

# КОНФОРМАЦИИ ШАРНИРНОЙ ЦЕПИ КАК МОДЕЛЬ СВОРАЧИВАНИЯ БЕЛКА

Вилкул Е.А., Иванов А.О., Тужилин А.А.

*Московский государственный университет им. М.В. Ломоносова, Москва*

elena.tuzhilina@mail.ru, aoiva@mech.math.msu.su, tuz@mech.math.msu.su

Поступила 11.08.2015

В настоящей работе мы изучаем геометрию простейшей модели, которая могла бы быть применима к исследованию конформации белков. А именно, мы рассматриваем пространственную ломаную (которую можно представлять, например, как углеродно-азотный остов белка), фиксируем в ней длины ребер и углы между смежными ребрами, задаем потенциальное взаимодействие между ее вершинами и изучаем конфигурации такой ломаной, соответствующие локальным минимумам суммарного потенциала. При некоторых допущениях на характер взаимодействия эта задача полностью решается. При менее жестких ограничениях удастся лишь проделать компьютерный эксперимент. Тем не менее, уже в рассмотренных случаях мы обнаружили ряд качественных явлений. Например, меняя потенциал так, чтобы его вид всегда оставался потенциальной ямой, мы наблюдали, как дискретные локальные минимумы включались в большие плоские области с очень маленькой вариацией значений: из-за наличия постоянного спонтанного движения атомов такие области также нужно относить к решениям. Кроме того, наши теоретические результаты и численные эксперименты позволили нам сформулировать гипотезу, оценивающую максимальное число строгих локальных минимумов суммарного потенциала, которое уже для 30 аминокислот оказалось астрономическим. Мы думаем, что такие оценки были бы полезны для осознания реальной ситуации в современной науке, изучающей конформации белков методами ядерно-магнитного резонанса и рентгено-структурного анализа.

УДК 514.8, 51-76, 57.087

## 1 Введение

Изучение конформации белков является одной из наиболее важных задач молекулярной биологии, так как физические, химические и биологические свойства белка в значительной мере зависят от пространственного расположения составляющих его

аминокислот. Например, одно из наиболее важных свойств белков — способность ускорять или замедлять химические реакции — реализуется обычно лишь некоторой конфигурацией близко лежащих в пространстве атомов белка, которые могут быть расположены далеко друг от друга вдоль аминокислотной цепи. Отметим, что сбои в процессе производства белка иногда приводят к “фатальному” изменению конформации, разрушающему те или иные свойства этого белка и, как следствие, вызывают серьезные заболевания. Последнее является одной из многочисленных мотиваций активного развития теории пространственных структур белков и поиска эффективных алгоритмов определения геометрической конфигурации (конформации) белка по последовательности его аминокислот.

В предыдущих наших работах (см. например [1]) мы пытались моделировать конформации белков с помощью функций, однозначно определяющих форму пространственных ломаных, вершины которых — альфа-углероды рассматриваемого белка. Мы собирались выявить различные геометрические закономерности, изучая белки с известными пространственными структурами, выложенными в знаменитой базе Protein Data Bank (PDB). Желая иметь “реальное” представление о том, сколь сильно могут отклоняться от среднего различные базисные характеристики, такие, как длины ковалентных связей, углы между смежными ковалентными связями и др., мы, используя информацию из PDB, проделали численный эксперимент и обнаружили, к нашему удивлению, многочисленные очевидные ошибки в имеющихся pdb-файлах [2]–[4]. Познакомившись ближе с тем, как были получены данные в PDB, мы узнали, что выложенные там координаты атомов уже сами являются результатом компьютерного моделирования. Все это мотивировало нас отложить изучение статистических характеристик, основанных на информации из PDB, и заняться теоретическими аспектами (общие сведения о молекулярной динамике могут быть найдены, например, в [5]).

В настоящей статье мы рассматриваем пространственные ломаные, у которых фиксированы длины звеньев и углы между соседними звеньями. Примером интерпретации таких ломаных может служить геометрическая модель углеродно-азотного остова белка. В качестве параметров, задающих ту или иную пространственную конфигурацию ломаной, мы берем ориентированные углы между звеньями, расположенными через одно. Мы предполагаем, что между вершинами ломаной имеются потенциальные взаимодействия, а искомые конфигурации минимизируют суммарный потенциал — сумму локальных потенциалов. Так как мы фиксировали длины звеньев и углы между соседними звеньями, то взаимодействия соседних вершин и вершин, расположенных через одну, можно не учитывать. Чтобы сделать задачу обозримой, мы разбили ее на ряд подзадач, каждая из которых ограничивает взаимодействие, устанавливая его только для пар вершин, между которыми не более чем  $k$  звеньев (идею такого разбиения нам подсказал профессор К.В. Шайтан).

Первоначально в качестве потенциалов мы брали произвольные функции, имеющие форму “потенциальной ямы”. Для  $k = 3$  и потенциалов, вообще говоря зависящих от пары вершин, задача полностью решается: всегда имеется конечное число решений, причем каждая пара взаимодействующих вершин дает вклад 2 или 1 в зависимости от того, является или нет расстояние, соответствующее минимуму потенциала, внутренней точкой отрезка возможных расстояний между этими вершинами. Если вклад равен 2, то локальные оптимальные конфигурации симметричны, так что оба значения локального потенциала одинаковы. Таким образом, при  $k = 3$  величины

всех локальных минимумом суммарного потенциала одинаковы, а максимальное количество локальных минимумов равно  $2^{n-3}$ , где  $n$  — число вершин.

Затем мы рассмотрели случай  $k = 4$ . Здесь все намного сложнее, даже для “коротких” ломаных и одной и той же функции локального потенциала. Оказалось, что на сей раз ответ существенно зависит от конкретной функции, задающей потенциал. Более того, уже здесь видно, что при некоторых потенциалах “строгими” решениями обойтись не удастся. Дело в том, что у результирующего потенциала могут возникать площадки, содержащие локальные минимумы и имеющие очень маленькие вариации значений. Если учесть, что в реальности атомы молекул всегда находятся в движении, то такие площадки стоит относить к искомым решениям. Именно такой феномен наблюдается для потенциала Леннард–Джонса, используемого при моделировании взаимодействия сферических неполярных молекул. Для белков, скорее всего, потенциал такого типа не подходит, так как считается, что белки сворачиваются вполне определенным образом.

Для  $k = 4$  мы ограничили число вершин ломаной. Если вершин 5, то имеется 2 параметра, и несложно визуализировать график суммарного потенциала. Если же вершин 6, то параметров 3, а график потенциала нужно рисовать в четырехмерном пространстве. Поэтому здесь мы, вместо графика, изображали представителей семейства эквипотенциальных поверхностей. Как уже отмечалось, для таких ломаных и  $k = 3$  суммарный потенциал имеет максимум 8 локальных минимумов. При переходе к  $k = 4$  для некоторых значений потенциала мы также обнаружили 8 локальных минимумов, однако теперь эти минимумы могут иметь разные значения. Больше 8 строгих минимумов нами не наблюдалось, даже в рассмотренном также случае  $k = 5$  (снова для 6 атомов).

Таким образом, наше исследование подняло ряд вопросов: верно ли, что для шести вершин всегда имеется не более 8 строгих минимумов? А что будет в общем случае: верно ли, что для ломаной с  $n$  вершинами максимальное число строгих локальных минимумов всегда (при любом  $k$  и любых локальных потенциалах) не больше  $2^{n-3}$ ? Или же возможность самопересечений у ломаной при больших  $n$  может нарушить эту закономерность? Какие потенциалы были бы наиболее подходящими для моделирования конформации белков? Заметим, что ответ на второй вопрос мог бы продемонстрировать степень надежности современных методов определения трехмерных структур белков: если ответ положительный, то уже для белка из 30 аминокислот имеем  $n \approx 90$ , поэтому количество локальных минимумов могло бы достигать  $10^{27}$ , а вероятность того, что компьютерная минимизация суммарного потенциала приводит к “правильному” решению, была бы ничтожно мала.

Работа выполнена при поддержке РФФ, соглашение № 14-50-00029.

## 2 Постановка общей задачи

Рассмотрим шарнирный механизм  $L$ , представляющий собой ломаную с фиксированными углами между последовательными звеньями. Таким образом, в изучаемом механизме допустимы вращения вокруг каждого неконцевого звена  $e$  каждой из двух ломаных, составляющих  $L \setminus e$ .

Введем необходимые обозначения, см. рис. 1. Пусть  $A_0, A_1, \dots, A_n$  — последовательные вершины ломаной  $L$ , а  $e_i = A_{i-1}A_i$  — ее  $i$ -ое вектор-звено. Длину звена  $e_i$

обозначим через  $a_i$ , а величину угла  $A_{i-1}A_iA_{i+1}$  — через  $\psi_i$ . В дальнейшем мы будем рассматривать всевозможные такие ломаные  $L$ , предполагая, что у них фиксированы  $a_i$  и  $\psi_i$ . Кроме того, будем считать, что  $0 < \psi_i < \pi$  при каждом  $i$ . Последнее приводит к тому, что векторы  $e_i$  и  $e_{i+1}$  линейно независимы.

