

Терабайт оперативной памяти в общем адресном пространстве: новые возможности для метода трансфер-матрицы*

С. С. Акименко¹, В. А. Горбунов¹, А. В. Мышлявцев^{1,2}, М. Д. Мышлявцева¹,
П. В. Стищенко¹, В. Ф. Фефелов¹

Омский государственный технический университет¹, Институт проблем переработки углеводородов СО РАН²

В модели многоцентровой адсорбции на гексагональной решетке ранее было обнаружено явление «чёртовой лестницы» фазовых переходов. Это явление затрудняет исследование фазового поведения при ненулевой температуре и требует использования детерминистского метода — метода трансфер-матрицы. Требования к объёму оперативной памяти (RAM) в этом методе растут экспоненциально с размером системы. С помощью вычислительного сервера на базе шины Intel QPI с 1 ТБ RAM в общем адресном пространстве нам удалось исследовать некоторые интересные аспекты данной модели, ранее недоступные для изучения из-за высоких требований к объёму RAM и скорости доступа к ней.

1. Введение

В работах [1, 2] нами была построена решёточная модель адсорбции гомоядерных димеров, учитывающая возможность различной ориентации молекул по отношению к поверхности твёрдого тела. В случае адсорбции на сотовой решётке было показано, что в основном состоянии ($T = 0$ К) в системе наблюдается явление «чёртовой лестницы» – при повышении давления в газовой фазе переход из первой упорядоченной фазы, с наименьшей плотностью, в последнюю фазу, максимально плотную, сопровождается каскадом фазовых переходов через бесконечную последовательность упорядоченных структур. Данное явление, помимо очевидной теоретической ценности, весьма интересно с практической точки зрения, поскольку открывает новые возможности для разработки процессов управляемой самосборки наноразмерных структур.

Для полноты изложения опишем решёточную модель монослойной адсорбции гомоядерных димеров на гексагональной решётке [2]. Поверхность представляет собой набор упорядоченных активных центров (АЦ). На **Рис. 1** показана структура поверхности.

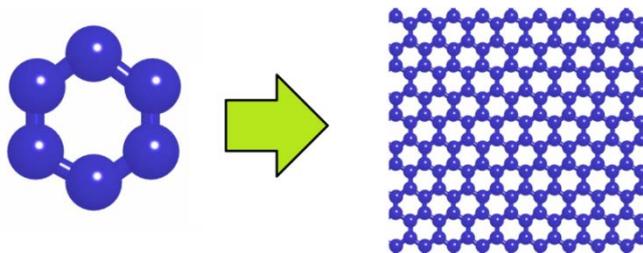


Рис. 1. Формирование гексагональной решетки

Димер – это молекула, состоящая из двух сегментов с фиксированной длиной связи равной постоянной решётки. Димер может адсорбироваться двумя различными способами: вертикально, занимая один активный центр, и горизонтально, занимая два активных центра (**Рис. 2**). Предполагается, что адсорбция димера на 2 АЦ более выгодна, чем на 1 АЦ, с точки зрения величины энергии адсорбции. В предложенной модели учитывалась разница h между энергиями адсорбции на 2 АЦ и 1 АЦ

* Данная работа поддержана Минобрнауки России (государственное задание №16.2413.2014/К) и президентским грантом для ведущих научных школ Российской Федерации (5998.2014.10).

В модель были введены бесконечно сильные отталкивающие взаимодействия между ближайшими адсорбированными молекулами, которые в первом приближении учитывают собственный размер и химическую структуру молекулы. Бесконечно сильное отталкивание проявляется в том, что ближайшие активные центры вокруг адсорбируемой молекулы являются недоступными для адсорбции других молекул.

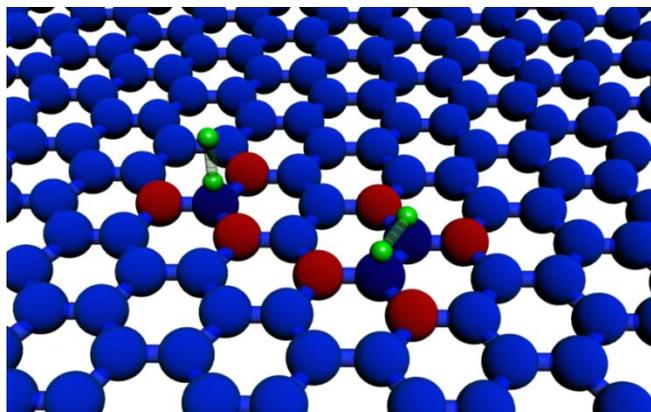


Рис. 2. Способы ориентации димера на поверхности: слева – вертикально, справа – горизонтально. Красным отмечены запреты на ближайшее соседство.

