Триплетная структура генетического кода,

Кодон — тройка  $C_1 C_2 C_3$  нуклеотидов,

Существует 64 кодона и 20 аминокислот — код вырожден.

Вырождение кода, химические свойства аминокислот — 2-адическая 2-мерная структура.



AAA Lys	UAA Ter	GAA Glu	CAA Gln
AAU Asn	UAU Tyr	GAU Asp	CAU His
AAG Lys	UAG Ter	GAG Glu	CAG Gln
AAC Asn	UAC Tyr	GAC Asp	CAC His
AUA Met	UUA Leu	GUA Val	CUA Leu
AUU Ile	UUU Phe	GUU Val	CUU Leu
AUG Met	UUG Leu	GUG Val	CUG Leu
AUC Ile	UUC Phe	GUC Val	CUC Leu
AGA Ter	UGA Trp	GGA Gly	CGA Arg
AGU Ser	UGU Cys	GGU Gly	CGU Arg
AGG Ter	UGG Trp	GGG Gly	CGG Arg
AGC Ser	UGC Cys	GGC Gly	CGC Arg
ACA Thr	UCA Ser	GCA Ala	CCA Pro
ACU Thr	UCU Ser	GCU Ala	CCU Pro
ACG Thr	UCG Ser	GCG Ala	CCG Pro
ACC Thr	UCC Ser	GCC Ala	CCC Pro

Митохондриальный генетический код.

## 2-Адическая плоскость кодонов

A.Yu. Khrennikov, S.V. Kozyrev, Genetic code on the diadic plane, Physica A: Statistical Mechanics and its Applications. 2007. V.381. P.265-272. arXiv:q-bio.QM/0701007

1) Нумерация нуклеотидов парами 0,1

$$\begin{array}{|c|c|} \hline A & G \\ \hline U & C \\ \hline \end{array} = \begin{array}{|c|c|} \hline 00 & 01 \\ \hline 10 & 11 \\ \hline \end{array} \quad (A)$$

1-ая строка — пурины

2-ая строка — пиримидины

1-й столбец — две водородные связи

2-й столбец — три водородные связи

2) Порядок нуклеотидов в кодоне

$$2 > 1 > 3 \quad (B)$$

3) 2-Адическая плоскость — группа  $\mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z}$  с координатами  $(x, y)$ :

$$x = (x_0 x_1 x_2) = x_0 + 2x_1 + 4x_2, \quad x_i = 0, 1,$$

$$y = (y_0 y_1 y_2) = y_0 + 2y_1 + 4y_2, \quad y_i = 0, 1,$$

и 2-мерной 2-адической метрикой

$$d_{1,q}((x, y), (x', y')) = \max(|x - x'|_2, q|y - y'|_2), \quad 1/2 < q < 1.$$

#### 4) Нумерация кодонов

Рассмотрим отображение

$$\rho : C_1 C_2 C_3 \mapsto (x, y) = (x_0 x_1 x_2, y_0 y_1 y_2),$$

$C_2$  порождает пару  $(x_0, y_0)$ ,

$C_1$  порождает пару  $(x_1, y_1)$ ,

$C_3$  порождает пару  $(x_2, y_2)$ .

Нуклеотиды определяют пары цифр по правилу (A),

Порядок нуклеотидов в кодоне определяется правилом (B).

## 5) Отображение Монна

Рассмотрим 2-мерное отображение Монна 2-адической плоскости на таблицу:

$$\eta : x \mapsto \tilde{x}, \quad y \mapsto \tilde{y};$$

$$\eta : x_0 + 2x_1 + 4x_2 \mapsto 1 + 4x_0 + 2x_1 + x_2;$$

$$\eta : y_0 + 2y_1 + 4y_2 \mapsto 1 + 4y_0 + 2y_1 + y_2.$$

Эквивалентно:

$$\eta : 0, 4, 2, 6, 1, 5, 3, 7 \mapsto 1, 2, 3, 4, 5, 6, 7, 8.$$

AAA	AAG	GAA	GAG	AGA	AGG	GGA	GGG
AAU	AAC	GAU	GAC	AGU	AGC	GGU	GGC
UAA	UAG	CAA	CAG	UGA	UGG	CGA	CGG
UAU	UAC	CAU	CAC	UGU	UGC	CGU	CGC
AUA	AUG	GUA	GUG	ACA	ACG	GCA	GCG
AUU	AUC	GUU	GUC	ACU	ACC	GCU	GCC
UUA	UUG	CUA	CUG	UCA	UCG	CCA	CCG
UUU	UUC	CUU	CUC	UCU	UCC	CCU	CCC