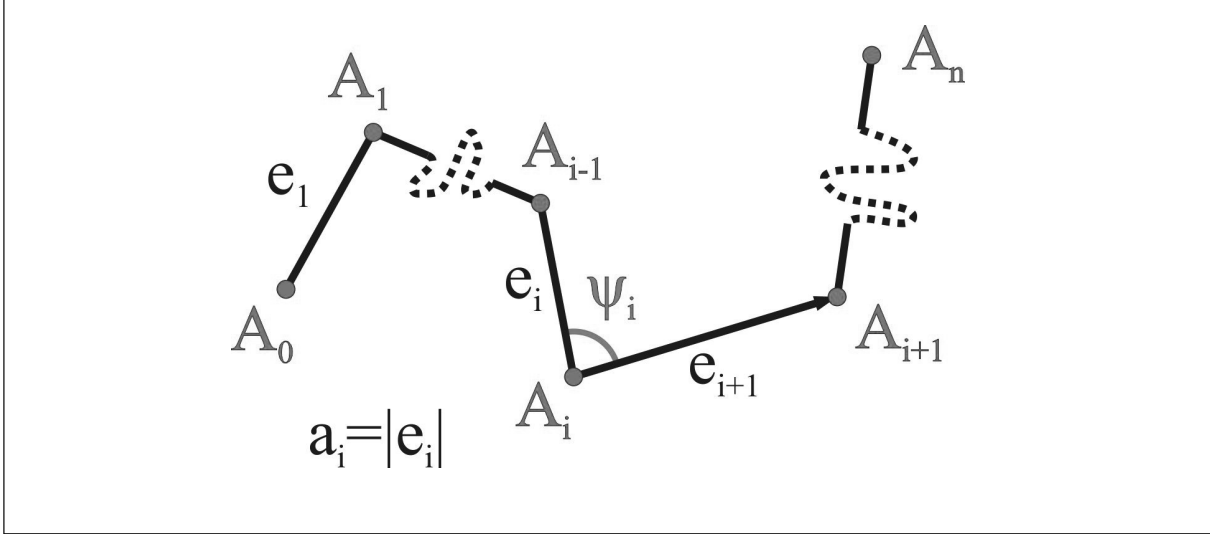


Рис. 1: Рассматриваемые шарнирные механизмы.

Зададим нормали  $N_i = A_iA_{i+1} \times A_iA_{i-1} = e_i \times e_{i+1}$  (к плоскостям, проходящим через треугольники  $A_{i-1}A_iA_{i+1}$ ), тогда  $e_i, e_{i+1}, N_i$  — положительно ориентированный базис в  $\mathbb{R}^3$  (по отношению к стандартной ориентации). Чтобы охарактеризовать положение звена  $e_{i+1}$  относительно звена  $e_{i-1}$ , введем угол  $\beta_i$  следующим образом. Положению, в котором нормали  $N_i$  и  $N_{i-1}$  сонаправлены, соответствует  $\beta_i = 0$ . Далее, плоскость вращения вектора  $N_i$  перпендикулярна  $e_i$ . Вектор  $N_{i-1}$  тоже перпендикулярен  $e_i$ . Чтобы задать положительное направление вращения, дополним вектор  $N_{i-1}$  до ориентированного базиса этой плоскости вектором  $M_{i-1} = N_{i-1} \times e_i$ . Если  $m_{i-1}$  и  $n_{i-1}$  обозначают нормированные векторы  $M_{i-1}$  и  $N_{i-1}$ , то направление вектора  $N_i$ , повернутого на угол  $\beta_i$ , равно  $n_{i-1} \cos \beta_i + m_{i-1} \sin \beta_i$ .

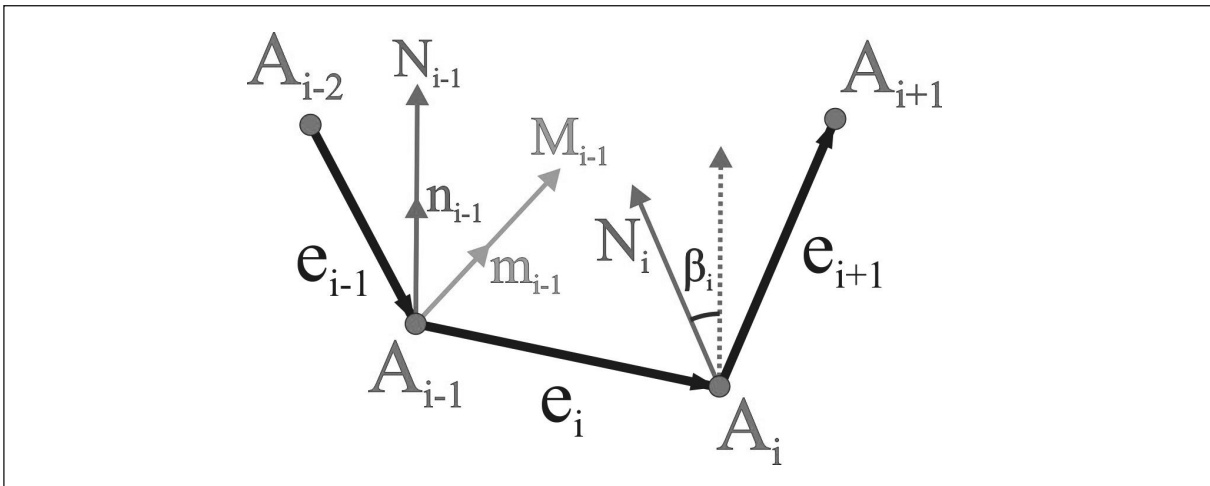


Рис. 2: Углы поворота звеньев шарнирного механизма.

Будем предполагать, что вершины  $A_i$  ломаной  $L$  представляют собой частицы, которые взаимодействуют с потенциалом  $U(x)$ , где  $x$  — квадрат расстояния между частицами. Потенциал  $\mathcal{U}(L)$  всей ломаной  $L$  будем считать равным сумме потенциалов  $U_{ij}(L)$  всех пар частиц — пар  $(A_i, A_j)$ ,  $i < j$ , вершин ломаной  $L$ . Таким образом,  $\mathcal{U}(L) = \sum_{0 \leq i < j \leq n} U_{ij}(L)$ . Наша задача: описать конфигурации, в которых потенциал  $\mathcal{U}(L)$  принимает наименьшее значение. Такие конфигурации будем называть *минимальными*. В более общей постановке, будем искать конфигурации, в которых  $\mathcal{U}(L)$  имеет локальный минимум (любые малые изменения ломаной  $L$  не уменьшают ее потенциал). Такие конфигурации будем называть *локально минимальными*.

Функцию  $U(x)$  выберем в форме “потенциальной ямы”. А именно, будем предполагать, что для некоторого  $x_0$  при  $0 < x \leq x_0$  функция  $U(x)$  строго монотонно убывает, а при  $x_0 \leq x < +\infty$  функция  $U(x)$  строго монотонно возрастает; при этом  $U(x_0) < 0$ , при  $x \rightarrow 0$  имеем  $U(x) \rightarrow +\infty$ , а при  $x \rightarrow +\infty$  выполняется  $U(x) \rightarrow 0$ , см. рис. 3.

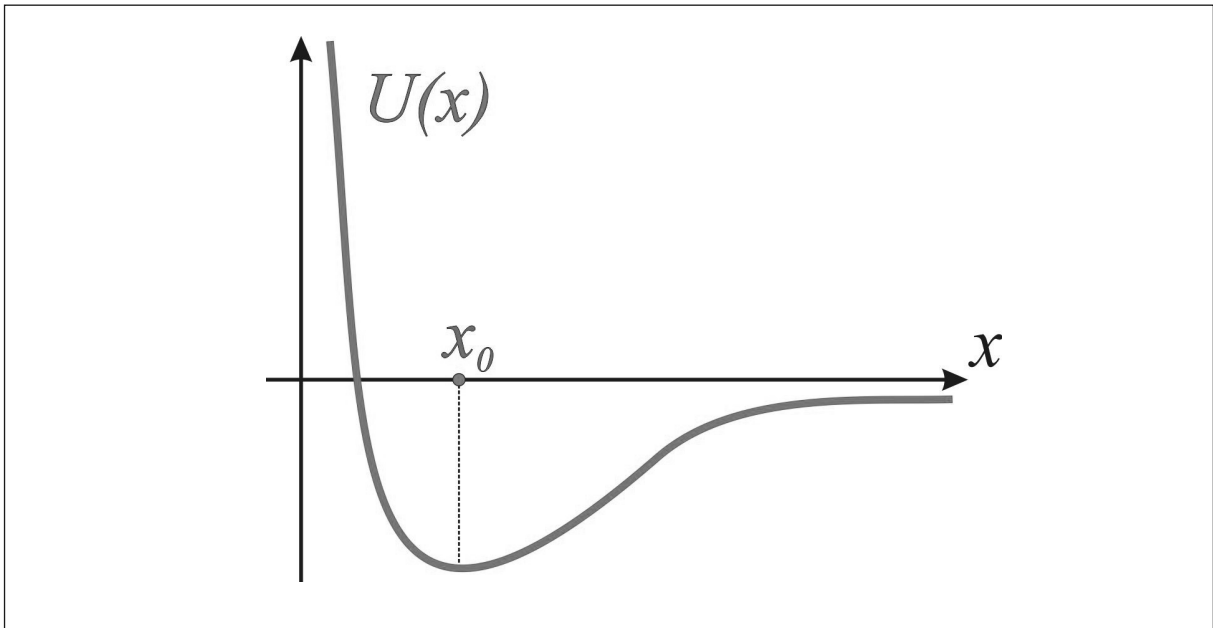


Рис. 3: Потенциал взаимодействия частиц.

Так как фиксированы длины звеньев  $e_i$  и углы между последовательными звеньями, то вклад в суммарный потенциал каждой пары последовательных вершин, в также вершин, взятых через одну, не меняется при допустимой деформации ломаной  $L$ . Поэтому в суммарном потенциале  $\mathcal{U}(L)$  эти пары можно не учитывать, что мы и будем делать. Таким образом, теперь

$$\mathcal{U}(L) = \sum_{\substack{0 \leq i \leq n-3 \\ i+3 \leq j \leq n}} U_{i(i+3)}(L).$$

Далее, чтобы сделать задачу обозримой, предположим для начала, что в суммарный потенциал дают вклады лишь те пары вершин, которые расположены не далее, чем на расстоянии  $k$  вдоль ломаной  $L$ . Таким образом, мы полагаем  $U_{ij}(L) = 0$  при  $j > i + k$ . Например, если  $k = 3$ , то взаимодействуют  $A_0$  и  $A_3$  с потенциалом  $U_{03}(L)$ ,

$A_1$  и  $A_4$  с потенциалом  $U_{14}(L)$ , и т.д. Если  $k = 4$ , то, помимо предыдущих взаимодействий, добавляются еще  $A_0 - A_4$  с потенциалом  $U_{04}(L)$ ,  $A_1 - A_5$  с потенциалом  $U_{15}(L)$  и т.д.