На **Рис. 3** представлена фазовая диаграмма системы в основном состоянии (при $T = 0$ К). Видно, что интервалы существования упорядоченных структур сужаются и стремятся к нулю по мере приближения химического потенциала к значению $3h/2$, с дальнейшим ростом химического потенциала на поверхности образуется максимально плотная структура, образованная только димерами, адсорбированными на 1 АЦ ($n = \infty$).

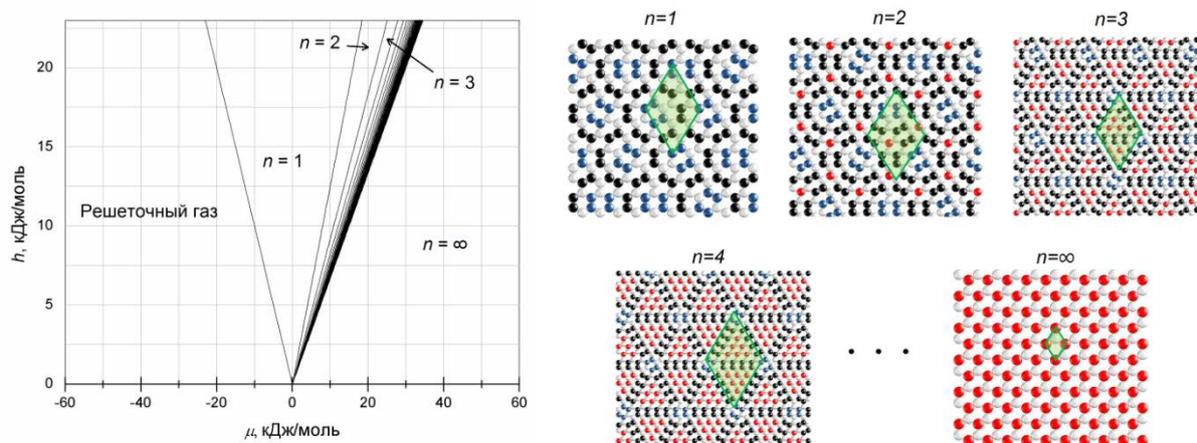


Рис. 3. Фазовая диаграмма в основном состоянии, и изображение нескольких упорядоченных структур, где n – номер упорядоченной структуры

Исследование описанной модели методом Монте-Карло при ненулевых температурах оказывается неэффективным в связи с высокой вырожденностью упорядоченных структур. Для того чтобы определить возможность существования найденных упорядоченных структур на поверхности мы обратились к детерминистским методам, в частности к методу трансфер-матрицы.

2. Метод трансфер-матрицы и мультипликативное разложение

С вычислительной точки зрения основным преимуществом метода трансфер-матрицы перед методом Монте-Карло является отсутствие зависимости от времени релаксации системы, а имеется лишь некоторая скорость сходимости к наперед заданной точности. Однако при ис-

пользовании метода трансфер-матрицы для расчёта равновесных характеристик предложенной модели также возникают существенные вычислительные затруднения. Дело в том, что размер элементарной ячейки каждой последующей фазы (за исключением последней) больше размера предыдущей. В связи с этим требуется изучать систему максимально возможного размера. Если же элементарная ячейка какой-либо фазы больше, чем размеры изучаемой системы, то такая фаза попросту не может образоваться.

Основная идея метода трансфер-матрицы заключается в замене задачи прямого вычисления большой статистической суммы на более простую задачу по вычислению собственного значения некоторой матрицы. Первоначально он был сформулирован для одномерных систем, но впоследствии выяснилось, что он применим и для описания многомерных случаев [3–7].

В двумерном случае метод трансфер-матрицы позволяет найти точное решение для полу-бесконечной системы – бесконечной в одном направлении и имеющей конечную ширину M в другом направлении, где M – количество элементарных ячеек моделируемой решётки активных центров. То есть метод трансфер-матрицы дает точное значение для системы размером $M \times \infty$, которая при периодических граничных условиях представляет собой бесконечный цилиндр, состоящий из колец, каждое из которых состоит из M элементарных ячеек. Элементы матрицы вычисляются по следующей формуле:

$$T_{ij} = \exp \left[\frac{1}{2} u_i + \frac{1}{2} u_j + v_{ij} \right] \quad (1),$$

