

С. В. Козырев

МИАН им. В.А.Стеклова

p -Адические параметры в теории сложных систем

6 июля, 2010, МФТИ, летняя Школа ПМФ 2010

Спиновые стёкла – p -адическая параметризация
матрицы Паризи

Динамика белка – уравнение p -адической диффузии

Генетический код – 2-адическая плоскость

Сложные системы

Спиновые стёкла — модель Шеррингтона–Киркпатрика

$$H = \sum_{i,j=1}^N J_{ij} \sigma_i \sigma_j$$

$\sigma_i = \pm 1$ – спины

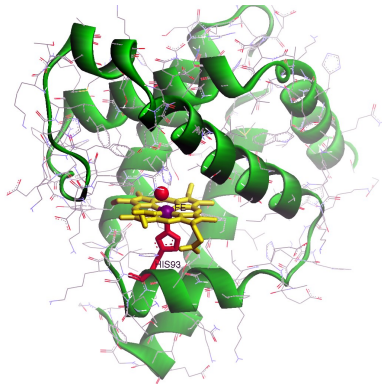
J_{ij} (беспорядок) – симметрическая вещественная гауссовская случайная матрица $J_{ij} = J_{ji}$,

нулевое среднее $\langle J_{ij} \rangle = 0$

матричные элементы независимы, то есть ненулевые квадратичные корреляционные функции только $\langle J_{ij}^2 \rangle = 1/N$

беспорядок — флуктуации в процессе приготовления образца
статистическая физика неупорядоченных систем

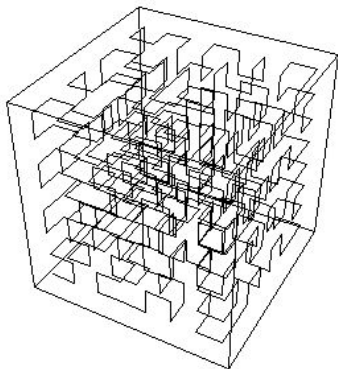
Белок — цепь аминокислот, свёрнутая в компактную глобулу
(нативное состояние)



Миоглобин

проблема устойчивости, проблема фолдинга

глобула белка схожа с кривой Пеано
(заполнение кривой пространства)

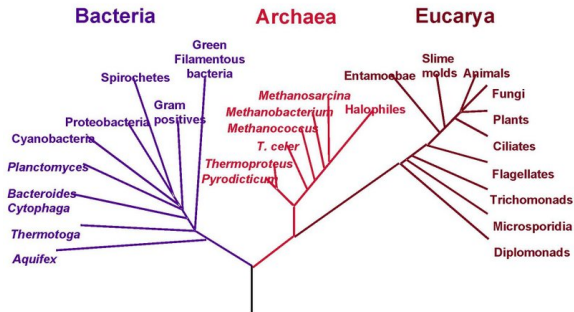


ДНК упакована также аналогично кривой Пеано

кривая Пеано устроена иерархически

Биология — сложные системы (уже на молекулярном уровне)

Phylogenetic Tree of Life



филогенетическое древо

p -Адические числа

p -Адическая норма рационального числа: для простого p

$$x = p^\gamma \frac{m}{n}, \quad |x|_p = p^{-\gamma}, \quad |0|_p = 0.$$

Поле \mathbb{Q}_p p -адических чисел —
пополнение \mathbb{Q} по p -адической норме
 p -адические числа — ряды

$$x = \sum_{i=\gamma}^{\infty} x_i p^i, \quad x_i = 0, \dots, p-1.$$

Мера Хаара: инвариантность относительно сдвигов,
нормировка:

$$\{x : |x|_p \leq 1\} = \left\{ \sum_{i=0}^{\infty} x_i p^i \right\}$$

шар с мерой единица.

Матрица Паризи и p -адическое расстояние

Блочнo–иерархическая матрица Паризи (Q_{ab})

(теория спиновых стёкол, метод реплик)

$$\begin{pmatrix} 0 & q_1 \\ q_1 & 0 \end{pmatrix}, \quad \left(\begin{array}{cc|cc} 0 & q_1 & q_2 & q_2 \\ q_1 & 0 & q_2 & q_2 \\ \hline q_2 & q_2 & 0 & q_1 \\ q_2 & q_2 & q_1 & 0 \end{array} \right),$$

$$\left(\begin{array}{cccc|cccc} 0 & q_1 & q_2 & q_2 & q_3 & q_3 & q_3 & q_3 \\ q_1 & 0 & q_2 & q_2 & q_3 & q_3 & q_3 & q_3 \\ q_2 & q_2 & 0 & q_1 & q_3 & q_3 & q_3 & q_3 \\ q_2 & q_2 & q_1 & 0 & q_3 & q_3 & q_3 & q_3 \\ \hline q_3 & q_3 & q_3 & q_3 & 0 & q_1 & q_2 & q_2 \\ q_3 & q_3 & q_3 & q_3 & q_1 & 0 & q_2 & q_2 \\ q_3 & q_3 & q_3 & q_3 & q_2 & q_2 & 0 & q_1 \\ q_3 & q_3 & q_3 & q_3 & q_2 & q_2 & q_1 & 0 \end{array} \right)$$

q_i суть вещественные параметры.

