

## КРАТКИЕ СООБЩЕНИЯ

УДК 547:541.6:51-7

## ИНВЕРСИИ ПРОЕКЦИЙ ЗАУЗЛЕННЫХ МОЛЕКУЛЯРНЫХ СТРУКТУР

© 2009 Л. И. Макаров\*

*Институт математики им. С.Л. Соболева СО РАН, Новосибирск**Статья поступила 16 сентября 2008 г.*

Рассмотрены некоторые преобразования проекций молекулярных структур, содержащих замкнутую цепочку элементов, топологической моделью которой служит узел. Описано графовое представление диаграмм узлов, изложены способы инверсии точек проекции узла для получения разных его проекций.

**Ключевые слова:** заузленная молекулярная структура, графовое представление диаграммы, инверсия.

Создание компьютерных систем исследования молекулярных структур неклассического строения (катенанов, ротоксанов, ДНК, макромолекул белков и т.д.) [1—8] требует решения ряда задач обработки данных об их топологических моделях.

Некоторые из этих молекулярных структур содержат замкнутые цепочки элементов, которые можно представить в трехмерном пространстве в виде узлов, т.е. замкнутых кривых без самопересечений. Каждый узел может быть задан с помощью диаграммы, т.е. проекции узла на плоскость, в которой для каждого самопересечения проекции указано, какой из отрезков кривой в данной точке проходит выше другого [9]. Диаграмму узла, имеющую минимально возможное число пересечений проекции, называют стандартной. Узел называют тривиальным, если его стандартная диаграмма не имеет пересечений.

Молекулярную структуру называют заузленной, если в ней существует хотя бы одна замкнутая цепочка элементов, топологической моделью которой является нетривиальный узел. Примеры заузленных молекулярных структур приведены на рис. 1. Вид проекции узла, полученной при исследовании реальной заузленной молекулярной структуры, например, с помощью электронной микроскопии, зависит от его положения относительно плоскости и направле-

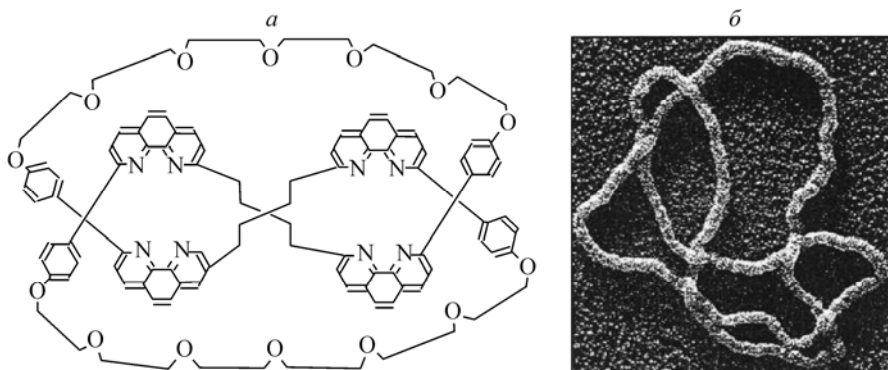


Рис. 1. Синтезированная заузленная молекулярная структура [7] — *a*, заузленная ДНК [6] — *б*

\* E-mail: makarov@math.nsc.ru

ния проектирования. При этом для одной и той же структуры могут быть получены разные виды проекций узла, в том числе и разные проекции стандартных диаграмм. Поэтому в базе данных таких структур необходимо иметь программные средства, позволяющие создавать разные виды проекций узла структуры, исходя из проекции, полученной при ее исследовании. При обработке данных о заузленных молекулярных структурах визуализация диаграмм их узлов используется для контроля данных о структурах, для их идентификации, для изображения структур, полученных при их компьютерном синтезе, и т.д.

В работе приведены некоторые средства представления и преобразования диаграмм узлов, соответствующих заузленным молекулярным структурам: графовое представление диаграмм узлов, получение разных проекций узла при инверсии заданных точек исходной его проекции и сглаживание кривой проекции узла с помощью сплайнов. При соответствующей модификации эти средства могут быть использованы и в процессе визуализации зацепленных структур.