Применим к таблице кодонов митохондриальный генетический код:

Lys	Glu	Ter	Gly
Asn	Asp	Ser	
Ter	Gln	Trp	
Tyr	His	Cys	Arg
Met	Val	Thr	Ala
Ile			
Leu	Leu	Ser	Pro
Phe			

В частности

$$\begin{array}{|c|c|} \hline AAA & AAG \\ \hline AAU & AAC \\ \hline \end{array} \rightarrow \begin{array}{|c|} \hline \text{Lys} \\ \hline \text{Asn} \\ \hline \end{array}, \quad \begin{array}{|c|c|} \hline CCA & CCG \\ \hline CCU & CCC \\ \hline \end{array} \rightarrow \begin{array}{|c|} \hline \text{Pro} \\ \hline \end{array}$$

Вырождение кода описывается метрикой  $d_{1,q}(\cdot, \cdot)$  на 2-адической плоскости: код локально постоянен на шарах относительно  $d_{1,q}(\cdot, \cdot)$ .

## Физико-химические свойства (гидрофобность)

также кластеризуются в 2-адической плоскости.

Гидрофобные аминокислоты образуют два шара

—	—	—	
—	—	Trp Cys	
Met Ile	Val		
Leu Phe	Leu		

Генетический код имеет 2-мерную 2-адическую структуру, связанную с нумерацией нуклеотидов.

Вырождение кода — локальное постоянство относительно 2-мерной метрики.

Химические свойства аминокислот (гидрофобность) также кластеризуются.

## Базис вещественных всплесков в $L^2(\mathbb{R})$ .

Базис всплесков имеет параметризацию сдвигами и растяжениями (фиксированной функции либо конечного набора функций)

$$\psi_{jn}(x) = 2^{j/2} \psi(2^j x - n), \quad x \in \mathbb{R}, \quad j, n \in \mathbb{Z}.$$

Функция  $\psi(x)$  называется всплеск. Пример – всплеск Хаара (1909) (разность двух характеристических функций)

$$\psi(x) = \chi_{[0,1/2)}(x) - \chi_{[1/2,1]}(x).$$

Пара индексов  $(j, n)$  параметризует вершины дерева.

Более общие базисы всплесков, кратномасштабный анализ, S.Mallat, Y.Meyer, I.Daubechies, см.

I. Daubechies, Ten Lectures on Wavelets, SIAM, 1992.

Базис  $p$ -адических всплесков в  $L^2(\mathbb{Q}_p)$ :

$$\psi_{k;jn}(x) = p^{j/2} \psi_k(p^{-j}x - n),$$

$$x \in \mathbb{Q}_p, \quad j \in \mathbb{Z}, \quad n \in \mathbb{Q}_p/\mathbb{Z}_p, \quad k \in \{1, \dots, p-1\}.$$

Мы рассматриваем пример:

$$\psi_k(x) = \psi(kx), \quad \psi(x) = \chi(p^{-1}x)\Omega(|x|_p),$$

где  $\Omega(x)$  есть характеристическая функция  $[0, 1]$   
(то есть  $\Omega(|x|_p)$  есть характеристическая функция  $\mathbb{Z}_p$ ),  
 $\chi$  есть характер

$$\chi(x) = e^{2\pi i \{x\}}, \quad \{x\} = \sum_{l=a}^{-1} x_l p^l, \quad x = \sum_{l=a}^{\infty} x_l p^l.$$

*C.B. Козырев, Теория всплесков как  $p$ -адический спектральный анализ. Известия РАН Серия Мат. 2002. Т.66. N.2. С.149–158, arXiv:math-ph/0012019*

# Отображение Монна

$$\mathbb{Q}_p \rightarrow \mathbb{R}_+,$$

$$\mathbb{Q}_p/\mathbb{Z}_p \rightarrow \mathbb{Z}_+,$$

$$\sum_{l=a}^{\infty} x_l p^l \mapsto \sum_{l=a}^{\infty} x_l p^{-l-1}.$$

Малые  $p$ -адические расстояния отображаются на малые вещественные. Взаимно однозначно почти всюду и сохраняет меру.

