Do we need Number Theory?

Václav Snášel

What is a number?

MCDXIX

||||||||||||||||||||

664554

0xABCD

01717

01010101011011100001

i^i

$$1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\cdots$$

СОВРЕМЕННАЯ МАТЕМАТИКА
ДЛЯ СТУДЕНТОВ

А. А. КИРИЛЛОВ

ЧТО ТАКОЕ ЧИСЛО?

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[А] Атья М. Лекции по К-теории.— М.: Мир, 1967.

[Б] Березин Ф. А. Метод вторичного квантования.— М.: Наука, 1965.—2-е изд.— М.: Наука, 1986.

[БР] Браттели О., Робинсон Д. В. Операторные алгебры и квантовая статистическая механика.— М.: Мир, 1985.

[В] Вейль А. Основы теории чисел.— М.: Мир, 1972.

[ВВ] Владимиров В. С., Волович И. В. Введение в суперанализ // Теорет. и мат. физика.— 1984.— Т. 59, № 1.— С. 3—27; Т. 60, № 2.— С. 169—198.

[ГГПШ] Гельфанд И. М., Граев М. И., Пятецкий-Шапиро И. И. Теория представлений и автоморфные функции.— М.: Наука, 1966.— (Обобщенные функции.— Вып. 6.)

[ГК] Гельфанд И. М., Кириллов А. А. Структура тела, связанного с полупростой расщепимой алгеброй Ли // Функцион. анализ и его прил.— 1969.— Т. 3, № 1.— С. 7—26.

[ГМ] Гельфанд С. И., Манин Ю. И. Методы гомологической алгебры. 1: Введение в теорию когомологий и производные категории.— М.: Наука, 1988.

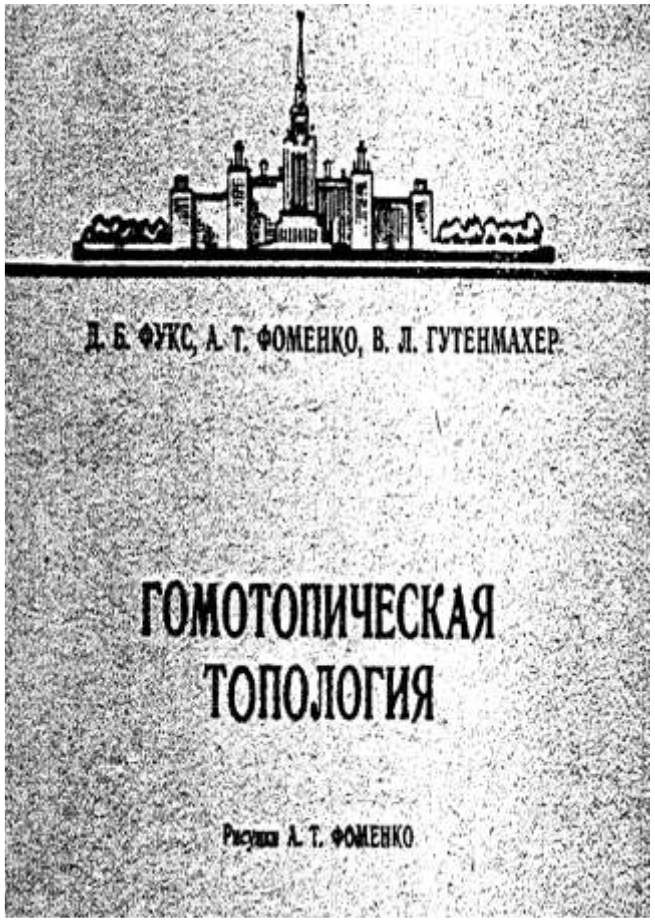
МОСКВА
ИЗДАТЕЛЬСКАЯ ФИРМА
«ФИЗИКО-МАТЕМАТИЧЕС
ВО «НАУКА»
1993

References

- Neal Koblitz, p-adic Numbers, p-adic Analysis, and Zeta-functions, Springer, 1977
- Z. I. Borevich, I. R. Shafarevich, Number theory, Academic Press, 1986
- Fernando Q. Gouvea, p-adic Numbers An Introduction, Springer 1997
- W.H.Schikhof, Ultrametric calculus, An introduction to p-adic analysis, Cambridge University Press, 1984
- Journal, p-Adic Numbers, Ultrametric Analysis and Applications,

  ISSN 2070-0466,

Images


Д. Б. ФУКС, А. Т. ФОМЕНКО, В. Л. ГУТЕНМАХЕР
ГОМОТОПИЧЕСКАЯ ТОПОЛОГИЯ
Рисунки А. Т. ФОМЕНКО

What Kind of Number is $\sqrt{2}^{\sqrt{2}}$

- Pythagoras noted that if the sides of the right-angle of a triangle have measures equal to 1, then the hypotenuse, measured by $\sqrt{2}$, is not a rational number.
- In 1900, in his 7th problem, Hilbert asked whether the following is true: if α is any algebraic number ($\alpha \neq 0,1$), and if β is any irrational algebraic number, then α^{β} is transcendental.

- In 1934, Gelfond and Schneider independently, and with different methods, solved Hilbert's 7th problem in the affirmative.

Fibonacci coding

- $F_n = F_{n-1} + F_{n-2}$ $\quad F_0 = 0, F_1 = 1$
- $N = \sum_{i=0}^{\infty} a_i F_{i+2}$ $\quad 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55$
- Zeckendorf's theorem states that every positive integer can be represented uniquely as the sum of one or more distinct Fibonacci numbers in such a way that the sum does not include any two consecutive Fibonacci numbers.

Fast Fibonacci decoding

Definition. Fibonacci shift operation Let $F(n)$ be a Fibonacci binary encoding, k be an integer, $k \geq 0$. The k th Fibonacci left shift $F(n) \ll_F k$ is defined as follows:

$$F(n) \ll_F k = 00 \ldots 0 a_0 a_1 a_2 \ldots a_p$$

Fibonacci right shift is defined as is follows:

$$F(n) \gg_F k = a_k a_{k+1} a_{k+2} \ldots a_p$$

Fast Fibonacci decoding

V is value of Fibonacci code.

Theorem. Calculation of the k-Fibonacci left shift. Let $F(n)$ be the Fibonacci binary encoding then

$$V(F(n) \ll_F k) = F_{k-1} \times V\big(F(n)\big) + F_{k-2}(F(n) \gg_F 1)$$

Jirí Walder, Michal Krátký, Radim Baca, Jan Platos, Václav Snásel: Fast decoding algorithms for variable-lengths codes. Inf. Sci. 183(1): 66-91 (2012)

Absolute Values on a Field

If we let $\mathbb{K}$ be a field and $\mathbb{R}_+ = \{x \in \mathbb{R}: x \geq 0\}$ then an absolute value on $\mathbb{K}$ is a function

$$|.| : \mathbb{K} \to \mathbb{R}_+$$

that satisfies

- $|x| = 0$ if and only if $x = 0$
- $|xy| = |x||y|$
- $|x + y| \leq |x| + |y|$ (Triangle inequality)

The Ordinary Absolute Value

The ordinary absolute value on $\mathbb{Q}$ is defined as follows:

$$|.| : \mathbb{Q} \to \mathbb{R}_+$$

$$|x| = \begin{cases} x : x \geq 0 \\ -x : x < 0 \end{cases}$$

This satisfied the required conditions.