Перестановка строк и столбцов матрицы Паризи $p^N \times p^N$
(в примере $p = 2$ и $N = 1, 2, 3$)

$$I : \{1, \dots, p^N\} \rightarrow p^{-N}\mathbb{Z}/\mathbb{Z},$$

$$I^{-1} : \sum_{j=-N}^{-1} x_j p^j \mapsto 1 + p^{-1} \sum_{j=-N}^{-1} x_j p^{-j}.$$

выражает матричный элемент Q_{ab} матрицы Паризи как
функцию p -адического расстояния между $I(a)$ и $I(b)$:

$$Q_{ab} = q(|I(a) - I(b)|_p),$$

где $q(p^k) = q_k$, $q(0) = 0$, $k = 1, \dots, N$.

Группа $p^{-N}\mathbb{Z}/\mathbb{Z}$ рассматривается как группа дробей
 $\{\sum_{j=-N}^{-1} x_j p^j\}$, $x_j = 0, \dots, p-1$ со сложением по модулю 1.

иерархия состояний спинового стекла

Динамика на ландшафтах энергии

Уравнение диффузии в потенциале

$$\frac{\partial}{\partial t} f(x, t) = \Delta f(x, t) + \beta \nabla f(x, t) \cdot \nabla U(x) + \beta f(x, t) \Delta U(x)$$

$f(x, t)$ – функция распределения

U – потенциал, $\beta = 1/kT$ – обратная температура.

Приближённое описание диффузии в потенциале
кинетическими уравнениями

переходы между потенциальными ямами

– локальными минимумами энергии

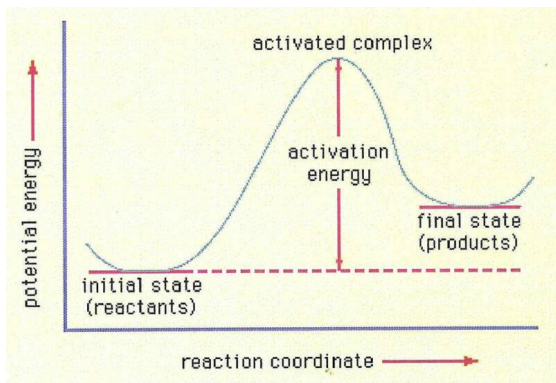
$$\frac{d}{dt} f(a, t) = \sum_b (Q_{ab} f(b, t) - Q_{ba} f(a, t)).$$

Скорость перехода – формула Аррениуса

– скорость пропорциональна

$$\exp(-\beta\Delta E), \quad \Delta E = E_1 - E_0$$

ΔE – барьер активации.



Сложные энергетические ландшафты

Сложный ландшафт — много локальных минимумов.
Как построить аррениусовскую кинетику?

Наблюдение:

три локальных минимума – два переходных состояния.
отсюда следует ультраметричность.

Пример: локальные минимумы A, B, C ,
переходное состояние с энергией E_1 между A и B ,
переходное состояние с энергией E_2 между (A, B) и C , $E_1 < E_2$
матрица переходных энергий

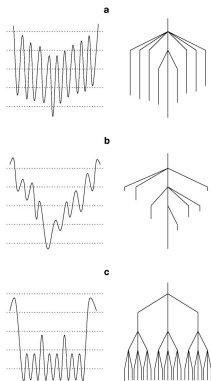
$$\left(\begin{array}{cc|c} 0 & E_1 & E_2 \\ E_1 & 0 & E_2 \\ \hline E_2 & E_2 & 0 \end{array} \right).$$

Пример: локальные минимумы A, B, C, D ,
переходное состояние с энергией E_1 между A и B ,
переходное состояние с этой же энергией между C и D ,
переходное состояние с энергией $E_2 > E_1$ между парами
минимумов (A, B) и (C, D)
матрица переходных энергий

$$\left(\begin{array}{cc|cc} 0 & E_1 & E_2 & E_2 \\ E_1 & 0 & E_2 & E_2 \\ \hline E_2 & E_2 & 0 & E_1 \\ E_2 & E_2 & E_1 & 0 \end{array} \right).$$

таблица расстояний для ультраметрики

Общий случай — дерево локальных минимумов и переходных состояний (**disconnectivity graph**)



иерархия бассейнов

барьеры активации — задают ультраметрическое расстояние между локальными минимумами

$$d(x, y) = E(\text{sup}(x, y))$$

Межбассейновая кинетика

Система кинетических уравнений

$$\frac{d}{dt}f(a, t) = \sum_b (Q_{ab}f(b, t) - Q_{ba}f(a, t)).$$

Сложный ландшафт — (Q_{ab}) есть матрица Паризи.