$p = 2$ , отображение Монна: номера  $(j, n)$  всплесков в базисе Хаара на положительной полуоси переводит в вершины в 2-адическом дереве шаров.

Базис Хаара на положительной полуоси переходит в 2-адический базис всплесков почти всюду.

Вещественный базис всплесков имеет иерархическую структуру.

$$f(x) \mapsto |a|_p^{-\frac{1}{2}} f\left(\frac{x-b}{a}\right).$$

Орбита всплеска  $\psi(x) = \chi(p^{-1}x)\Omega(|x|_p)$  совпадает с набором произведений векторов из базиса  $\psi_{k;jn}$  на корни степени  $p$  из единицы.

Параметризация базиса всплесков возникает автоматически. Обобщение — фреймы всплесков.

**Фреймом**  $\{e_n\}$  в гильбертовом пространстве  $\mathcal{H}$  называется набор векторов в  $\mathcal{H}$ , такой, что:  $\exists A, B > 0: \forall g \in \mathcal{H}$

$$A\|g\|^2 \leq \sum_n |\langle g, e_n \rangle|^2 \leq B\|g\|^2.$$

Константы  $A, B$  называются нижней и верхней границей фрейма соответственно. Если  $A = B$ , то фрейм называется жёстким. Если  $\|e_n\| = \text{const}$ , то фрейм называется однородным.

## Фреймы $p$ -адических всплесков как орбиты (системы когерентных состояний)

Пусть  $f$  есть основная функция с нулевым средним на  $\mathbb{Q}_p$   
(основная – локально постоянная с компактным носителем).

Пусть также эта функция имеет общее положение  
(стабилизатор функции относительно действия аффинной  
группы есть пересечение стабилизаторов всплесков в  
разложении функции по всплескам).

Тогда орбита функции  $f$  относительно действия аффинной  
группы есть жёсткий однородный фрейм, причём фрейм имеет  
параметризацию сдвигами и растяжениями как у базиса  
 $p$ -адических всплесков.

Также есть явное выражение для границы фрейма.

Параметризация базиса (и фрейма) всплесков в  
кратномасштабном анализе всплесков в  $p$ -адическом случае  
возникает автоматически из вида системы когерентных  
состояний для группы автоморфизмов дерева шаров.

## Метрики в многомерных $p$ -адических пространствах

Метрика  $s_{q_1, \dots, q_d}(x, y)$  в  $\mathbb{Q}_p^d$  определена нормой  $N_{q_1, \dots, q_d}(z)$

$$s_{q_1, \dots, q_d}(x, y) = N_{q_1, \dots, q_d}(x - y),$$

$$N_{q_1, \dots, q_d}(z) = \max_{i=1, \dots, d} (q_i |z_i|_p), \quad q_i > 0.$$

Норма ( $A$ -вращение  $N_{q_1, \dots, q_d}$ ,  $A$  есть матрица из  $\mathrm{Gl}_d(\mathbb{Q}_p)$ )

$$N_{q_1, \dots, q_d}^A(z) = N_{q_1, \dots, q_d}(Az).$$

Растяжения  $p^k \mathbb{Z}_p^d$ ,  $k \in \mathbb{Z}$  являются шарами относительно всех норм  $N_{q_1, \dots, q_d}$ , если  $p^{-1} < q_i \leq 1$ . Для

$$p^{-1} < q_1 < \dots < q_d \leq 1$$

набор промежуточных  $N_{q_1, \dots, q_d}$ -шаров между  $p\mathbb{Z}_p^d$  и  $\mathbb{Z}_p^d$

$$B_j = \mathbb{Z}_p \times \dots \times \mathbb{Z}_p \times p\mathbb{Z}_p \times \dots \times p\mathbb{Z}_p$$

с  $j$  компонентами  $\mathbb{Z}_p$  и  $d - j$  компонентами  $p\mathbb{Z}_p$ ,  $j = 0, \dots, d$ .

## Многомерные базисы всплесков как орбиты. Общая схема:

- 1) Метрика в  $\mathbb{Q}_p^d$ .
- 2) Группа автоморфизмов дерева шаров относительно метрики (линейные преобразования плюс сдвиги).
- 3) Базисы и фреймы всплесков как орбиты такой группы.