Metric Spaces

A metric is a „distance“ function, defined as follows:

If X is a set, then a metric on X is a function d

$$d: X \times X \to \mathbb{R}_+$$

which satisfied the following properties:

- $d(x, x) \geq 0$
- $d(x, y) = d(y, x)$
- $d(x, y) + d(y, z) \geq d(x, z)$ (Triangle inequality)

(X, d) is called metric space.

The Rationals as a Metric Space

$\mathbb{Q}$ forms a metric space with the ordinary absolute value as our distance function.

We write this metric space as $(\mathbb{Q}, |.|)$

If X is a set, then a metric on X is a function d

The metric, d, is defined in the obvious way:

$$d: \mathbb{Q} \times \mathbb{Q} \to \mathbb{R}_+$$
$$d(x, y) = |x - y|$$

Cauchy Sequences

A Cauchy sequence in a metric space is a sequence whose elements become „close“ to each other.

A sequence

$$x_1, x_2, x_3, x_4 \cdots$$

is called Cauchy if for every positive (real) number ε, there is a positive integer N such that for all natural numbers $n, m > N$,

$$d(x_m, x_n) = |x_m, x_n| < \varepsilon$$

Complete Metric Space

We call a metric space (X, d) complete if every Cauchy sequence in (X, d) converges in (X, d)

Concrete example: the rational numbers with the ordinary distance function, $(\mathbb{Q}, |.|)$ is not complete.

Example: $(\sqrt{2})$

$$1, 1.4, 1.41, 1.414, \ldots$$

Completing $\mathbb{Q}$ to get $\mathbb{R}$

If a metric space is not complete, we can complete it by adding in all the „missing“ points.

For $(\mathbb{Q}, |.|)$, we add all the possible limits of all the possible Cauchy sequences.

We obtain $\mathbb{R}$.

It can be proven that the completion of field gives a field.

Since $\mathbb{Q}$ is a field, $\mathbb{R}$ is field.

The p-adic Absolute Value

For each prime p, there is associated p-adic absolute value $|.|_p$ on $\mathbb{Q}$.

Definition. Let p be any prime number. For any nonzero integer a, let $ord_p a$ be the highest power of p which divides a , i.e., the greatest m such that $a \equiv 0 \ (mod\ p^m)$.

$$ord_p ab = ord_p a + ord_p b, \qquad ord_p\ a/b = ord_p a - ord_p\ b,$$

Examples:

$ord_5 35 = 1, ord_5 77 = 0, ord_2 32 = 5$

The p-adic Absolute Value

Further define absolute value $|.|_p$ on $\mathbb{Q}$ as follows: ($a \in \mathbb{Q}$)

$$|a|_p = \begin{cases} p^{-ord_p a}, & a \neq 0 \\ 0, & a = 0 \end{cases}$$

Proposition. $|.|_p$ is a norm on $\mathbb{Q}$.

Example: $|\frac{968}{9}|_{11} = |11^2.\ \frac{8}{9}|_{11} = 11^{-2}$

Completing $\mathbb{Q}$ a different way

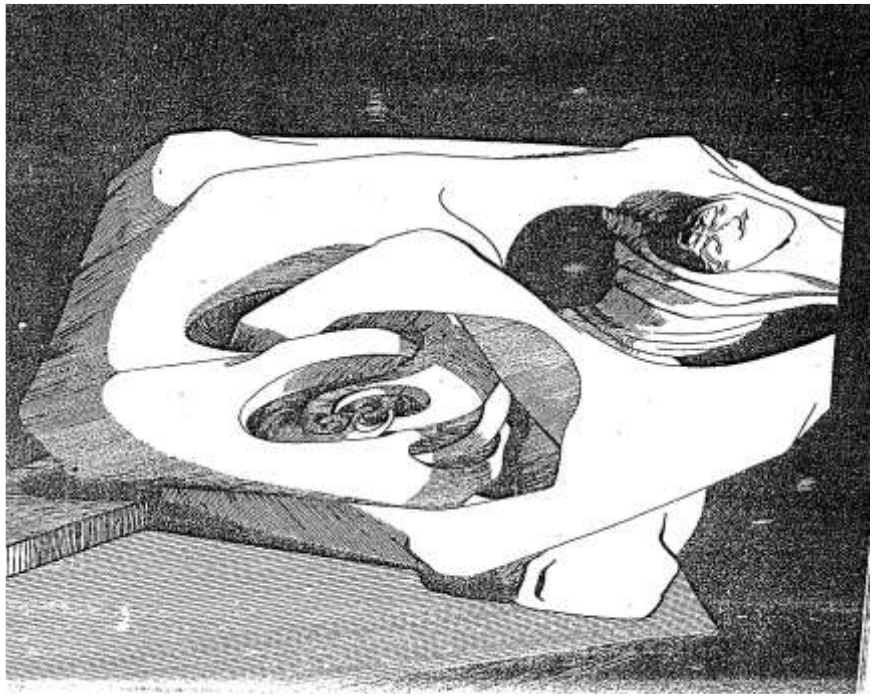
The p-adic absolute value give us a metric on $\mathbb{Q}$ defined by

$$d: \mathbb{Q} \times \mathbb{Q} \to \mathbb{R}_+$$
$$d(x, y) = |x - y|_p$$

When $p = 7$ we have that 7891 and 2 are closer together than 3 and 2

$$|7891 - 2|_7 = |7889|_7 = |7^3 \times 23|_7 = 7^{-3} = 1/343$$
$$|3 - 2|_7 = |1|_7 = |7^0|_7 = 7^0 = 1 > 1/343$$

Completing $\mathbb{Q}$ a different way



The p-adic absolute value give us a metric on $\mathbb{Q}$ defined by

$$d: \mathbb{Q} \times \mathbb{Q} \to \mathbb{R}_+$$
$$d(x, y) = |x - y|_p$$

When $p = 7$ we have that 7891 and 2 are closer together than 3 and 2

$$|7891 - 2|_7 = |7889|_7 = |7^3 \times 23|_7 = 7^{-3} = 1/343$$
$$|3 - 2|_7 = |1|_7 = |7^0|_7 = 7^0 = 1 > 1/343$$

Completing $\mathbb{Q}$ a different way

$\mathbb{Q}$ is not complete with respect to p-adic metric $d(x, y) = |x - y|_p$.

Example:

Let $p = 7$. The infinite sum

$$1 + 7 + 7^2 + 7^3 + 7^4 + 7^5 + \cdots$$

is certainly not element of $\mathbb{Q}$ but sequence

$$1, 1 + 7, 1 + 7 + 7^2, 1 + 7 + 7^2 + 7^3, \ldots$$

is a Cauchy sequence with respect to the 7-adic metric.

Completion of $\mathbb{Q}$ by $|x - y|_p$ gives field $\mathbb{Q}_p$: field of p-adic number.