Обозначим через  $L_{ij}$  часть ломаной  $L$  между вершинами  $A_i$  и  $A_j$ .

### 3 Случай $k = 3$

Заметим, что расстояния  $|A_i A_{i+3}|$  могут быть выбраны независимо друг от друга, причем для минимального или максимального  $|A_i A_{i+3}|$  возможна ровно одна конфигурация  $L_{i(i+3)}$  (являющаяся плоской), а для остальных значений  $|A_i A_{i+3}|$  — две, приводящие к двум зеркально симметричным треугольным пирамидам  $L_{i(i+3)}$ . Таким образом, минимальное значение потенциала  $\mathcal{U}(L)$  соответствует минимальным значениям  $U_{i(i+3)}(L)$  всех подломанных  $L_{i(i+3)}$ .

Итак, рассмотрим некоторую подломаную  $L_{i(i+3)}$ , например  $L_{03}$ . Тогда положение между  $e_1$  и  $e_3$  задается углом  $\beta_2$ . Непосредственные вычисления показывают, что

$$|A_0 A_3|^2 = a_1^2 + a_2^2 + a_3^2 - 2a_1 a_2 \cos(\psi_1) - 2a_2 a_3 \cos(\psi_2) + 2a_1 a_3 \cos(\psi_1) \cos(\psi_2) - 2a_1 a_3 \sin(\psi_1) \sin(\psi_2) \sin(\beta_2),$$

поэтому при каждом значении параметров  $a_i > 0$  и  $0 < \psi_j < \pi$  величина  $|A_0 A_3|^2$  строго монотонно убывает при  $-\pi/2 \leq \beta_2 \leq \pi/2$ , и строго монотонно возрастает при  $\pi/2 \leq \beta_2 \leq 3\pi/2$ . При этом в точках  $\beta_2$  и  $\pi - \beta_2$  величина  $|A_0 A_3|^2$  принимает одинаковые значения.

Пусть  $M$  и  $m$  обозначают соответственно максимальное и минимальное значение величины  $|A_i A_{i+3}|^2$ . Если  $m < x_0 < M$ , то минимальное значение  $U_{i(i+3)}(L)$  достигается на двух зеркально симметричных конфигурациях  $L_{i(i+3)}^0$  и  $L_{i(i+3)}^1$  (при некоторых  $\beta_{i+2}^0$  и  $\pi - \beta_{i+2}^0$ ). Конфигурации  $L_{i(i+3)}^k$  назовем *невыврожденными*. Если же  $x_0 \leq m$  или  $x_0 \geq M$ , то минимум и максимум потенциала  $U_{i(i+3)}(L)$  достигаются при  $\beta_{i+2}$ , равном  $\pi/2$  и  $-\pi/2$  соответственно (на плоской конфигурации). Для удобства, будем считать, что и здесь минимум потенциала достигается на двух совпадающих конфигурациях  $L_{i(i+3)}^0 = L_{i(i+3)}^1$ . В этом случае конфигурации  $L_{i(i+3)}^k$  назовем *вырожденными*.

В силу независимости расстояний  $|A_i A_{i+3}|$  друг от друга, каждая минимальная конфигурация ломаной  $L$  “склеивается” из минимальных конфигураций  $L_{i(i+3)}^k$ ,  $k = 0, 1$  и, поэтому, однозначно (с точностью до движения) задается последовательностью  $(k_2, k_3, \dots, k_{n-2})$  из нулей и единиц, соответствующей выбору минимальных конфигураций  $L_{i(i+3)}^{k_{i+2}}$ , составляющих минимальную  $L$ .

**Замечание 3.1.** Ясно, что если все  $a_i$  равны между собой, все  $\psi_i$  равны между собой, и последовательность  $k_i$  состоит из чередующихся 0 и 1, то минимальная конфигурация  $L$  — спираль. В общем случае, минимальные конфигурации  $L$ , соответствующие таким последовательностям  $k_i$ , будем называть *обобщенными спиральями*.

**Замечание 3.2.** Если некоторое  $k_{i+2}$  равно  $k_{i+3}$ , то треугольные пирамиды, построенные на вершинах ломаных  $L_{i(i+3)}$  и  $L_{(i+1)(i+4)}$ , лежат в одном полупространстве относительно плоскости, проходящей через треугольник  $A_{i+1} A_{i+2} A_{i+3}$ , и такое крепление фрагментов  $L_{i(i+3)}$  и  $L_{(i+1)(i+4)}$  естественно назвать *цис-креплением*; если же  $k_{i+2} \neq k_{i+3}$ , то соответствующее крепление естественно назвать *транс-креплением*.

Если запретить цис-крепления, то минимальные конфигурации для ломаной  $L$  будут всегда обобщенными спиралями. Таким образом, все “изломы” обобщенной спирали в рассматриваемом случае возникают благодаря цис-креплениям.

**Замечание 3.3.** Все приведенные выше выводы не зависят от выбора потенциала  $U(x)$  и параметров  $a_i$  и  $\psi_j$  ломаной  $L$ .

## 4 Случай $k = 4$

В этом случае уже никакой независимости нет. В качестве начального шага мы рассмотрим ломаную  $L$ , состоящую из пяти вершин  $A_0, \dots, A_4$ , причем для простоты будем считать, что все  $a_i$  равны между собой и равны  $a$ , и все  $\psi_i$  также равны между собой и равны  $\psi$ . Так как изменения масштаба для нас не важны, положим  $a = 1$ .

Кроме этого, мы изменим параметризацию вращений вокруг ребер: мы будем считать, что два средних звена ломаной  $L$  фиксированы, а углы  $\beta_i$  описывают вращения первого и последнего звеньев. Таким образом, репер  $e_2, e_3, N_2$  неподвижен, угол  $\beta_1$  описывает вращение первого звена ломаной  $L$ , а угол  $\beta_3$  — последнего. Дадим формальное определение этих углов.

Для этого мы переопределим  $M_i$ , положив  $M_2 = N_2 \times e_2$ ,  $M_3 = N_2 \times e_3$ , и если  $m_i$  — направление так определенного  $M_i$ , то  $n_1 = n_2 \cos(\beta_1) + m_2 \sin(\beta_1)$ , а  $n_3 = n_2 \cos(\beta_3) + m_3 \sin(\beta_3)$ .

Вычислим теперь  $|A_0A_3|^2$ ,  $|A_1A_4|^2$  и  $|A_0A_4|^2$  через параметры ломаной и углы  $\beta_1$  и  $\beta_3$ . Приведем результаты вычисления для  $\psi = 2\pi/3$ :

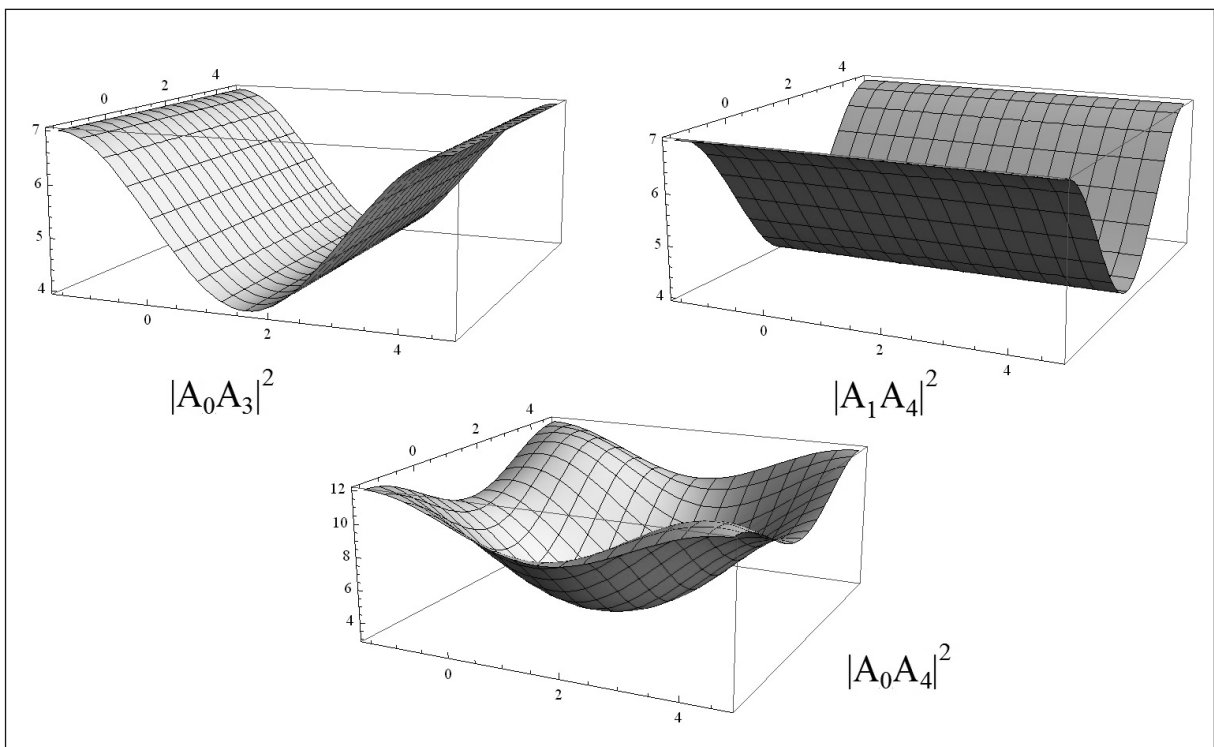


Рис. 4: Графики квадратов расстояний.

$$|A_0 A_3|^2 = \frac{11}{2} - \frac{3}{2} \sin(\beta_1),$$

$$|A_1 A_4|^2 = \frac{11}{2} - \frac{3}{2} \sin(\beta_3),$$

$$|A_0 A_4|^2 = \frac{33}{4} - \frac{9}{4} \sin(\beta_1) - \frac{9}{4} \sin(\beta_3) - \frac{3}{2} \cos(\beta_1) \cos(\beta_3) - \frac{3}{4} \sin(\beta_1) \sin(\beta_3).$$

На рис. 4 показаны графики этих величин.