Пример: норма

$$\|z\|_p = \max_{i=1,\dots,d}(|z_i|_p).$$

Группа автоморфизмов дерева шаров — сдвиги; однородные растяжения; линейные преобразования, сохраняющие норму (матрицы с матричными элементами из  $\mathbb{Z}_p$  и  $|\det(\cdot)|_p = 1$ .)

Всплески — произведения характеров  $\mathbb{Q}_p^d$  на характеристические функции шаров, причём характер на шаре делает одно колебание. Таким образом, всплески нумеруются шарами и волновыми векторами (максимальными подшарами в единичной сфере).

**Теорема** Набор функций  $\{\psi_{k;jn}\}$  является ортонормированным базисом в  $L^2(\mathbb{Q}_p^d)$ :

$$\psi_k(x) = \chi(p^{-1}k \cdot x)\Omega(\|x\|_p), \quad x \in \mathbb{Q}_p^d, \quad k \cdot x = \sum_{l=1}^d k_l x_l,$$

$$k = (k_1, \dots, k_d), \quad k_l = 0, \dots, p-1,$$

где хотя бы один из  $k_l$  не равен нулю;

$$\psi_{k;jn}(x) = p^{-\frac{dj}{2}} \psi_k(p^j x - n), \quad x \in \mathbb{Q}_p^d, \quad j \in \mathbb{Z}, \quad n \in \mathbb{Q}_p^d / \mathbb{Z}_p^d,$$

$$n = (n^{(1)}, \dots, n^{(d)}), \quad n^{(l)} = \sum_{i=\beta_l}^{-1} n_i^{(l)} p^i,$$

$$n_i^{(l)} = 0, \dots, p-1, \quad \beta_l \in \mathbb{Z}_-.$$

Альтернативные выборы метрики — базисы всплесков с матричными растяжениями (quincunx basis, etc.)

Базис  $p$ -адических всплесков есть базис из собственных векторов оператора Владимира

$$D^\alpha \psi_{k;jn} = p^{\alpha(1-j)} \psi_{k;jn}.$$

$$D^\alpha f(x) = \int_{\mathbb{Q}_p} \frac{f(x) - f(y)}{|x - y|_p^{1+\alpha}} d\mu(y), \quad \alpha > 0,$$

$$\Gamma_p(-\alpha) = \frac{p^\alpha - 1}{1 - p^{-1-\alpha}}.$$

Спектры более общих операторов Оператор на локально компактном ультраметрическом пространстве  $X$  с борелевской мерой  $\mu$

$$Tf(x) = \int_X t(\sup(x, y))(f(x) - f(y)) d\mu(y).$$

$\sup(x, y)$  есть минимальный шар, содержащий точки  $x, y$ .

A.Yu.Khrennikov, S.V.Kozyrev, Wavelets on ultrametric spaces, Applied and Computational Harmonic Analysis. 2005. V.19. P. 61-76.

Пусть  $X = \mathbb{Q}_p$  и  $\mu$  есть мера Хаара.

**Теорема** Пусть следующий ряд сходится:

$$\sum_{j=0}^{\infty} p^j |t(I_{j0})| < \infty.$$

Тогда оператор  $T$  имеет плотную область определения в пространстве  $L^2(\mathbb{Q}_p)$ , диагонализуется в базисе  $p$ -адических всплесков и имеет в этом базисе собственные значения

$$\lambda_{jn} = p^j t(I_{jn}) + (1 - p^{-1}) \sum_{j'=j+1}^{\infty} p^{j'} t(I_{j', p^{j'-j} n}).$$

Здесь  $I_{jn}$  есть шар с характеристической функцией  $\Omega(|p^j x - n|_p)$ . Суммирование идёт по возрастающей последовательности шаров, содержащих  $I_{jn}$ .

Собственное значение  $\lambda_{jn}$  оператора  $T$  в базисе всплесков зависит от двух индексов  $j$  и  $n$  (диаметра и положения шара, являющегося носителем всплеска, то есть от самого шара).

## Заключение

Рассмотрены примеры многомерных иерархических систем.

$p$ -Адические всплески — деревья шаров, группы автоморфизмов таких деревьев, базисы всплесков — системы когерентных состояний.

Анализ данных — кластеризация, мультикластеризация, сети кластеров, аналогия с аффинными билдингами, размерность общих сетей кластеров.

Простой пример многомерной иерархии — 2-адическая плоскость генетического кода.