The p-adic Absolute Value

Definition. A norm is called non-Archimedean if

$$|x + y| \leq \max(|x|, |y|)$$

always holds. A metric is called non-Archimedean if

$$d(x, z) \leq \max(d(x, y), d(y, z))$$

in particular, a metric is non-Archimedean if it is induced by a non-Archimedean norm.

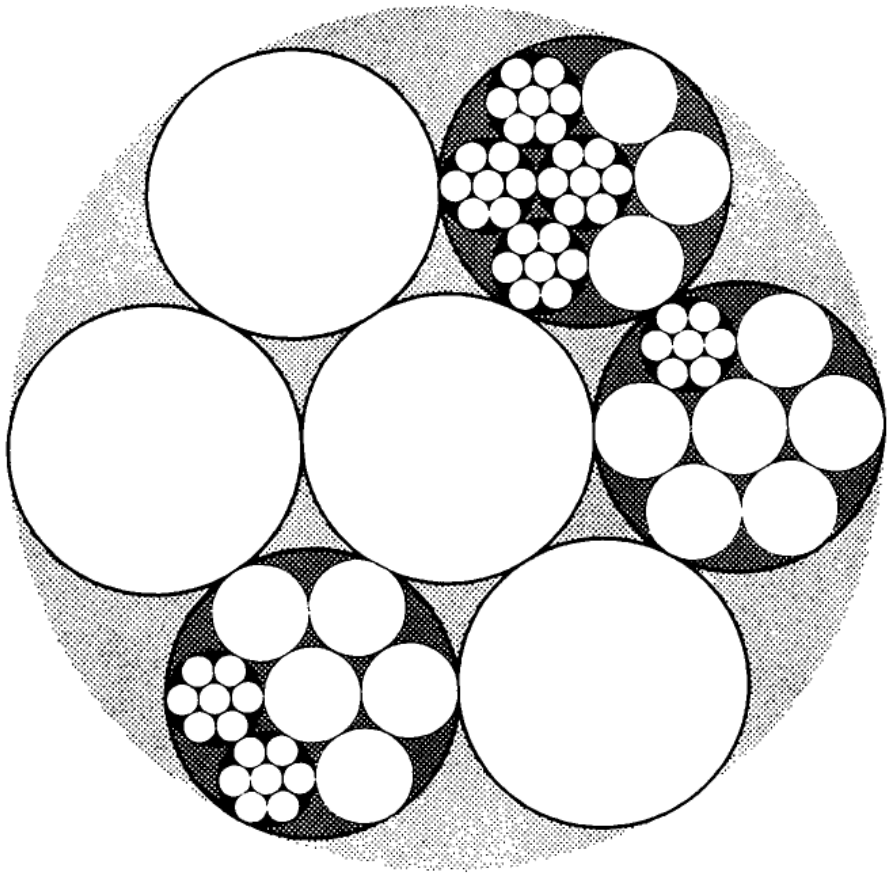
Thus, $|.|_p$ is a non-Archimedean norm on $\mathbb{Q}$.

Theorem (Ostrowski). Every nontrivial norm $|.|$ on $\mathbb{Q}$ is equivalent to $|.|_p$ for some prime p or the ordinary absolute value on $\mathbb{Q}$.

Basic property of a non-Archimedean field

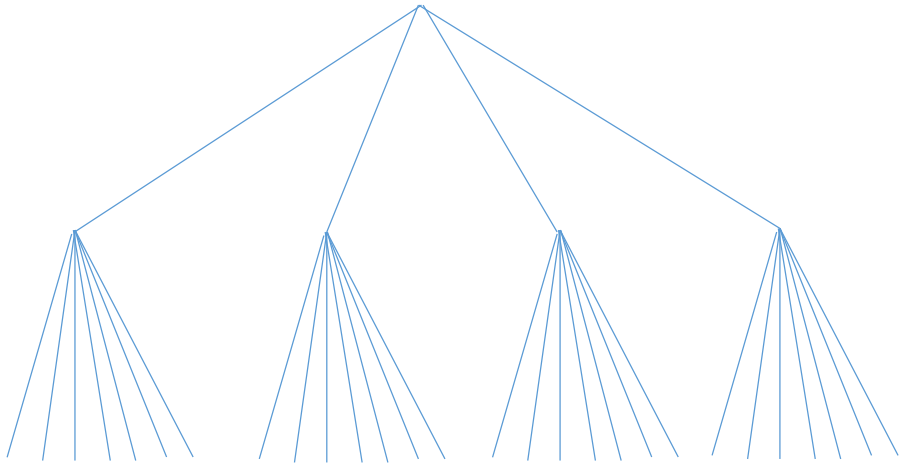
- Every point in a ball is a center!
- Set of possible distances are „small“
  $\{p^n; n \in \mathbb{Z}\}$
- every triangle is isosceles