Заметим, что первые две функции имеют целый отрезок минимумов, а третья — ровно один минимум, причем при движении от точки минимума к границе области определения углов  $\beta_i$  все три функции монотонно возрастают.

Заметим также, что графики первых двух величин пересекаются по четырем кривым, встречающимся в точке общего минимума, рис. 5.

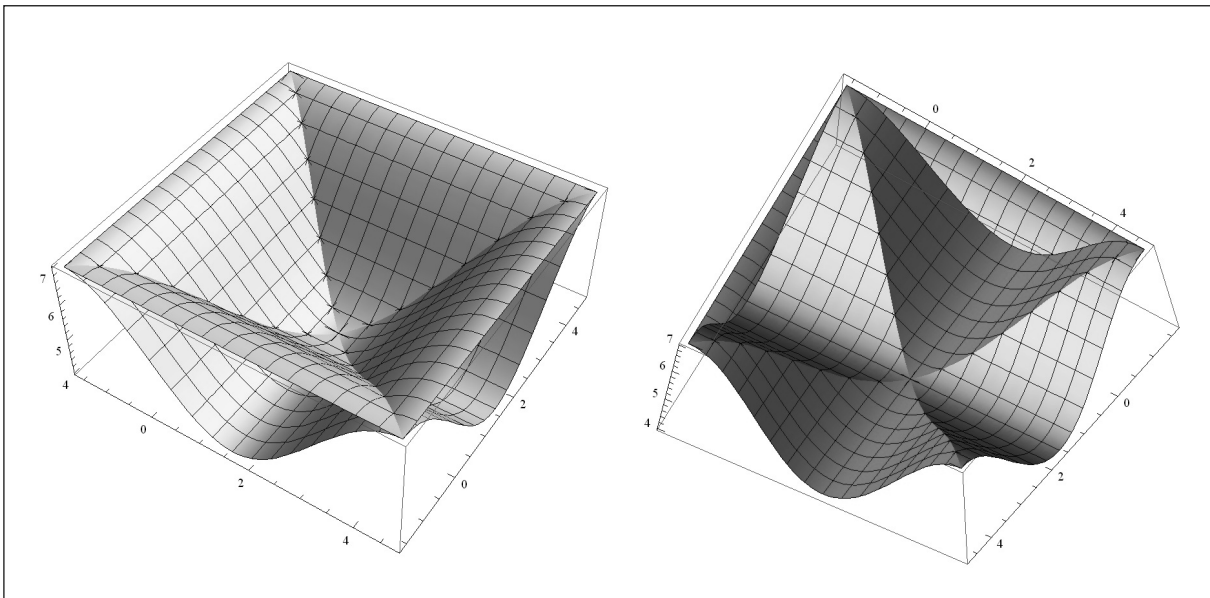


Рис. 5: Графики квадратов двух расстояний, изображенные одновременно.

Изобразим теперь все три графика одновременно, см. рис. 6. На рисунке отчетливо видно, что имеются четыре точки пересечения всех трех графиков. Эти четыре точки разбиваются на две пары симметричных относительно вертикальной прямой, проходящей через середину области определения углов  $\beta_i$ . При этом высоты, на которых находятся пары точек, различны. Обозначим эти высоты через  $h_1$  и  $h_2$ .



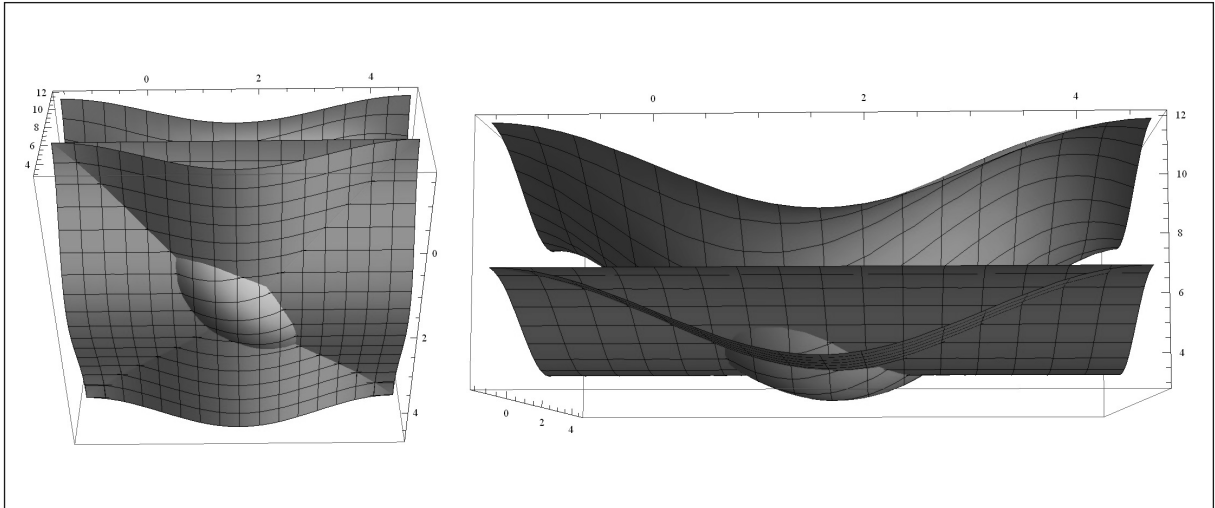


Рис. 6: Графики квадратов двух расстояний, изображенные одновременно.

Ясно, что если подобрать потенциал  $U(x)$  так, чтобы  $x_0$  равнялось  $h_1$  или  $h_2$ , то минимальная конфигурация  $L$  будет реализовываться при тех углах  $\beta_i$ , которые порождают соответствующую пару. Если же  $x_0$  отлично от  $h_1$  и  $h_2$ , то возможны различные варианты.

Мы решили провести численный эксперимент с различными  $U(x)$ . Для этого мы определили однопараметрическое семейство таких функций

$$U(x, a) = \frac{1}{(\sqrt{a} - 1)^2} \left( \frac{1}{x} - \frac{a}{x + 1} \right).$$

Заметим, что у этих функций минимальное значение всегда равно  $-1$ , а при изменении  $a$  от 0 до плюс бесконечности  $x_0$  принимает все возможные положительные значения.

Рассмотрим функцию  $\mathcal{U}(L, a) = U(|A_0A_3|^2, a) + U(|A_1A_4|^2, a) + U(|A_0A_4|^2, a)$ . Будем менять  $a$  и смотреть, как изменяется множество минимумов. Это множество минимумов соответствует минимальным конфигурациям ломаной  $L$ .

На рис. 7 изображен случай  $a = 1.41$ . Отчетливо видно наличие одной симметричной пары, соответствующей минимальной конфигурации, а другой — локально минимальной конфигурации с большим значением потенциала. Таким образом, здесь имеется 4 локальных минимума.

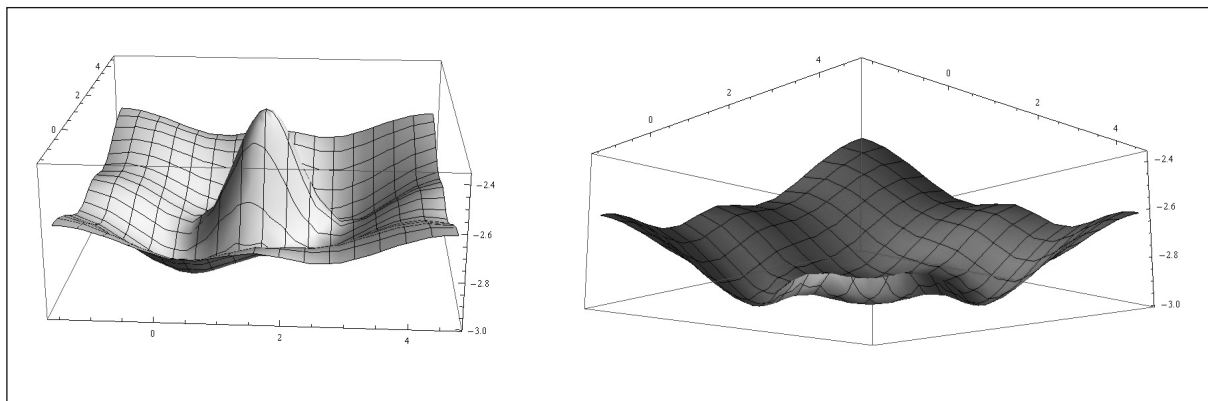


Рис. 7: Графики потенциала  $\mathcal{U}(L, 1.41)$ .

Приведем теперь вид потенциала  $\mathcal{U}(L, a)$  при  $a = 1.56$ , рис. 8.