где $i, j \in [1, p^M]$ – индексы возможных состояний колец, p – количество возможных состояний элементарной ячейки моделируемого адсорбционного слоя, u_i описывает энергию взаимодействий внутри одного кольца, а v_{ij} – взаимодействия между кольцами. Пусть λ_1 – наибольшее по модулю собственное значение матрицы \mathbf{T} , w – количество АЦ в элементарной ячейке ($w = 2$ для гексагональной решётки), а $W = w \times M$ – количество АЦ в одном кольце. Тогда большой термодинамический потенциал, приходящийся на один АЦ, может быть вычислен по формуле:

$$\Omega = -\frac{1}{W} \ln \lambda_1 \quad (2).$$

Таким образом, задача исследования термодинамических характеристик адсорбционной системы сводится к нахождению собственного значения λ_1 . Все алгоритмы поиска собственного значения матрицы требуют многократного умножения этой матрицы на вектор или другую матрицу соответствующего размера. Трансфер-матрица \mathbf{T} является плотной и её размер $p^M \times p^M$ экспоненциально зависит от M , поэтому её умножение при больших M требует много вычислительного времени. Для решения этой проблемы был разработан алгоритм мультипликативного разложения [8].

Трансфер-матрица \mathbf{T} может быть представлена в виде произведения трёх матриц:

$$\mathbf{T} = \mathbf{DAD} \quad (3),$$

где $A_{ij} = \exp(v_{ij})$, $D_{ii} = \exp(\frac{1}{2} u_i)$. Матрица \mathbf{D} является диагональной, поэтому её можно хранить в виде вектора. Матрица \mathbf{A} имеет размер $p^M \times p^M$ и не является разреженной. Поэтому для больших p и M её хранение требует очень много памяти. Алгоритм мультипликативного разложения позволяет представить матрицу \mathbf{A} в виде тензорного произведения матриц \mathbf{t} , существенно меньшего размера $p \times p$:

$$\mathbf{A} = \mathbf{t} \otimes \mathbf{t} \otimes \dots \otimes \mathbf{t} \quad (4).$$

Благодаря такому представлению не требуется хранить в оперативной памяти всю матрицу \mathbf{A} , а достаточно знать \mathbf{t} .

Кроме того, умножение матрицы \mathbf{A} на вектор может быть заменено на M последовательных умножений сильно разреженных матриц. В результате алгоритмическая сложность умножения матрицы \mathbf{A} и, следовательно, матрицы \mathbf{T} на вектор снижается с $O(p^{2M})$ до $O(Mp^{M+1})$.

Таким образом, главным фактором, ограничивающим область применимости метода, является объём оперативной памяти, требуемой для хранения вектора, необходимого для вычисления собственного значения матрицы \mathbf{T} . Дадим оценку объёма оперативной памяти требуемого для моделирования исследуемой адсорбционной системы димеров на гексагональной решётке. Количество возможных состояний элементарной ячейки p равно 8 (пустая ячейка, два способа вертикальной адсорбции и 5 – горизонтальной). Значение M определяет размер элементарной ячейки упорядоченной фазы, которая может быть исследована в ходе моделирования. Особый интерес представляет $M = 12$, поскольку это значение является наименьшим общим кратным для линейных размеров элементарных ячеек первых двух смешанных упорядоченных фаз (3 и 4), кроме того, такая ширина захватывает структуры с размерами элементарных решеток 6 и 12. Во-первых, $M=12$ позволяет исследовать первый фазовый переход между упорядоченными структурами в системе. Во-вторых, позволяет оценить возможность появления в системе многократно вырожденных структур с элементарными ячейками (6×6) и (12×12) . Для вычисления собственного значения трансфер-матрицы \mathbf{T} такой модели потребуется хранить вектор из $p^M = 8^{12} = 2^{36}$ элементов. Если каждый элемент вектора занимает 8 байт, то для его хранения потребуется 512 гигабайт оперативной памяти.

Реализация алгоритма мультипликативного разложения на системах с распределённой памятью не может быть эффективной из-за слабой локализации работы с памятью и вытекающей из этого потребностью часто синхронизировать большие объёмы данных по принципу "каждый-с-каждым". Поэтому, до недавнего времени не существовало практически применимого способа решения данной задачи.