Balls in $\mathbb{Q}_7$

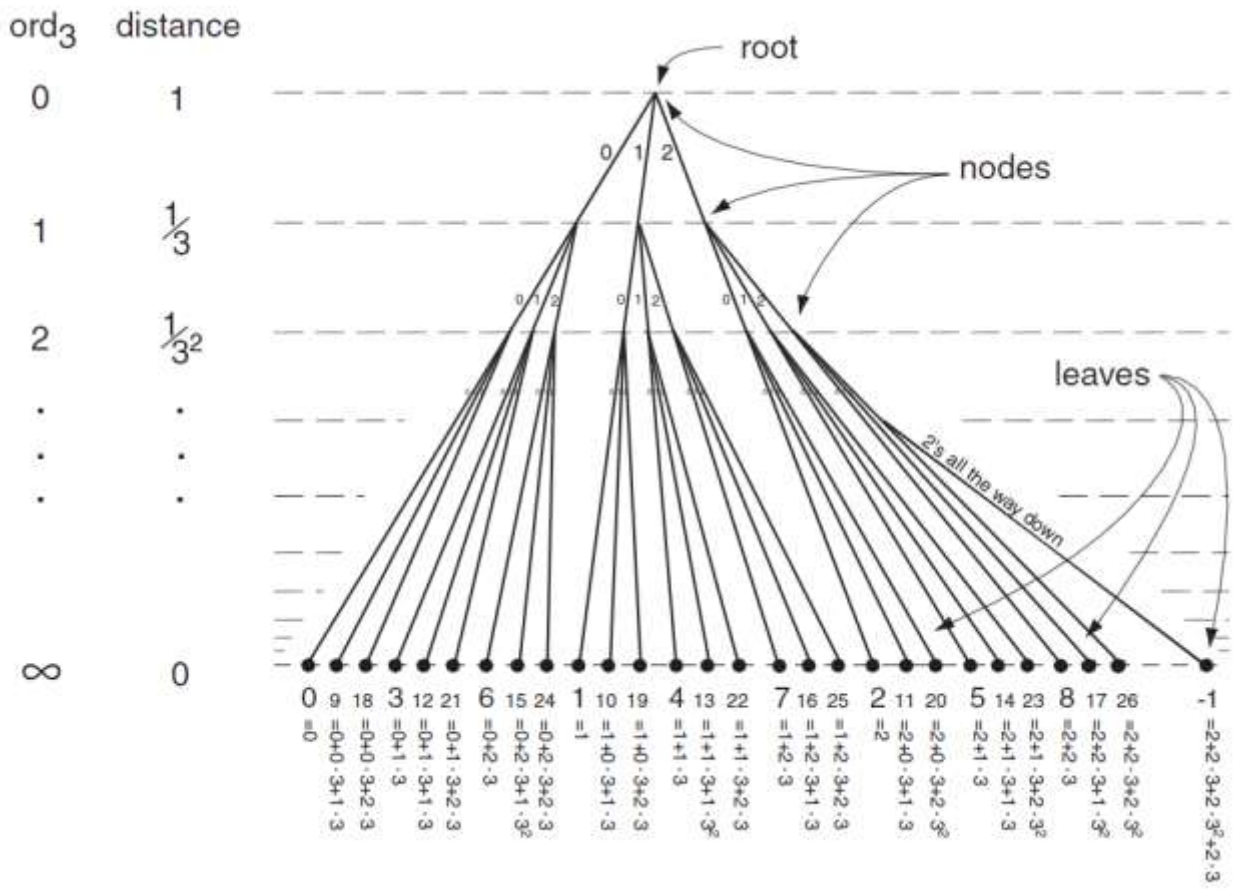


Definition. A metric space (X, d) is an ultrametric space if the metric d satisfies the strong triangle inequality $d(x, z) \leq \max(d(x, y), d(y, z))$.

Vizialization of ultrametrics



Visualization of $\mathbb{Q}_7$



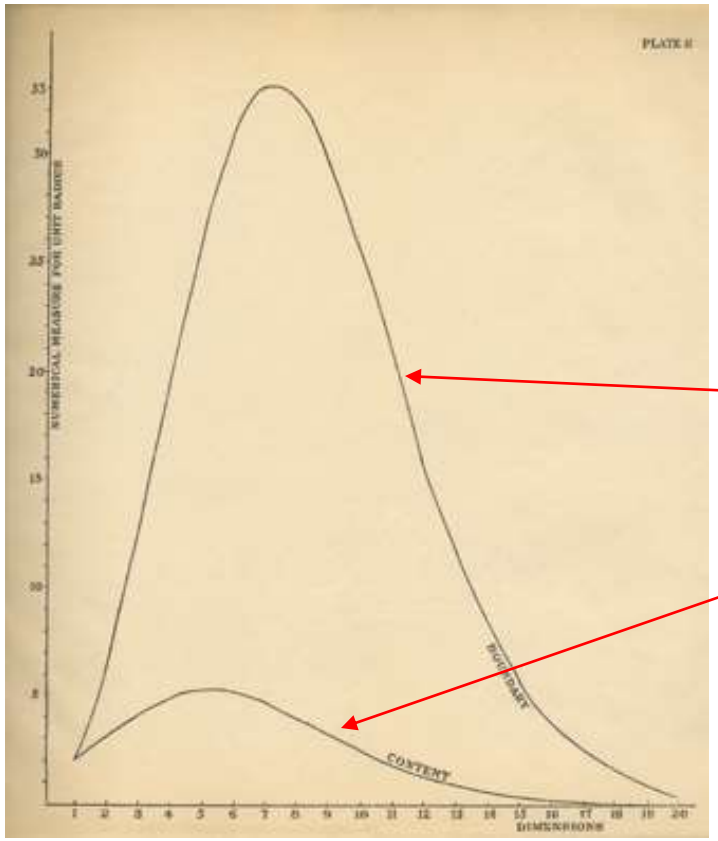
Curse of Dimensionality

Curse of Dimensionality

- The curse of dimensionality is a term coined by Richard Bellman to describe the problem caused by the exponential increase in volume associated with adding extra dimensions to a space.

- Bellman, R.E. 1957. Dynamic Programming. Princeton University Press, Princeton, NJ.

N - Dimensions



The graph of n-ball volume as a function of dimension was plotted more than 117 years ago by Paul Renno Heyl.

The upper curve gives the ball's surface area, for which Heyl used the term "boundary."

The volume graph is the lower curve, labeled "content."
The illustration is from Heyl's 1897 thesis, "Properties of the locus r = constant in space of n dimensions."

Brian Hayes: An Adventure in the *N*th Dimension. *American Scientist*, Vol. 99, No. 6, November-December 2011, pages 442-446.

N - Dimensions

n	$V(n,1)$
1	2
2	$\pi \approx 3.1416$
3	$\frac{4}{3}\pi \approx 4.1888$
4	$\frac{1}{2}\pi^2 \approx 4.9348$
5	$\frac{8}{15}\pi^2 \approx 5.2638$
6	$\frac{1}{6}\pi^3 \approx 5.1677$
7	$\frac{16}{105}\pi^3 \approx 4.7248$
8	$\frac{1}{24}\pi^4 \approx 4.0587$
9	$\frac{32}{945}\pi^4 \approx 3.2985$
10	$\frac{1}{120}\pi^5 \approx 2.5502$

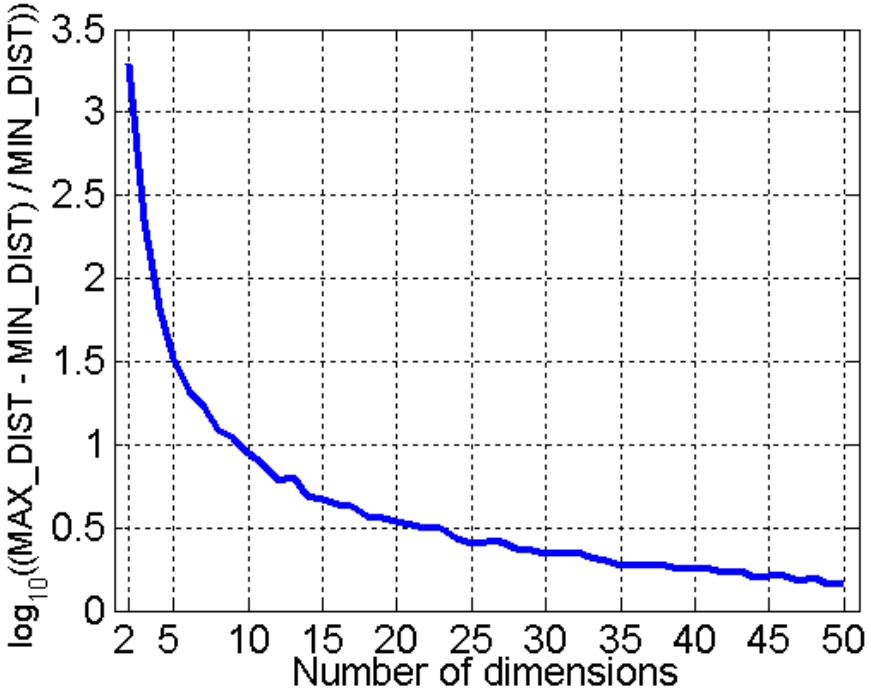
Beyond the fifth dimension, the volume of unit n-ball decreases as n increases. When we compute few larger values of n, finding that V(20,1) = 0,0258 and V(100,1) = 10^{-40}

Brian Hayes: An Adventure in the *N*th Dimension. *American Scientist*, Vol. 99, No. 6, November-December 2011, pages 442-446.