Блоки матрицы Паризи — переходы между бассейнами

— межбассейновая кинетика

p -адическая параметризация матрицы Паризи даёт

$$\frac{\partial}{\partial t} f(a, t) = \int_{p^M \mathbb{Z}_p / p^N \mathbb{Z}_p} q(|a-b|_p) (f(b, t) - f(a, t)) d\mu(b), \quad M < N,$$

интегрирование (суммирование)

по мере Хаара на группе $p^M \mathbb{Z}_p / p^N \mathbb{Z}_p$,

нормировка меры: $\mu(\mathbb{Z}_p / p^N \mathbb{Z}_p) = 1$.

Предел $M \rightarrow -\infty$, $N \rightarrow +\infty$

бесконечно много локальных минимумов

интегральное уравнение межбассейновой кинетики

$$\frac{\partial}{\partial t} f(x, t) = \int_{\mathbb{Q}_p} q(|x-y|_p) (f(y, t) - f(x, t)) d\mu(y).$$

p -Адическая диффузия.

Пример:

$$q(|x|_p) = \Gamma_p^{-1}(-\alpha)|x|_p^{-1-\alpha},$$

уравнение p -адической диффузии

$$\frac{\partial}{\partial t} f(x, t) + D_x^\alpha f(x, t) = 0$$

порождаемой дробным оператором Владимирова

$$D_x^\alpha f(x, t) = \Gamma_p^{-1}(-\alpha) \int_{\mathbb{Q}_p} \frac{f(x, t) - f(y, t)}{|x - y|_p^{1+\alpha}} d\mu(x)$$

α пропорционально обратной температуре: $\alpha = \beta k$.

— описывает динамику белка.

x — конформации белка

Воспроизводит результаты экспериментов Фрауенфельдера по динамике миоглобина

Генетический код

ДНК (РНК) — конечные последовательности
(линейные полимеры) нуклеотидов

белок — конечная последовательность аминокислот

Нуклеотиды: С, А, Т (или U), G

Cytosine, Adenine, Thymine (или Uracil), Guanine

Триплетная структура генетического кода (Гамов)

Кодон — тройка $C_1 C_2 C_3$ нуклеотидов

Генетический код — переводит кодоны в аминокислоты

20 аминокислот и 1 символ остановки

4 нуклеотида, $4 \times 4 \times 4 = 64$ кодонов

$64 > 21$ — проблема вырождения генетического кода

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

митохондриальный генетический код

2-Адическая плоскость кодонов

1) Занумеруем нуклеотиды парами 0,1

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

Химический смысл:

1-ая строка — пурины

2-ая строка — пиримидины

1-ый столбец — слабая Н-связь

2-ой столбец — сильная Н-связь

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) :

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

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

с 2-мерной 2-адической метрикой

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

4) Отображение кодонов

ρ переводит кодон в точку 2-адической плоскости

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

C_2 определяет пару (x_0, y_0) ,

C_1 определяет пару (x_1, y_1) ,

C_3 определяет пару (x_2, y_2) .

Нуклеотиды определяют пары цифр по правилу (A),
порядок пар задаётся правилом (B).

5) Перенумеровка строк и столбцов

Занумеруем строки и столбцы 2-адической плоскости:

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

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

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

Эквивалентно:

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

(чтобы близкие 2-адически элементы
были расположены в соседних клетках плоскости)

2-Адическая плоскость кодонов

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

с метрикой

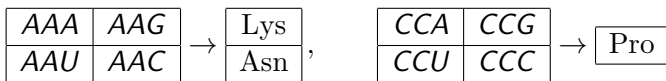
$$d(C_1 C_2 C_3, C'_1 C'_2 C'_3) = \max(|x - x'|_2, q|y - y'|_2), \quad 1/2 < q < 1,$$

$$(x, y) = \rho(C_1 C_2 C_3), \quad (x', y') = \rho(C'_1 C'_2 C'_3).$$

Применим к плоскости кодонов митохондриальный код:

$\frac{\text{Lys}}{\text{Asn}}$	$\frac{\text{Glu}}{\text{Asp}}$	$\frac{\text{Ter}}{\text{Ser}}$	Gly
$\frac{\text{Ter}}{\text{Tyr}}$	$\frac{\text{Gln}}{\text{His}}$	$\frac{\text{Trp}}{\text{Cys}}$	Arg
$\frac{\text{Met}}{\text{Ile}}$	Val	Thr	Ala
$\frac{\text{Leu}}{\text{Phe}}$	Leu	Ser	Pro

В частности



Вырождение генетического кода описывается метрикой $d(\cdot, \cdot)$: кодоны, отображающиеся на одинаковые аминокислоты — шары относительно метрики $d(\cdot, \cdot)$.

Иерархические методы для сложных систем

Спиновые стёкла – p -адическая параметризация матрицы Паризи

Динамика белка – уравнение p -адической диффузии

Генетический код – 2-адическая плоскость

непостижимая эффективность математики
в естественных науках

непостижимая неэффективность математики в биологии

возможно, нужны новые математические методы