**Графовое представление диаграмм узлов.** Проекцией узла на плоскость, не пересекающую узел, называют такую его проекцию, в которой в каждое ее самопересечение проектируется не более двух точек узла. Проекцию узла, в которой для каждого ее пересечения указано, какая из точек является при проектировании верхней, а какая — нижней, называют диаграммой узла. Для задания диаграммы используют разные графовые представления [ 10, 11 ], но каждое из них предназначено для конкретных целей исследования свойств узлов, например, для построения полиномиальных инвариантов узлов. Здесь предлагается графовое представление диаграммы узла, позволяющее применять эффективные алгоритмы построения изображений разных его проекций.

Для изображения проекции узла необходимо задать в плоскости проекции  $(X,Y)$  координаты всех пересечений, и для каждой дуги кривой, соединяющей пару пересечений, должны быть заданы координаты хотя бы одной ее точки (точки дуги). Верхнюю  $t$  и нижнюю  $u$  точки пересечения и точку дуги  $m$  можно указать, задавая по оси проектирования  $Z$ , перпендикулярной к плоскости  $(X,Y)$ , условные значения их координаты (одинаковые для каждого типа точек проекции):  $z_t > z_m > z_u$ .

Кроме пересечений и дуг кривой, их соединяющих, элементами проекции являются ее грани — части плоскости, ограниченные дугами кривой и не содержащие внутри никаких точек проекции. Бесконечную грань проекции называют внешней, а остальные грани — внутренними. Тогда диаграмму узла можно представить в виде двух помеченных графов:  $CG$  — контур-графа диаграммы и  $GG$  — графа граней его проекции. Вершины контур-графа соответствуют заданным точкам пересечений и дуг, а ребра — отрезкам дуг кривой, соединяющим эти точки. Контур-граф  $CG$  задается последовательным по направлению обхода списком координат заданных точек диаграммы, и последняя точка списка совпадает с первой. Пример такого списка приведен на рис. 2. Для контур-графа диаграммы узла, имеющей  $n > 0$  пересечений, длина списка составляет  $2n + n_0 + 1 \geq 4n + 1$ , где  $n_0$  — число заданных точек дуг.

Грани проекции узла задаются перечислением их граничных дуг кривой. Вершинам графа граней  $GG$  сопоставлены грани проекции и две вершины соединены между собой ребром, если соответствующие грани имеют общую дугу кривой.

Послойным видом графа граней (ПВГГ) назовем такое его описание [ 12 ], в котором первым слоем является вершина, соответствующая внешней грани, вторым слоем являются все вершины, смежные с ней, и в каждом следующем слое находятся все вершины, смежные с вершинами предыдущего слоя.

На рис. 2 показаны диаграмма узла, имеющая 6 пересечений, список  $CG$  координат ее заданных точек и послойный вид графа граней проекции.

**Инверсии точек проекции узла.** Две проекции узла назовем одинаковыми, если послойные виды их графов граней одинаковы, т.е. существует такая подстановка на множестве вершин графа граней одной проекции, которая переводит его ПВГГ в ПВГГ другой проекции, сохраняя отношение смежности. Две проекции узла назовем разными, если такая подстановка не существует.

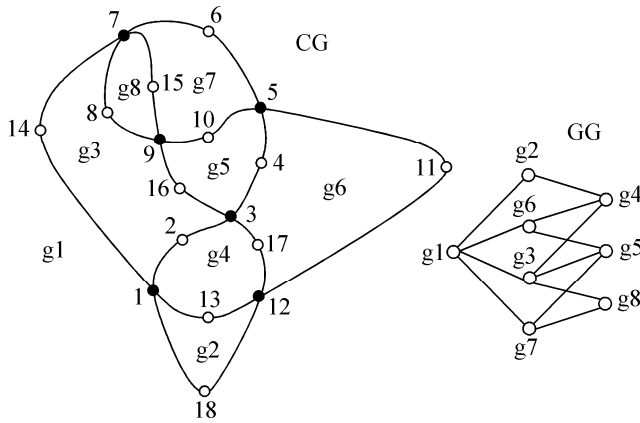


Рис. 2. Диаграмма узла, имеющая 6 пересечений, и список координат точек диаграммы (вершин контур-графа CG). Пересечения обозначены черными кружками, а заданные точки дуг — белыми. Грани проекции обозначены символами  $g$ . GG — послойный вид графа грани проекции