Curse of Dimensionality

When dimensionality increases difference between max and min distance between any pair of points is being uniform

Randomly generate 500 points

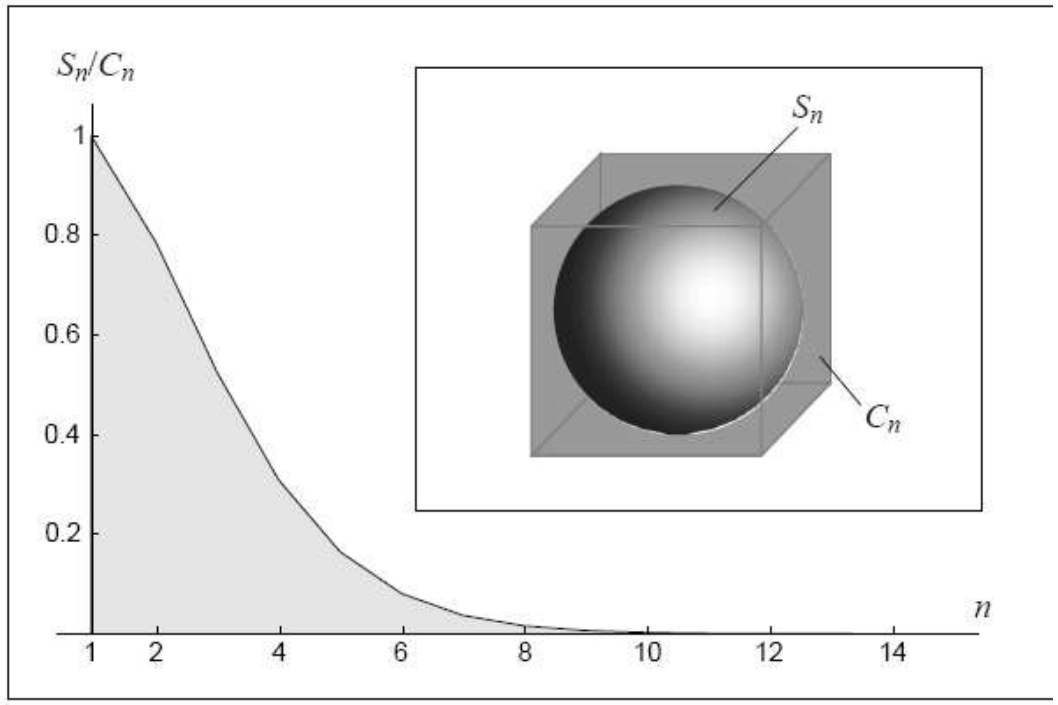


Fionn Murtagh, Hilbert Space Becomes Ultrametric in the High Dimensional Limit: Application to Very High Frequency Data Analysis

Curse of Dimensionality

The volume of an n-dimensional sphere with radius r is

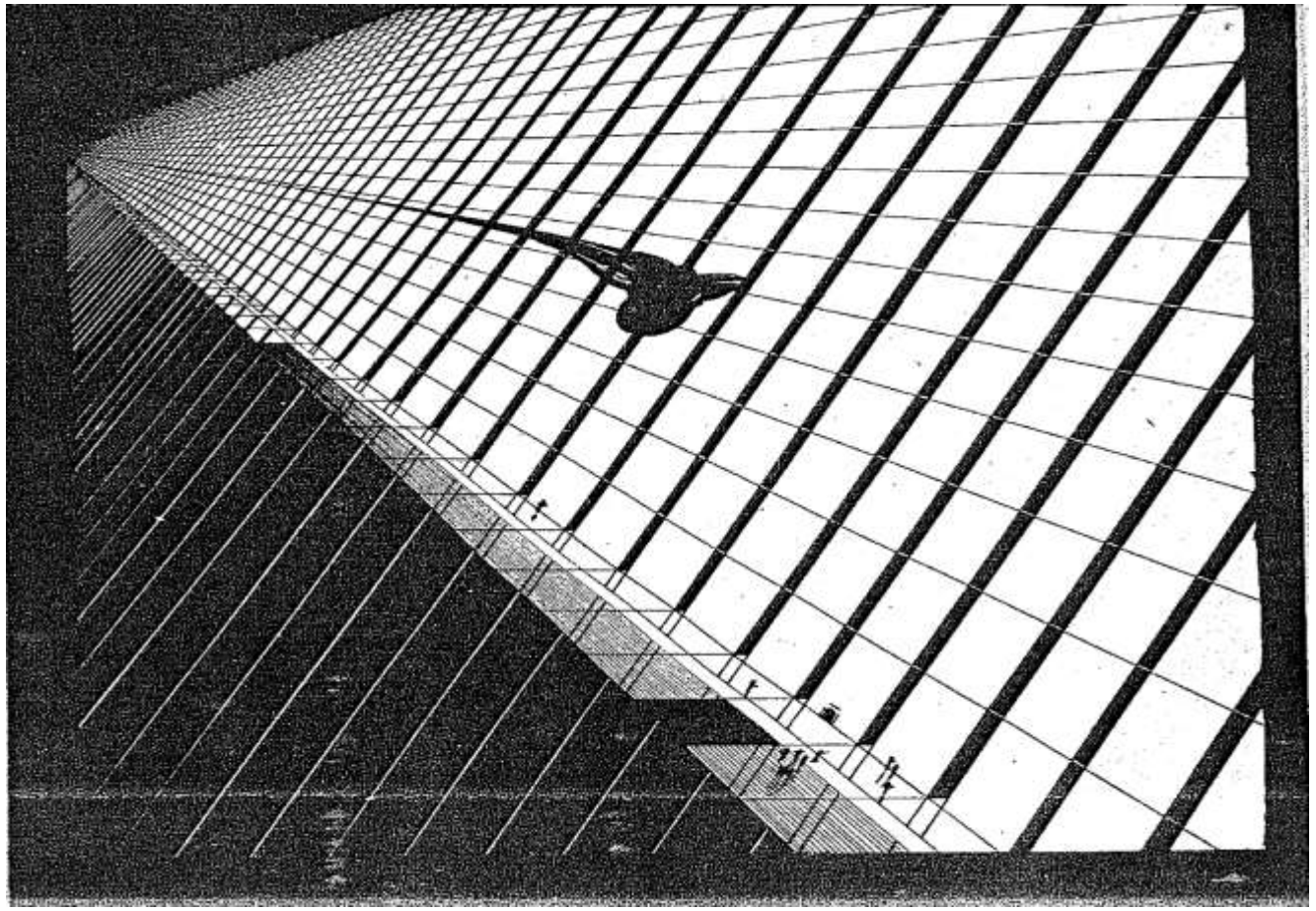
$$V_n(r) = \frac{\pi^{\frac{n}{2}} r^n}{\Gamma\left(\frac{n}{2} + 1\right)}$$



dimension

Ratio of the volumes of unit sphere and embedding hypercube of side length 2 up to the dimension 14.