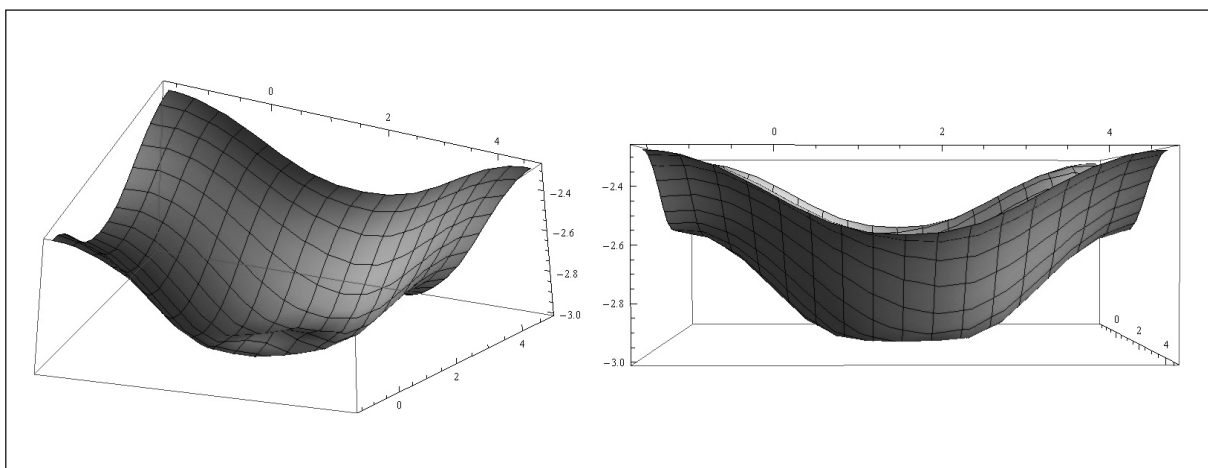


Рис. 8: Графики потенциала  $\mathcal{U}(L, 1.56)$ .

Множество минимумов здесь выглядит как окружность. Конечно, настоящие минимумы скорее всего устроены по-другому, но, при наличии спонтанного (температурного) движения, конфигурация, минимизирующая потенциал, будет принимать все значения из найденной окружности.

Мы также рассмотрели случай 6 атомов: теперь имеется три угловых параметра, поэтому вместо графиков приходится рассматривать поверхности уровня потенциала. На рис. 9 показана поверхность уровня для  $k = 3$  и потенциала  $U(x, 1.2)$ . Хорошо видны 6 из 8 симметричных областей, внутри которых находятся минимумы.

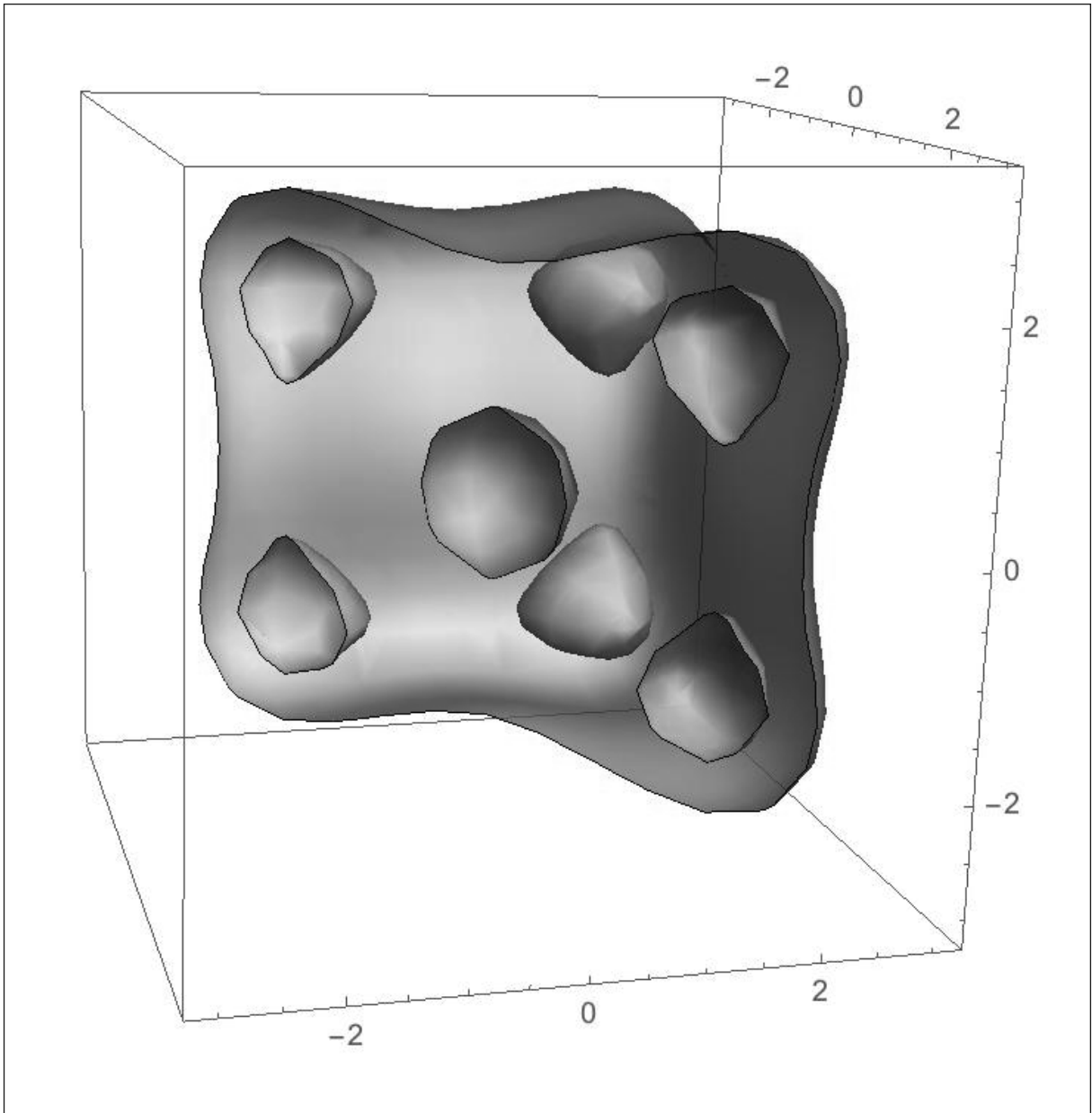


Рис. 9: Поверхности уровня потенциала  $\mathcal{U}(L, 1.2)$  для  $k = 3$ . Значения функции  $-2.8$  и  $-2.95$ .

Для  $k = 4$  значения потенциала в точках локальных минимумов становятся разными. На рис. 10 показаны все 8 областей, внутри которых находятся минимумы. Хорошо видно, что размеры областей отличаются: чем меньше область, тем больше значение минимума внутри.

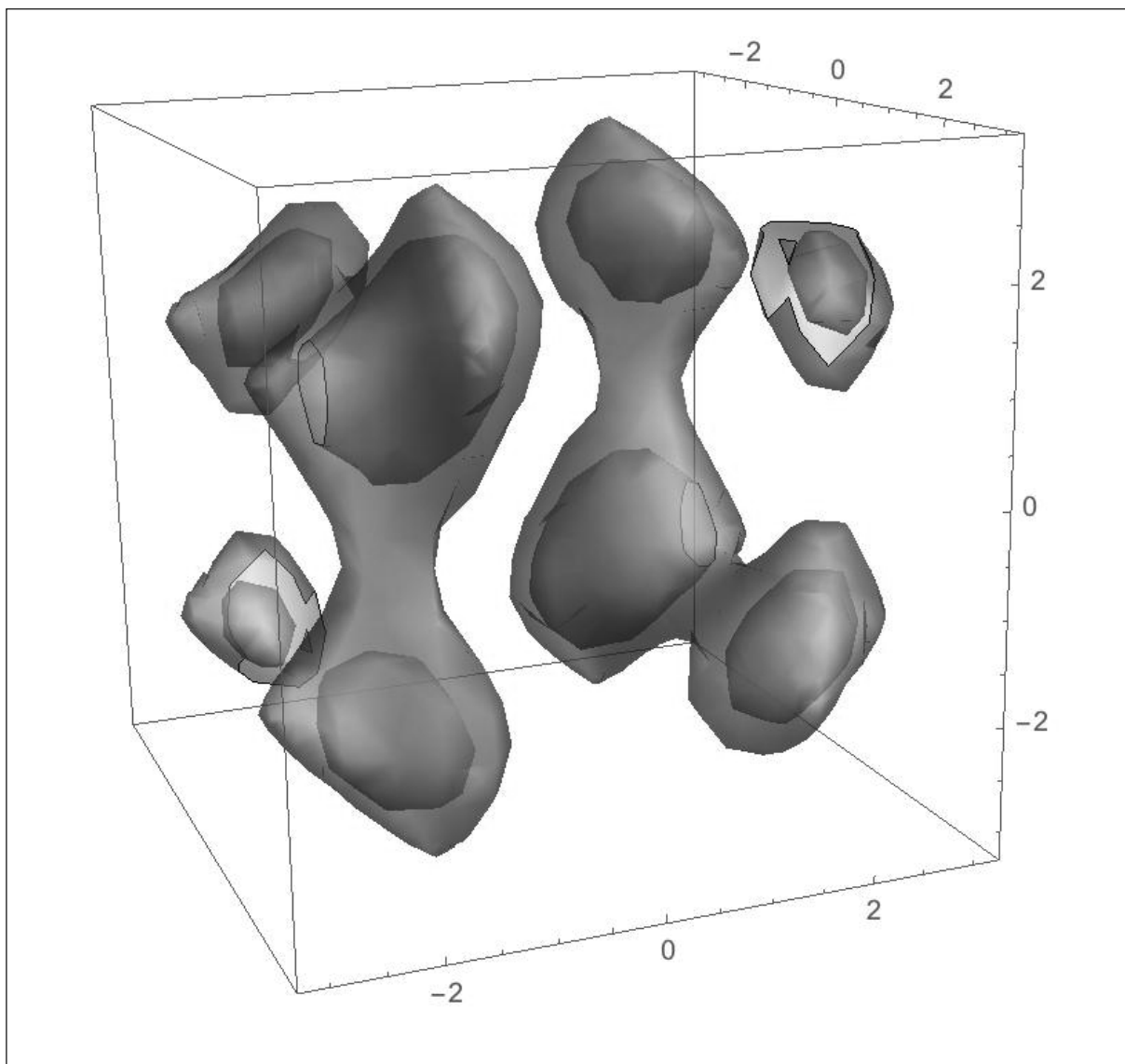


Рис. 10: Поверхности уровня потенциала  $\mathcal{U}(L, 1.2)$  для  $k = 4$ . Значения функции  $-4.65$  и  $-4.75$ .