CG:  $(x_1 y_1 z_t) (x_2 y_2 z_m) (x_3 y_3 z_u) (x_4 y_4 z_m) (x_5 y_5 z_t) (x_6 y_6 z_m) (x_7 y_7 z_u) (x_8 y_8 z_m) (x_9 y_9 z_t) (x_{10} y_{10} z_m) (x_{11} y_{11} z_u) (x_{12} y_{12} z_t) (x_{13} y_{13} z_m) (x_{14} y_{14} z_m) (x_{15} y_{15} z_m) (x_{16} y_{16} z_m) (x_{17} y_{17} z_m) (x_{18} y_{18} z_m) (x_1 y_1 z_t)$

При изменении положения кривой узла в трехмерном пространстве и проектировании его на одну плоскость можно получать разные его проекции. Например, внутренняя грань одной проекции может стать внешней гранью другой. Поэтому, хотя для разных проекций их графы граней могут быть изоморфны, послойные виды этих графов могут отличаться друг от друга (см. рис. 2, 3).

Любую внутреннюю грань  $g$  заданной проекции узла можно преобразовать во внешнюю грань другой его проекции с помощью инверсии [ 13 ] точек плоскости относительно окружности радиуса  $R$  с центром  $C$ , находящимся в грани  $g$ . Пусть  $r(C, p) > 0$  — расстояние между центром инверсии  $C$  и точкой плоскости  $p$ . Для заданных точек пересечений и дуг кривой  $p_i$  проекции узла обозначим  $r^h = \min_i r(C, p_i)$ ,  $r^e = \max_i r(C, p_i)$ . Рассмотрим следующие типы инверсий, радиусы окружностей которых зависят от величин  $r^h$  и  $r^e$ .

*Гиперболическая инверсия IG.* Инверсия IG — это такое отображение точек  $p$  плоскости, что  $IG(p) = p'$  и  $p'$  принадлежит лучу  $\langle C, p \rangle$ , а  $r(C, p') = R^2/r(C, p)$ , где  $R^2 = r^h \cdot r^e$ .

*Стереографическая инверсия (проекция) IS.* При инверсии IS отображение точки  $p$  в точку  $p'$  производится следующим образом. Задается радиус сферы инверсии  $R = r^e$ . Сфера располагается так, что ее экваториальная плоскость совпадает с плоскостью проекции и ось полюсов  $NS$  проходит через центр инверсии  $C$ . Точки проекции проектируются на верхнюю полусферу. Каждая полученная точка проектируется по лучу из верхнего полюса  $N$  в инвертированную точку  $p'$  пересечения его с экваториальной плоскостью. При этом  $r(C, p') = r(C, p)R/(R - (R^2 - r^2(C, p))^{1/2})$  по лучу  $\langle C, p \rangle$ ,  $0 < r(C, p) \leq R$ .

*Экваториальная инверсия IE.* При инверсии IE отображение точки  $p$  в точку  $p'$  производится следующим образом. Задается радиус сферы инверсии из соотношения  $R^2 = (r^e)^2 + (r^h)^2$ . Сфера располагается так, что ее экваториальная плоскость совпадает с плоскостью проекции

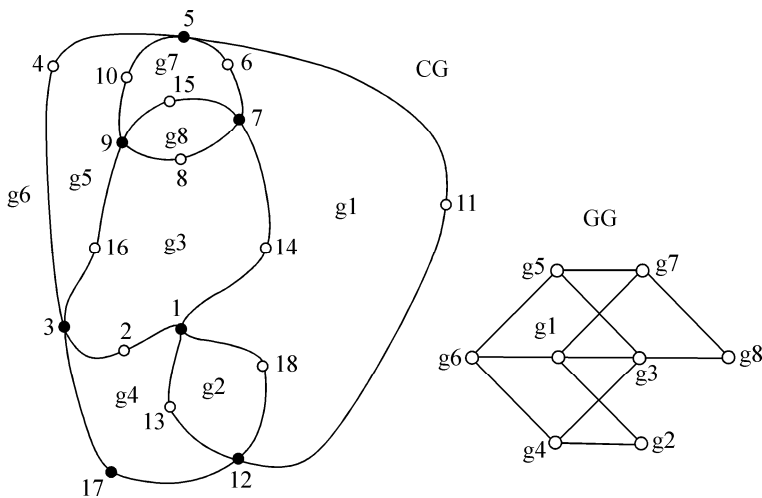


Рис. 3. Сглаженная кривая линейно инвертированной проекции CG узла (см. рис. 2) и послойный вид ее графа грани GG. Инверсия произведена с центром в грани  $g_6$  исходной проекции узла