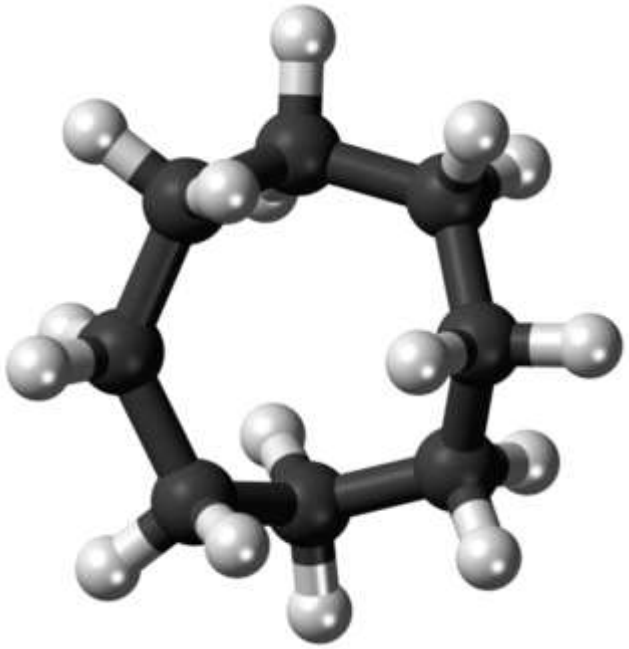
High dimensional space



Cyclooctane

Cyclooctane is molecule with formula C_8H_{16}
To understand molecular motion we need characterize the molecule‘s possible shapes.

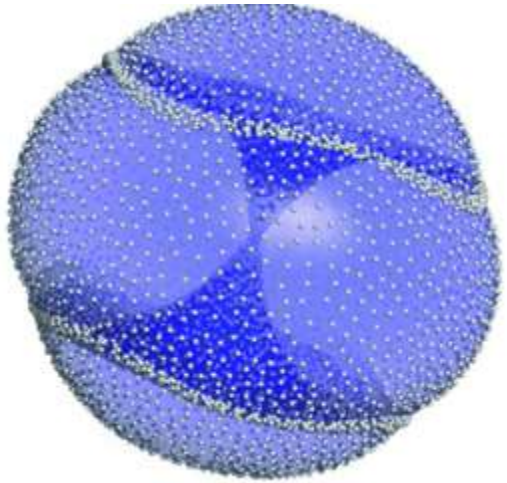
Cyclooctane has 24 atoms and it can be viewed as point in 72 dimensional spaces.



A. Zomorodian. Advanced in Applied and Computational Topology, Proceedings of Symposia in Applied Mathematics, vol. 70, AMS, 2012

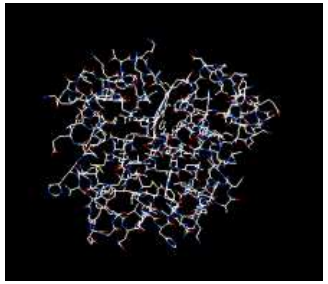
Cyclooctane‘s space

- The conformation space of cyclooctane is a two-dimensional surface with self intersection.

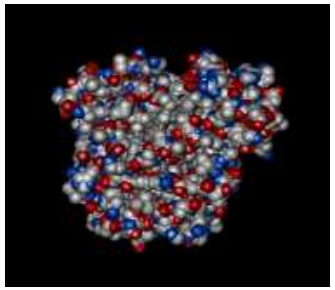


W. M. Brown, S. Martin, S. N. Pollock, E. A. Coutsias, and J.-P. Watson. Algorithmic dimensionality reduction for molecular structure analysis. Journal of Chemical Physics, 129(6):064118, 2008.

How to define protein dynamics



Protein is a macromolecule



protein states

Protein states are defined by means of conformations of a protein macromolecule.

A conformation is understood as the spatial arrangement of all "elementary parts" of a macromolecule.

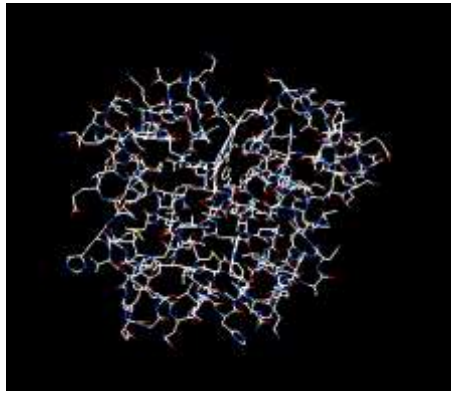
Atoms, units of a polymer chain, or even larger molecular fragments of a chain can be considered as its "elementary parts". Particular representation depends on the question under the study.

protein dynamics

Protein dynamics is defined by means of conformational rearrangements of a protein macromolecule.

Conformational rearrangements involve fluctuation induced movements of atoms, atomic groups, and even large macromolecular fragments.

Protein dynamics



To study protein motions on the subtle scales, say, from ~10^{-9} sec, it is necessary to use the atomic representation of a protein molecule.

Protein molecule consists of ~10^{3} atoms.

Protein conformational states:

number of degrees of freedom : ~ 10^3

dimensionality of (Euclidian) space of states : ~ 10^3

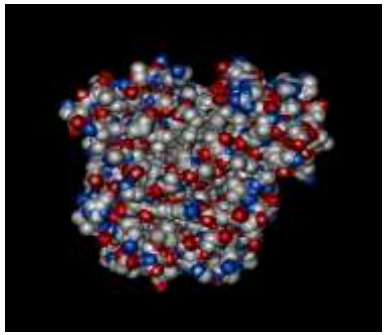
In fine-scale presentation, dimensionality of a space of protein states is very high.

Protein dynamics

Protein dynamics over high dimensional conformational space is governed by complex energy landscape.

protein energy landscape

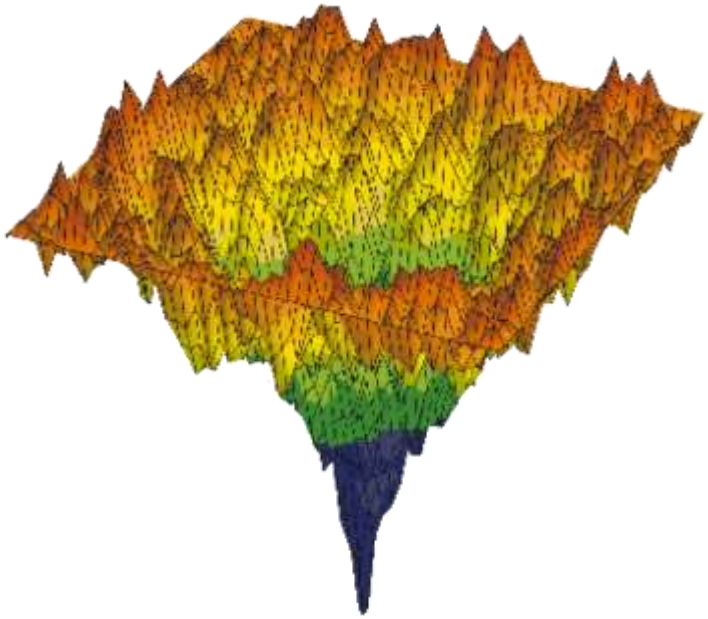
Given the interatomic interactions, one can specify the potential energy of each protein conformation, and thereby define an energy surface over the space of protein conformational states. Such a surface is called the protein energy landscape.