**ВЫВОД:** для того, чтобы разобраться в реальном поведении белков, нужно знать точный вид потенциала  $U(x)$ , а не только его качественную форму.

## 5 Потенциал Леннард–Джонса

В качестве конкретного потенциала мы взяли известный потенциал Леннард–Джонса, известный также как потенциал 6–12. Он широко используется при моделировании парных взаимодействий, в том числе, при изучении углеродных цепочек. Напомним, что он имеет вид

$$U(x) = 4\epsilon \left( \frac{\sigma^{12}}{x^{12}} - \frac{\sigma^6}{x^6} \right),$$

где  $x$  — расстояние между взаимодействующими частицами, а  $\varepsilon$  и  $\sigma$  — некоторые параметры, которые для случая углеродных цепочек принимают значения  $\varepsilon = 0.6$  и  $\sigma = 3.39967$ . Мы проанализировали цепочки из пяти и шести атомов для  $k = 4$ .

### 5.1 Случай пяти атомов

У конфигурации имеется два угловых параметра. Полученный график для всевозможных значений углов приведен на рис. 11. На графике видно большое плоское “плато,” каждая точка которого претендует на то, чтобы быть минимумом.

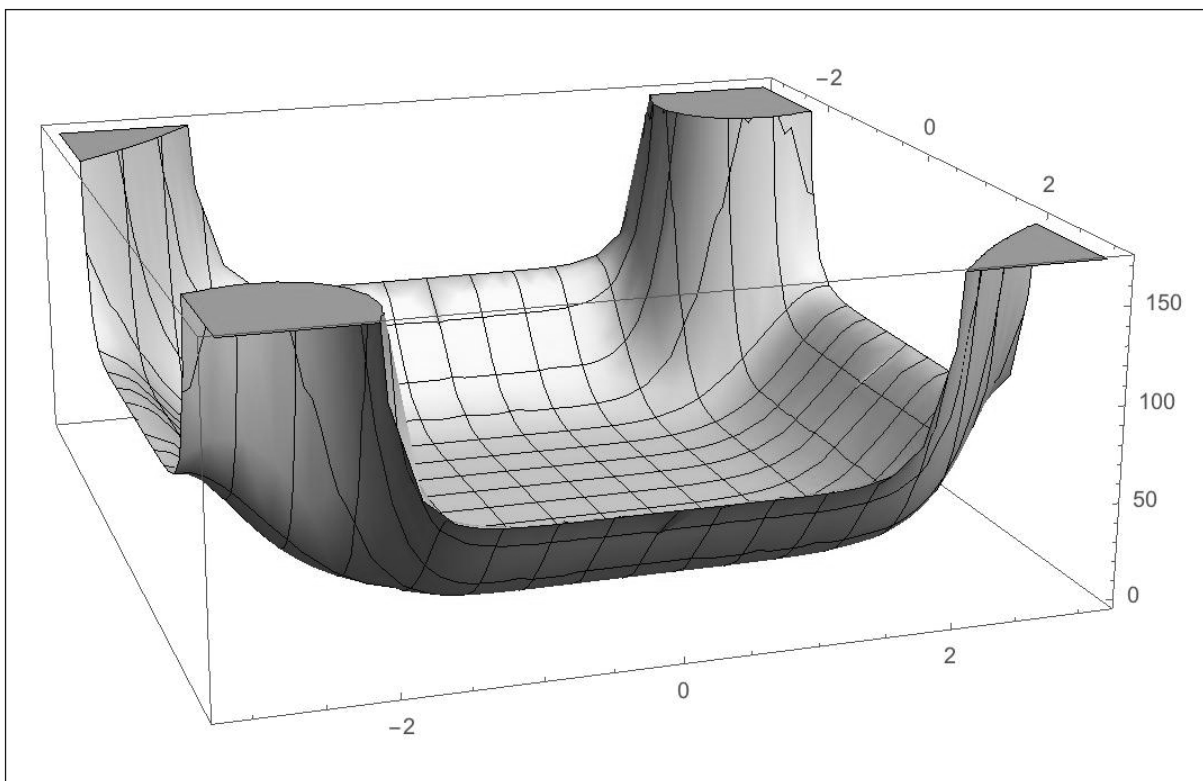


Рис. 11: График потенциала Леннарда–Джонса для пяти точек и  $k = 4$ .

Впрочем, удастся подобрать значения параметров так, чтобы настоящие минимумы стали более заметны, см. рис. 12. Здесь отчетливо видны два неглубоких минимума.

### 5.2 Случай шести атомов

У рассматриваемой конфигурации имеется уже 3 угловых параметра. Кроме того, точная формула для потенциала содержит много слагаемых с малыми коэффициентами. Для упрощения вычислений мы выбросили слагаемые с коэффициентами меньшими (по абсолютной величине) чем  $10^{-10}$ . Мы рассмотрели сначала случай  $k = 3$ . Нам удалось поймать 8 точек минимума, см. рис. 13, однако, разница значений настолько невелика, что, скорее, речь идет о целом критическом множестве (см. рис. 14). Разница значений потенциала на этих двух рисунках 0.0001.

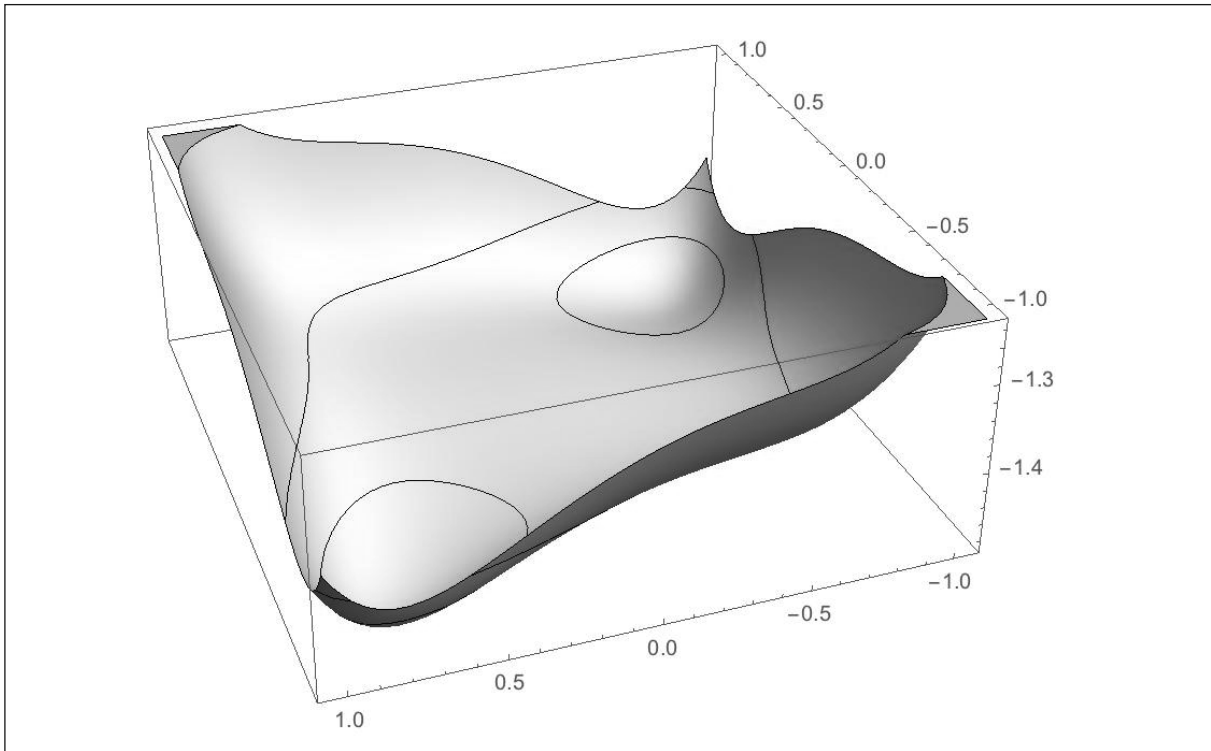


Рис. 12: Срезка графика потенциала Леннарда-Джонса для пяти точек и  $k = 4$ .