и ось полюсов проходит через центр инверсии  $C$ . Точки  $p$  проектируются на верхнюю полусферу. Каждая полученная точка полусферы перемещается на  $90^\circ$  в нижнюю полусферу по окружности сечения сферы плоскостью, проходящей через эту точку и ось полюсов сферы. Затем эти точки проектируются в точки  $p'$  на экваториальную плоскость. При этом  $r(C, p') = (R^2 - r^2(C, p))^{1/2}$  по лучу  $\langle C, p \rangle$ ,  $0 < r(C, p) < R$ .

*Линейная инверсия ИЛ.* Инверсия ИЛ это такое отображение точек  $p$  плоскости, что  $ИЛ(p) = p'$  и  $p'$  принадлежит лучу  $\langle C, p \rangle$ , а  $r(C, p') = 2R - r(C, p)$ , где  $R = (r^e + r^h)/2$ ,  $0 < r(C, p) < 2R$ .

Анализ выражений для преобразования расстояний точек  $p$  плоскости по лучу  $\langle C, p \rangle$  от центра инверсии  $C$ , лежащего внутри некоторой грани  $g$  проекции, показывает следующее. При инверсиях ИГ, ИЛ и выбранных значениях радиуса  $R$  точки плоскости, принадлежащие отрезку  $[r^h, R]$  луча  $\langle C, p \rangle$ , отображаются в точки отрезка  $[r^e, R]$  этого луча и наоборот. Аналогично, при инверсии ИЕ точки отрезка  $[r^h, r^e]$  луча  $\langle C, p \rangle$  отображаются в точки отрезка  $[r^e, r^h]$  этого луча и наоборот. Поэтому точка исходной проекции, имеющая в ней расстояние  $r^h$ , переходит в инвертированной проекции в точку, имеющую расстояние  $r^e$ , и наоборот. Таким образом, в инвертированной проекции величины  $r^h$  и  $r^e$  относительно центра  $C$  инверсии совпадают с этими же величинами исходной проекции. Точки, имеющие  $r(C, p) = R$  для инверсий ИГ, ИЛ, ИС и  $r(C, p) = R\sqrt{2}$  для инверсии ИЕ, остаются на месте. Для инверсии ИС при  $k = r^h/r^e$  справедливо  $r^e \leq k \max_i r(C, p'_i) \leq 2r^e$ , где  $k$  — масштабный коэффициент, при котором  $k \min_i r(C, p'_i) = r^h$ .

Легко видеть, что для размера  $D$  исходной проекции узла (наибольшего расстояния между ее заданными точками пересечений и переходов) выполняется  $D \leq 2r_0^e$ , где  $r_0^e = \max_j r_j^e$ , и величина  $r_j^e$  соответствует инверсии с центром в грани  $g_j$ . Таким образом, размер инвертированной проекций для инверсий ИГ, ИЛ, ИЕ не превышает размера  $D$  исходной проекции, а для инверсии ИС он может быть вдвое больше этой величины при  $k = r^h/r^e$ .

Нелинейность выражений для преобразования расстояний точек от центра для инверсий ИГ, ИЕ, ИС приводит к увеличению вероятности появления дополнительных точек пересечений при сглаживании дуг кривой инвертированной проекции с помощью сплайна.

Если проекция узла содержит грани, эквивалентные в смысле их одинакового расположения относительно всех остальных ее граней, то граф граней проекции содержит соответствующие эквивалентные вершины. При этом множество вершин графа граней может быть разбито на орбиты — классы эквивалентных вершин [14]. Для эквивалентных вершин одной орбиты "все одинаково": степени равны, графы, порожденные окружением, изоморфны и т.д. Таким образом, послонные виды графов граней, соответствующие вершинам одной орбиты, одинаковы и инверсия с центрами в эквивалентных гранях приводит к созданию одинаковых проекций. Отсюда следует, что для получения разных проекций узла достаточно производить инверсии только с центрами в гранях, соответствующих вершинам — представителям разных орбит графа граней, и число разных проекций узла при таких инверсиях равно числу этих орбит. Для узла рис. 2 орбитами множества вершин графа граней являются 4 их класса:  $(g1, g3)$ ,  $(g2, g8)$ ,  $(g4, g7)$ ,  $(g5, g6)$ .