As far as the protein polymeric chain is folded into a condensed globular state, high dimensionality and ruggedness are assumed to be characteristic to the protein energy landscapes

Protein energy landscape: dimensionality: ~ 10^3;
number of local minima ~10^{100}

Protein dynamics



While modeling the protein motions on many time scales (from ~10^{-9} sec up to ~10^{0} sec), we need the simplified description of protein energy landscape that keeps its multi-scale complexity.

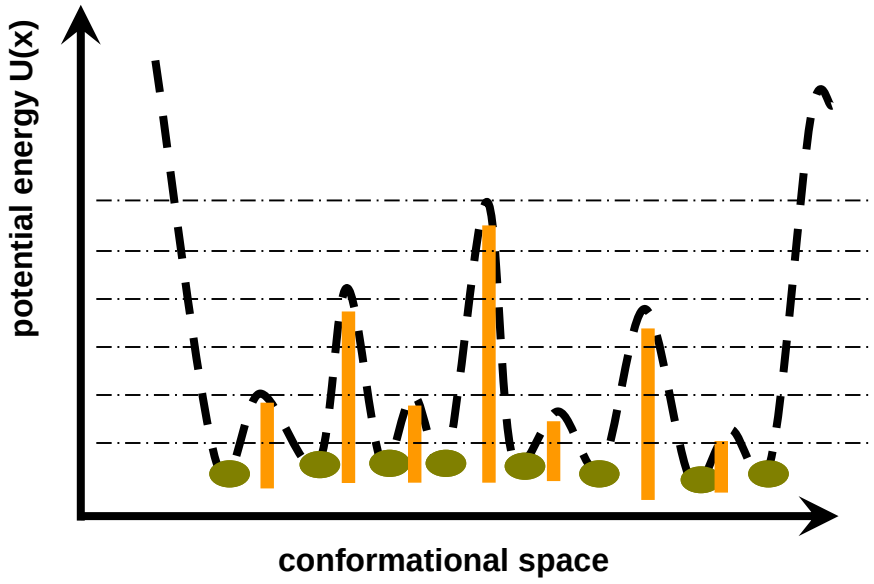
How such model can be constructed?

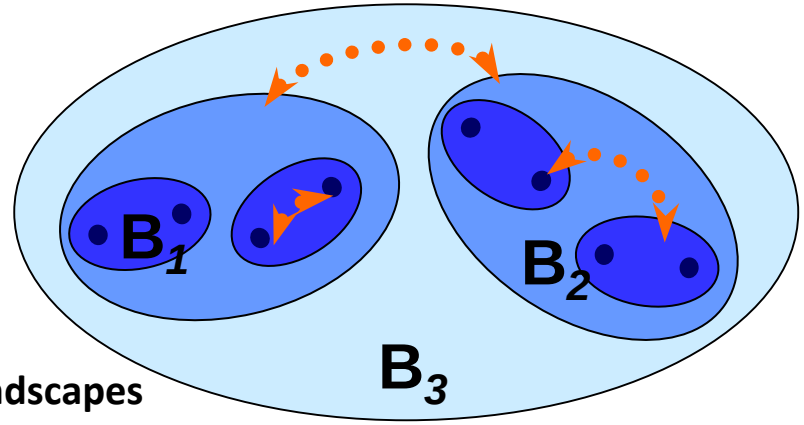
Computer reconstructions of energy landscapes of complex molecular structures suggest some ideas.

Protein dynamics

Method

1. **Computation of local energy minima and saddle points on the energy landscape using molecular dynamic simulation;**
2. **Specification a topography of the landscape by the energy sections;**
3. **Clustering the local minima into hierarchically nested basins of minima.**
4. **Specification of activation barriers between the basins.**



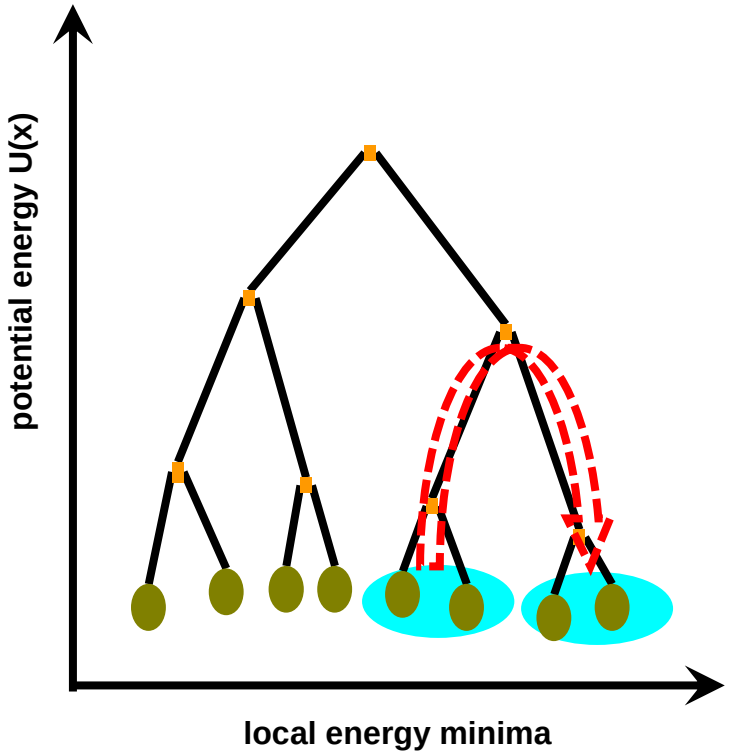


O.M.Becker, M.Karplus, Computer reconstruction of complex energy landscapes *J.Chem.Phys*. 106, 1495 (1997)

Protein dynamics

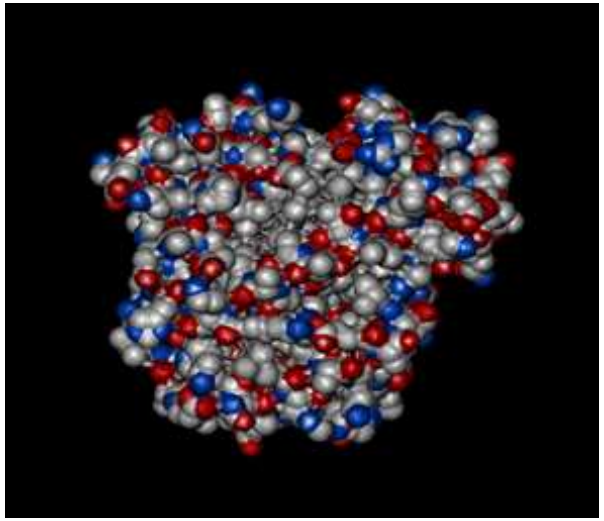
The relations between the basins embedded one into another are presented by a tree-like graph.