3. Метод трансфер-матрицы для больших систем

Создание высокоскоростных шин межпроцессорного взаимодействия и доступа к общей оперативной памяти, таких как Intel QuickPath Interconnect и AMD HyperTransport, привело к появлению относительно доступных серверов высокопроизводительных вычислений, объединяющих в общем адресном пространстве достаточное количество оперативной памяти и процессорных ядер. С помощью одного из таких серверов, а именно Fujitsu RX900 S2, установленного в Омском государственном техническом университете, нам удалось решить поставленную задачу изучения фазового поведения адсорбционной системы димеров на гексагональной решётке при ненулевых температурах. В данном сервере установлены 8 восьмиядерных процессоров Intel Xeon E7-8800 и 1 терабайт оперативной памяти, объединённые в общем адресном пространстве с помощью технологии QuickPath Interconnect.

Для решения задачи о поиске собственного значения λ_1 мы воспользовались модифицированным методом простых итераций:

$$\mathbf{x}^{(k)} = \mathbf{T} \cdot \mathbf{x}^{(k-1)} \quad (5).$$

Методы с более быстрой сходимостью, такие как метод Арнольди, оказались неприменимы, поскольку требуют хранения нескольких векторов длиной p^M , что в данном случае невозможно из-за ограниченного объёма оперативной памяти. В качестве условия сходимости итераций удобно пользоваться сходимостью по 2-норме:

$$\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_2 < \varepsilon \quad (6).$$

Но такой критерий требует хранения как минимум двух векторов длины p^M , что в данном случае невозможно из-за ограниченного объёма оперативной памяти. Поэтому оценка сходимости выполнялась не по всему вектору \mathbf{x} , а по его проекции в случайное подпространство меньшей размерности p^{M-1} , которое задавалось на каждом шаге итерации путем выбора множества \mathbf{R} , состоящего из p^{M-1} случайных индексов базисных векторов исходного вектора \mathbf{x} . Таким образом, критерий сходимости (6) был преобразован к следующему виду:

$$\|\mathbf{x}^{(k)} - \mathbf{P} \cdot \mathbf{x}^{(k-1)}\|_2 < \varepsilon \quad (7),$$

где \mathbf{P} – диагональная матрица, такая, что $P_{ii} = 1$ если $i \in \mathbf{R}$ и $P_{ii} = 0$ если $i \notin \mathbf{R}$. Иными словами, мы оценивали сходимость последовательности векторов $\mathbf{x}^{(k)}$ по случайно выбранному подмножеству его элементов. Далее будем называть такой критерий сходимости *сходимостью по случайной 2-норме*. Мы не делали теоретической оценки распределения вероятностей отклонений результатов при использовании случайной 2-нормы и обычной 2-нормы, но сравнение результатов вычислений для исследуемой системы при $M < 12$ показали, что применение случайной 2-нормы незначительно увеличивает число итераций, необходимых для сходимости до заданного ε и не оказывает заметного влияния на собственно результаты вычислений. Для хранения множества \mathbf{r} и проекции $\mathbf{P} \mathbf{x}^{(k-1)}$ потребовалось по 64 гигабайта памяти (в p раз меньше, чем для всего вектора \mathbf{x}). Ещё 256 гигабайт потребовалось для хранения диагональной матрицы \mathbf{D} – для хранения её элементов использовались числа с одинарной точностью, занимающие по 4 байта. Всего было задействовано 896 гигабайт оперативной памяти из одного терабайта.

Метод простой итерации имеет плохую сходимость в области фазовых переходов из-за близких собственных значений, соответствующих различным собственным векторам. В комбинации с особенностями моделируемой системы это приводит к резкому увеличению числа итераций, необходимых для достижения требуемой сходимости в области фазовых переходов: приблизительно до 3000 итераций при $\varepsilon = 10^{-10}$.

Для ускорения сходимости нами был использован приём, сообщённый нам Файзуллинским Р. Т., использованный им в работах [9–10]. Суть этого приема заключается во введении некоторой “инерции” в итеративную процедуру. Таким образом, преобразование (5) превращается в:

$$\mathbf{x}^{(k)} = \mathbf{T} \cdot \mathbf{x}^{(k-1)}(1 - \alpha) + \mathbf{P} \cdot \mathbf{x}^{(k-1)}\alpha \quad (8).$$

где $\alpha \in (0, 1)$ – коэффициент инерции.