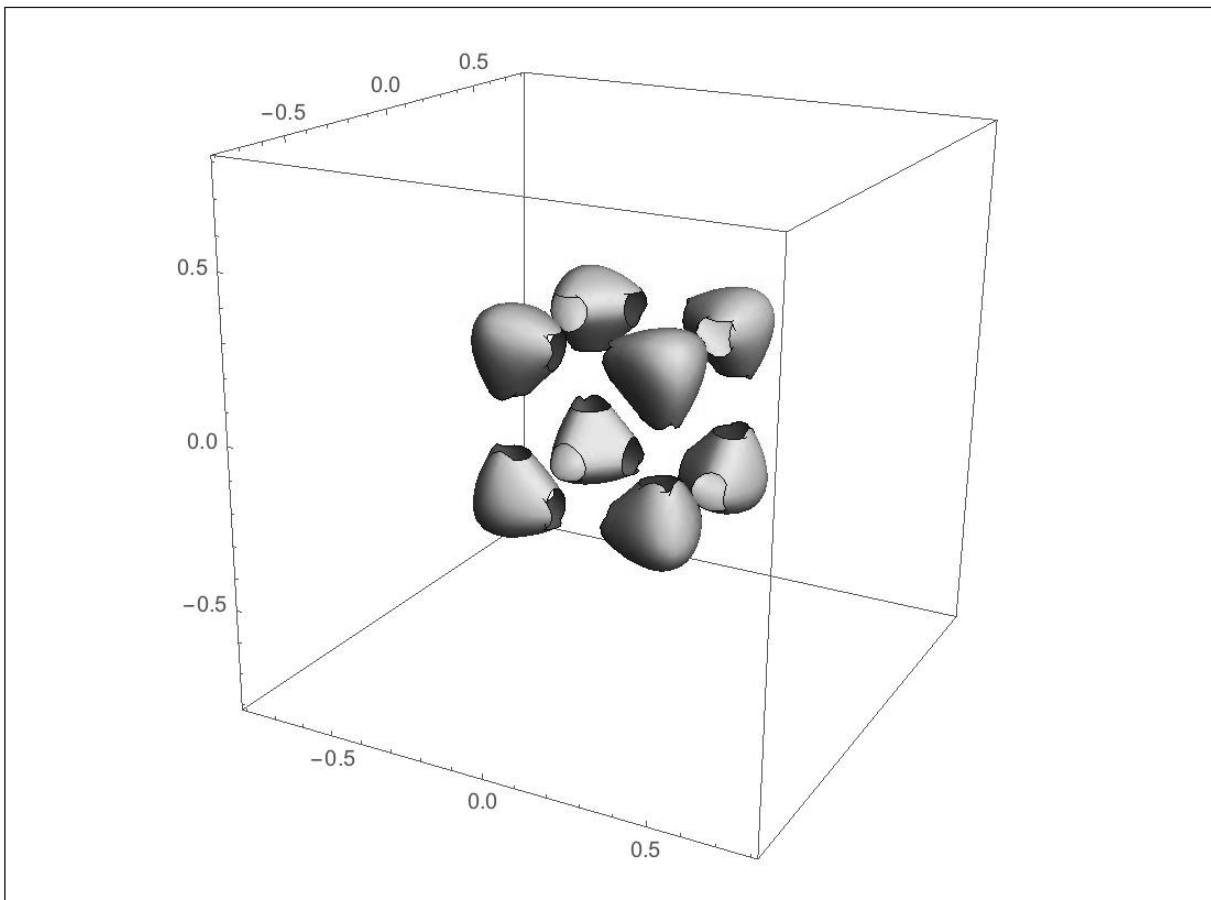


Рис. 13: Поверхность уровня потенциала Леннарда-Джонса для шести точек и  $k = 4$ , значение потенциала  $-1.7996$ .

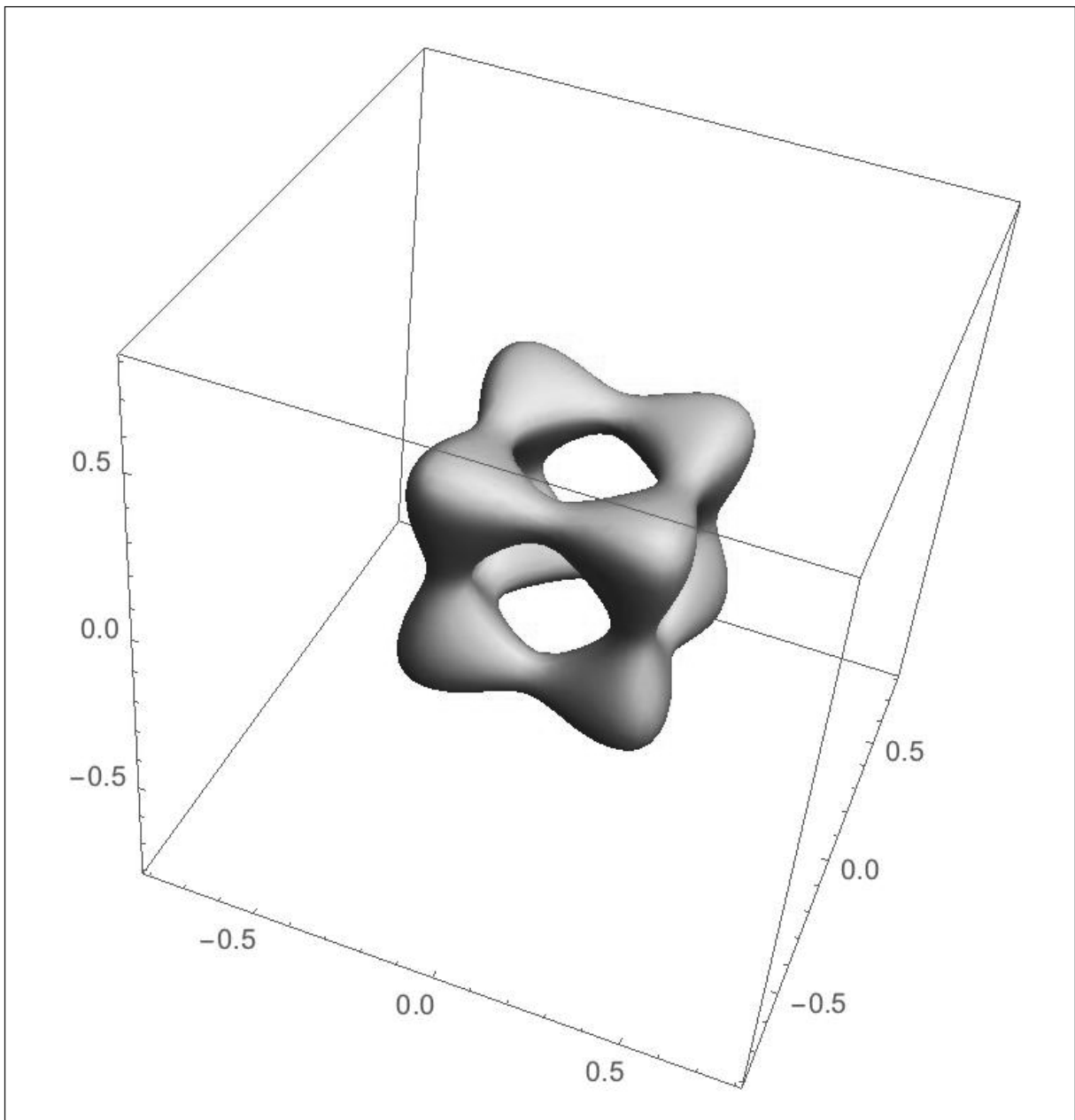


Рис. 14: Поверхность уровня потенциала Леннард-Джонса для шести точек и  $k = 4$ , значение потенциала  $-1.7995$ .

Поверхности уровня потенциала для  $k = 4$  показаны на рис. 15. Снова видна большая область (зеленый “куб”), внутри которой функция меняется мало. Нам удалось выделить четыре слабо выраженных минимума, расположенных в синих областях. Более отчетливо они видны на рис. 16. Впрочем, разность значений снова очень мала, порядка  $0.001$ , поэтому снова можно говорить о критическом множестве.

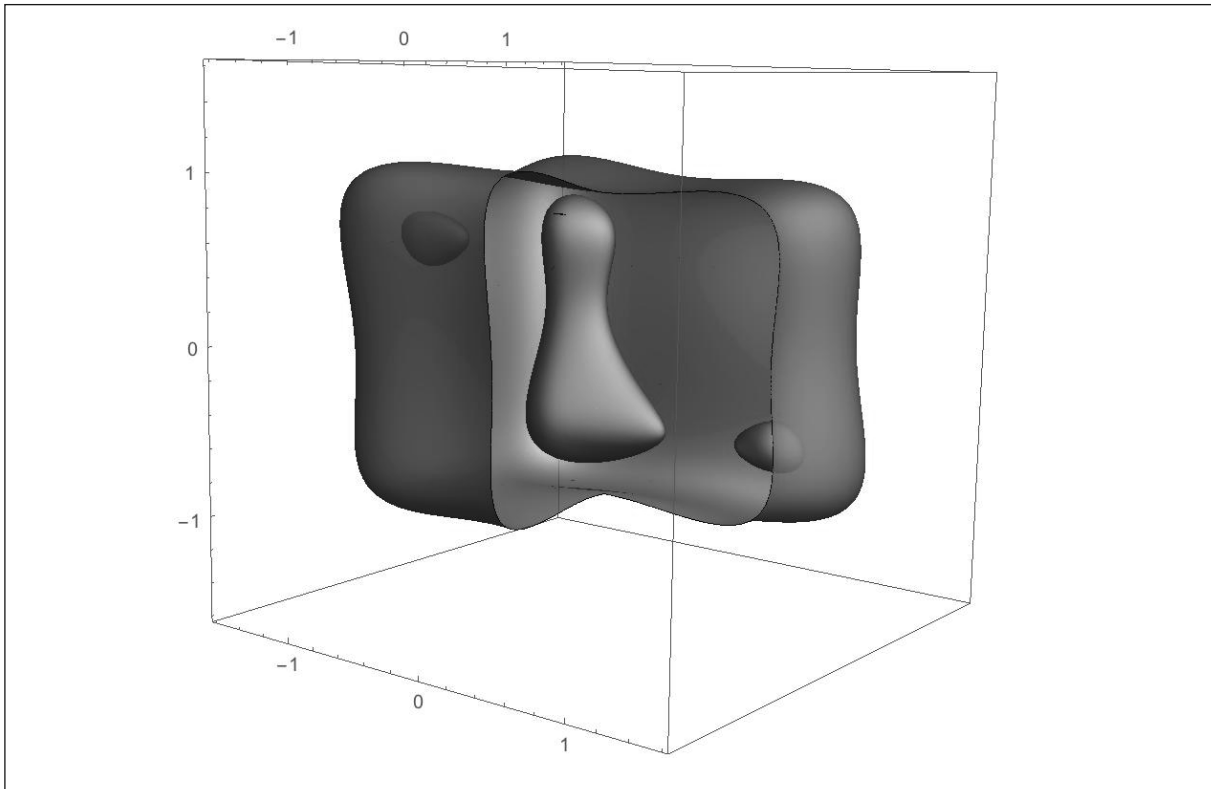


Рис. 15: Поверхности уровня потенциала Леннарда-Джонса для шести точек и  $k = 4$ .