Such a tee is interpreted as a "skeleton" of complex energy landscape. The nodes on the border of the tree (the "leaves") are associated with local energy minima (quasi-steady conformational states). The branching vertexes are associated with the energy barriers between the basins of local minima.



O.M.Becker, M.Karplus, Presentation of energy landscapes by tree-like graphs
***J.Chem.Phys*. 106, 1495 (1997)**

Complex energy landscapes : a protein

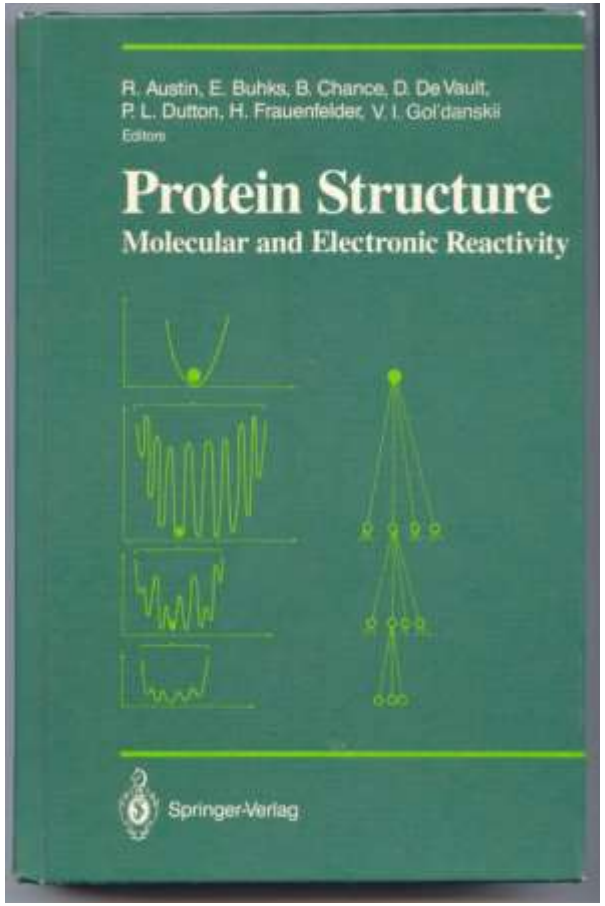


The total number of minima on the protein energy landscape is expected to be of the order of ~10^{100}.

This value exceeds any real scale in the Universe. Complete reconstruction of protein energy landscape is impossible for any computational resources.

Protein Structure

25 years ago, Hans Frauenfelder suggested a tree-like structure of the energy landscape of myoglobin



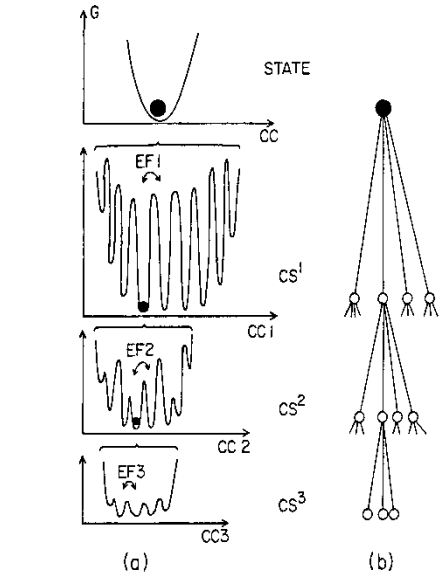


Figure 5. Hierachical arrangement of the conformational substates in myoglobin. (a) Schematized energy surfaces. (b) Tree diagram. G is the Gibbs energy of the protein, CC(1-4) are conformational coordinates. After [31].

Hans Frauenfelder, in *Protein Structure* (N-Y.:Springer Verlag, 1987) p.258.

10 years later, Martin Karplus suggested the same idea

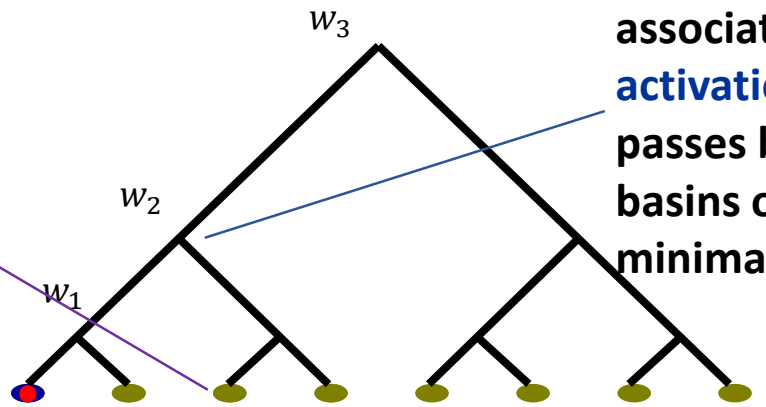
"In <…> proteins, for example, where individual states are usually clustered in "basins", the interesting kinetics involves basin-to-basin transitions. The internal distribution within a basin is expected to approach equilibrium on a relatively short time scale, while the slower basin-to-basin kinetics, which involves the crossing of higher barriers, governs the intermediate and long time behavior of the system."

Becker O. M., Karplus M. *J. Chem. Phys.*, 1997, 106, 1495

This is exactly the physical meaning of protein ultrameticity !

Random walk on the boundary of a Cayley tree

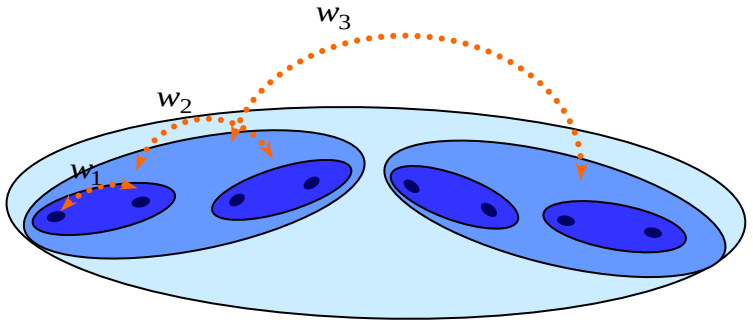
Cayley tree is understood as a hierarchical skeleton of protein energy landscape. The leaves are the local energy minima, and each subtree of the Cayley tree is a basin of local minima.



The branching vertexes are associated with the activation barriers for passes between the basins of local minima.

Master equation

$$\frac{df_i(t)}{dt} = \sum_{j \neq i}^{n} w_{ji} f_j(t) - \sum_{j \neq i}^{n} w_{ij} f_i(t)$$



f_i is the transition probability, i.e. the probability to find a walker in a state i at instant t, and w_{ji} is the rate of transition from j to i. The energy landscape is represented by the transition rates w_{ji}

p-Adic description of ultrametric random walk

The basic idea:

In the basin-to-basin approximation, the distances between the protein states are ultrametric, so they can be specified by the *p*-adic numerical norm, and transition rates can be indexed by the p-adic numbers.

Conclusion