Контур-граф диаграммы узла задает кривую его проекции с помощью небольшого числа заданных точек пересечений и дуг, что дает грубое представление о расположении ее на плоскости. Для более гладкого изображения кривой можно применить интерполяцию ее сплайном. Сплайн является приближенным аналитическим выражением непрерывной функции, описывающей кривую, по заданным ее значениям в конечном множестве точек независимых переменных. Для изображения диаграмм узлов использован алгоритм построения интерполяционного параметрического кубического (ИПК) сплайна [15]. Быстродействие алгоритма линейно зависит от числа заданных точек. На рис. 3 изображена сглаженная кривая линейно инвертированной проекции узла (см. рис. 2), построенная по его контур-графу. Линейная инверсия произведена с центром в грани  $g6$  исходной проекции узла.

В работе предложено представление диаграммы узла в виде пары графов (контур-графа и графа граней), которое обеспечивает удобный и надежный ввод данных о диаграммах узлов и позволяет использовать эффективные алгоритмы их изображения.

Из анализа выражений для преобразования расстояний в рассмотренных типах инверсий и из проведенных сравнительных экспериментов можно рекомендовать линейную инверсию для применения при изображении проекций. При этом для получения из исходной проекции других проекций с помощью инверсии достаточно производить ее только с центрами в гранях, соответствующих вершинам — представителям разных орбит графа граней.

Для интерполяции кривой проекции узла по последовательности обхода ее заданных точек использован эффективный алгоритм построения ИПК сплайна, создающий достаточно гладкую кривую проекции.

Кроме рассмотренных выше в Институте математики СО РАН разработаны и другие средства обработки данных о заузленных молекулярных структурах: экранный редактор, позволяющий вводить и корректировать графы диаграмм узла, средства идентификации узлов с использованием полиномиальных инвариантов [ 11 ] и т.д. Приведенные алгоритмы и программные средства могут быть использованы при создании компьютерных систем регистрации, экспертных систем и систем автоматизации научных исследований в химии и биологии, предназначенных для накопления и обработки данных о соединениях, имеющих заузленные молекулярные структуры.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант № 08-01-00673).

#### СПИСОК ЛИТЕРАТУРЫ

1. Шилл Г. Катенаны, ротаксаны и узлы. – М.: Мир, 1973.
2. Liu L., Depew R.E., Wang J.C. // J. Mol. Biol. – 1976. – **106**. – P. 439.
3. Liu L.F., Liu C.C., Alberts B.M. // Cell. – 1980. – **19**. – P. 697.
4. Mizuuchi K., Fisher L.M., O'Dea M.H., Gellert M. // Proc. Nat. Acad. Sci. USA. – 1980. – **77**. – P. 1847.
5. Франк-Каменецкий М.Д., Вологодский А.В. // Успехи физ. наук. – 1981. – **134**, № 4. – С. 641.
6. Wasserman S.A., Cozzarelli N.R. Biochemical topology: Applications to DNA recombination and replication. – Science. – 1986. – **232**. – P. 951 – 960.
7. Dietrich-Buchecker C.D., Sauvage J.P. // Angew. Chem. (Int. Eng. Ed.) – 1989. – **28**. – P. 189 – 192.
8. New Scientific Applications of Geometry and Topology / Proc. Sympos. Appl. Math. (Ed. D.W. Sumners) – Vol. 45, Amer. Math. Soc., Providence, RI, 1994.
9. Reidemeister K. Knotentheorie, Ergebnisse der Mathematik. – 1932. – **1**. – P. 1.
10. Liang C., Mislow K. // J. Math. Chem. – 1994. – **15**. – P. 1 – 34.
11. Веснин А.Ю., Добрынин А.А. Вычислительные системы. Вып. 155. – Новосибирск, 1996. – С. 37.
12. Скоробогатов В.А. Вычислительные системы. Вып. 69. – Новосибирск, 1977. – С. 6 – 9.
13. Александров А.Д., Нецветаев Н.Ю. Геометрия. – М.: Наука, 1990.
14. Емеличев В.А., Мельников О.И., Сарванов В.И., Тышкевич Р.И. Лекции по теории графов. – М.: Наука, 1990.
15. Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л. Методы сплайн-функций. – М.: Наука, 1980.