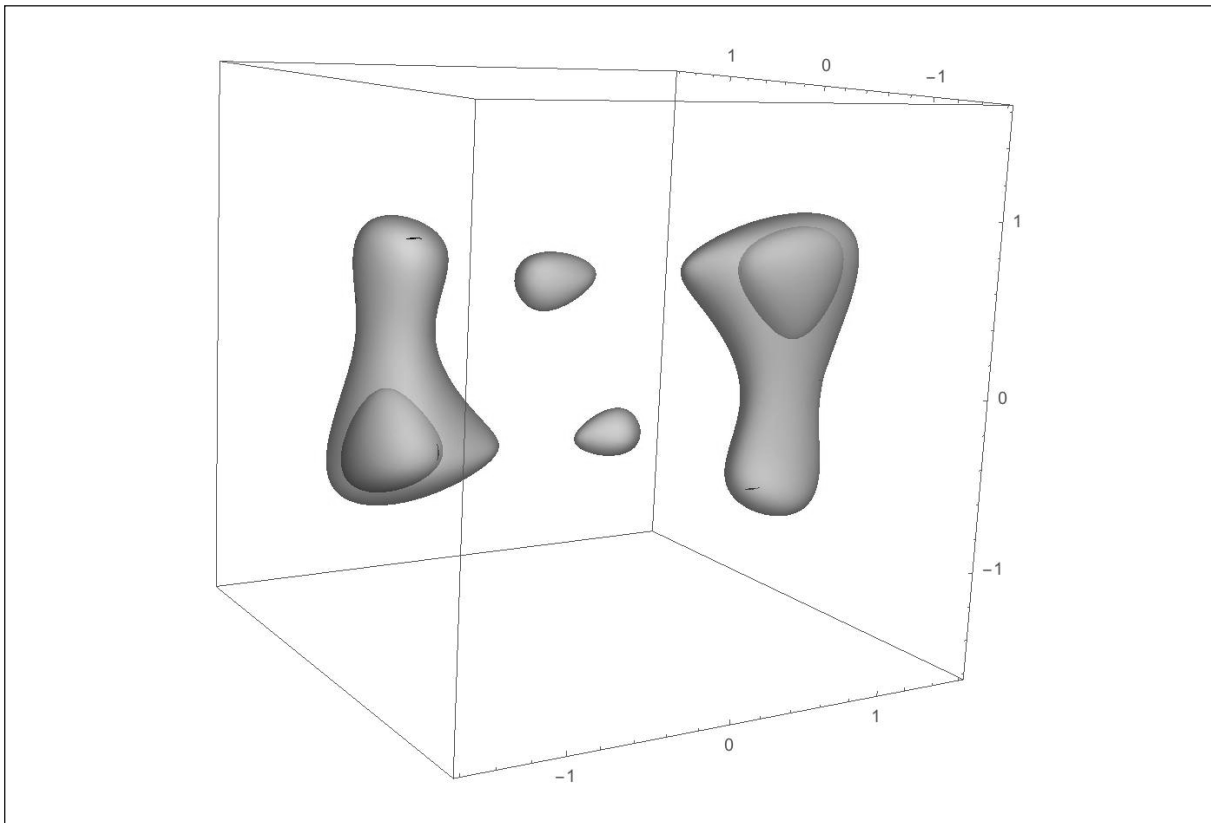


Рис. 16: Более плотные уровни потенциала Леннарда-Джонса для шести точек и  $k = 4$  вблизи минимального значения.



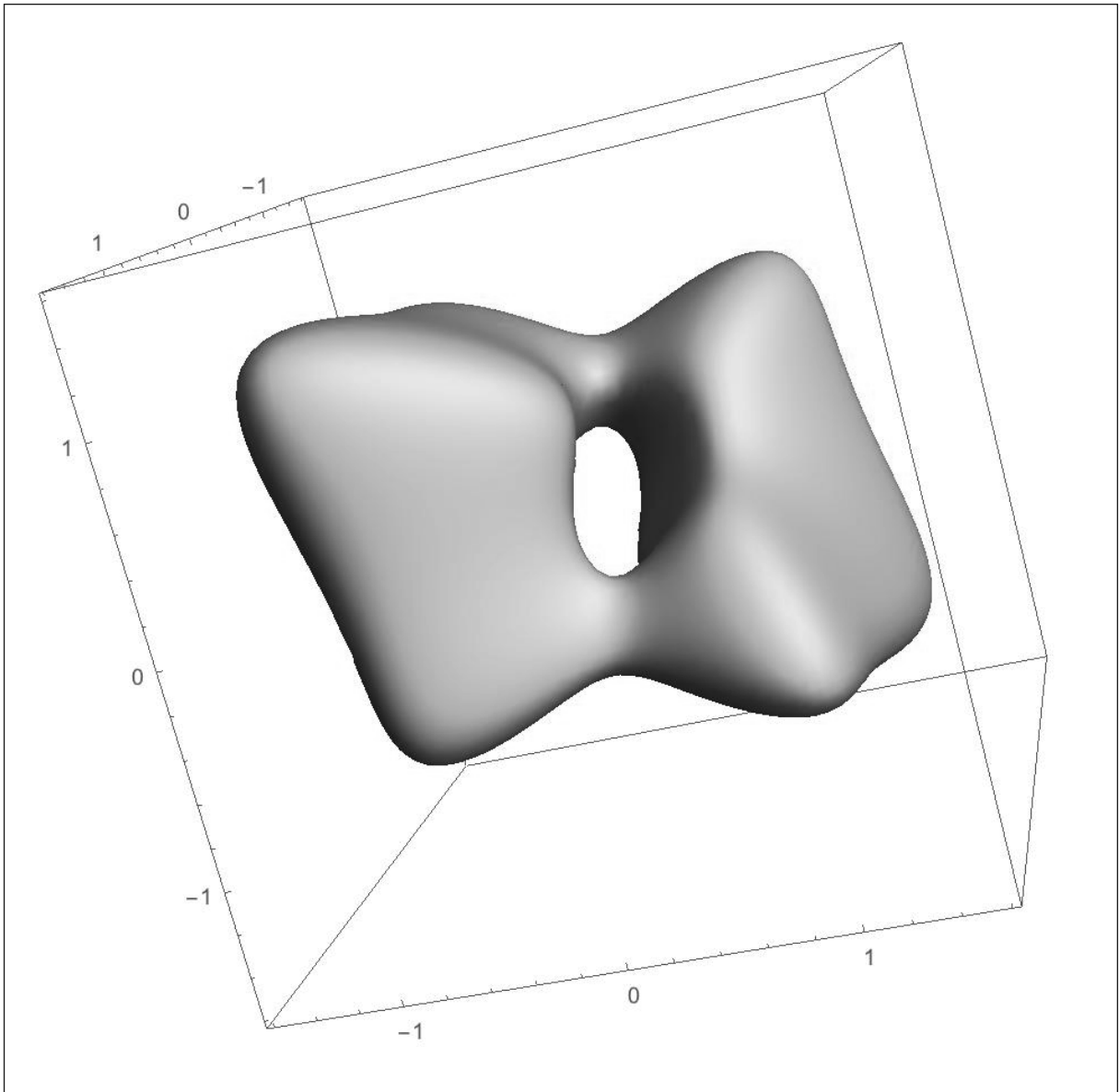


Рис. 17: Поверхность уровня потенциала Леннард–Джонса для шести точек и  $k = 4$ , ограничивающая точки минимума.

**ВЫВОД.** Модель углеродной цепочки, основанная на потенциале Леннард–Джонса, неустойчива (этот эффект известен — углеродная цепочка легко меняет свою пространственную конфигурацию). Так как реальные белки как правило не обладают такой “гибкостью”, то для их моделирования нужен другой потенциал, дающий более “глубокие” минимумы.

## Список литературы

- [1] *Ivanov A.O., Mishchenko A.S., Tuzhilin A.A.* Geometry of space curves and applications to polymers conformations investigation // Educational Internet-Journal “Computer Graphics & Geometry”, 2007, **9**(2), 43–63.
- [2] *Иванов А.О., Мищенко А.С., Тужилин А.А.* Геометрия ломаных и полипептидов // Наноструктуры. Математическая физика и моделирование, 2014, **10**(1), 39–48; <http://nano-journal.ru>
- [3] *Иванов А.О., Мищенко А.С., Тужилин А.А.* Геометрия аминокислот и полипептидов // Наноструктуры. Математическая физика и моделирование, 2014, **10**(1), 49–76; <http://nano-journal.ru>
- [4] *Ivanov A.O., Mishchenko A.S., Tuzhilin A.A.* Critical analysis of amino acids and polypeptides geometry // In Continuous and Distributed Systems: Theory and Applications, Springer, 2015, **2**, 29–74; <http://springer.com>
- [5] *Неелов И.М.* Введение в молекулярное моделирование биополимеров // СПб: НИУ ИТМО, 2014, 101 стр.

## CONFORMATION OF ARTICULATED-LINK CHAIN MODELING PROTEINS FOLDING

A. O. Ivanov, A. A. Tuzhilin, E. A. Vilkul

*Lomonosov Moscow State University, Moscow, Russia*

[elena.tuzhilina@mail.ru](mailto:elena.tuzhilina@mail.ru), [aoiva@mech.math.msu.su](mailto:aoiva@mech.math.msu.su), [tuz@mech.math.msu.su](mailto:tuz@mech.math.msu.su)

Received 11.08.2015

In the present paper we investigate the geometry of the simplest model which may be useful in studying the proteins conformation. Namely, we consider a spacial polygonal line (one can imagine, for example, the protein backbone), fix its edges lengths, fix the angles between adjacent edges, assign potential interaction between its vertices, and study the configurations of such a polygonal line corresponding to the local minima of the total potential. Under some assumptions on the interaction type, this problem can be solved completely. For less restrictions, we could only conduct computer simulations. However, the cases in consideration gave us the possibility to find a few quality phenomena. For example, during a change of the potential function preserving its potential well form, we observed how discrete local minima were included in big flat domains with very small variation of the values: permanent spontaneous atoms motion obliges to refer such domains to solutions as well. Besides, both our theoretical results and computer experiments lead to a conjecture estimating the maximal number of strict local minima of the total potential, which already for 30 amino acids turned out to be very large. We think that such estimations may be useful for adequate understanding the situation in the modern science investigating proteins folding by means of nuclear magnetic resonance and x-ray analysis.