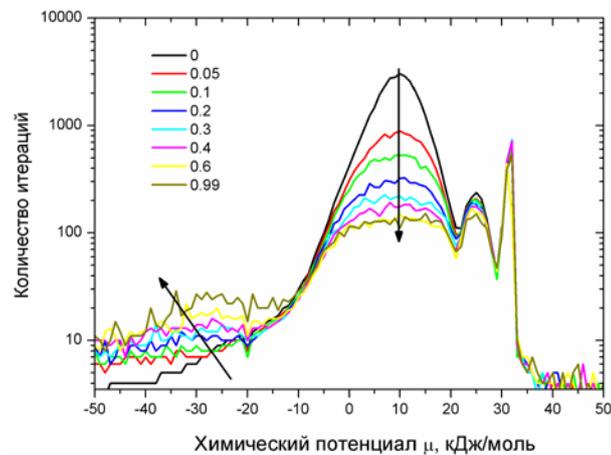


Рис. 4. Зависимость скорости сходимости итераций (8) от коэффициента инерции α на исследованном диапазоне химического потенциала μ при $M = 6$. Стрелки указывают направление увеличения α .

Данный метод можно рассматривать как приближенное разложение собственного вектора по последним двум векторам подпространства Крылова. Такое приближение требует значительно меньше оперативной памяти, по сравнению с другими методами, в частности с методом Арнольди. Влияние коэффициента α на скорость сходимости было изучено для $M = 6$ и представлено на **Рис. 4**. Этот прием позволил сократить максимально требуемое число итераций с 3000 до 200.

При реализации алгоритма нами использовался компилятор C++ из GNU Compiler Collection версии 4.8.2. Для генерации случайных чисел использовалась реализация вихря Мерсенна из библиотеки генераторов случайных чисел Агнера Фогга [11].

4. Результаты

С использованием вышеописанной модификации метода трансфер-матрицы нами была рассчитана изотерма адсорбции $\rho(\mu)$ и изменение энтропии $S(\rho)$ для $M = 12$ и химического потенциала $\mu \in [-50, 31]$ кДж/моль при температуре $T = 300$ К (Рис. 5). По графику $\rho(\mu)$ можно судить об интервалах существования упорядоченных фаз (горизонтальные плато), а по графику $S(\rho)$ – о точках их наибольшей устойчивости (локальные минимумы энтропии).

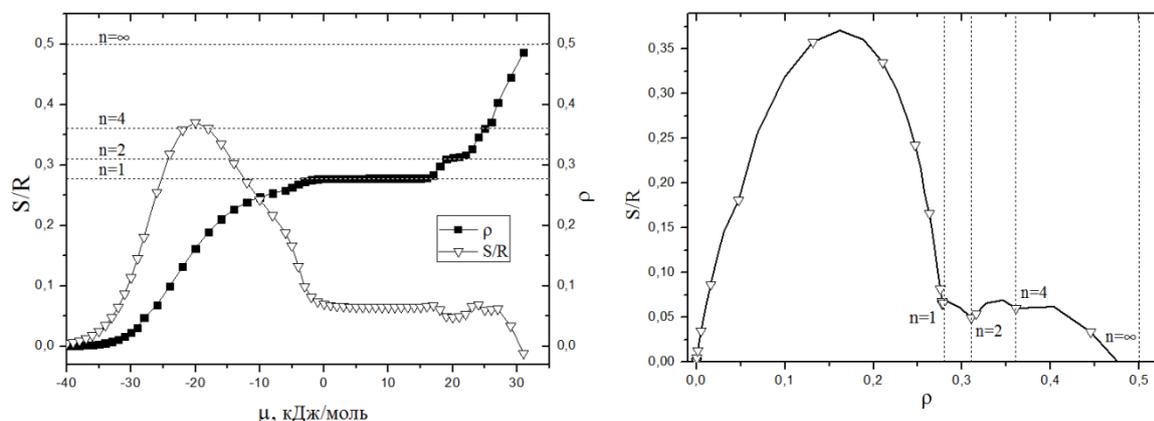


Рис. 5. Изотерма адсорбции и изменение энтропии системы.

Видно, что в исследованном интервале значений химического потенциала в системе образуются 4 упорядоченные фазы: $\rho \approx 0,2(7)$, $\rho \approx 0,31$, $\rho = 0,36$ и $\rho = 0,5$. Приведенные значения указывают на образование фаз $n = 1$, $n = 2$, $n = 4$ и $n = \infty$, обнаруженных в основном состоянии (Рис. 3). Другими словами, при конечных температурах существует, по крайней мере, данная последовательность фаз. По-видимому, отсутствие промежуточных фаз ($n = 3$ и $5 \leq n \leq \infty$) связано с размером исследуемой системы ($M = 12$).

Также следует отметить, что данная система обладает большой вырожденностью упорядоченных фаз и, соответственно, энтропия вносит большой вклад в свободную энергию. Как известно, влияние энтропийного фактора увеличивается линейно с ростом температуры. Поэтому, можно предположить, что количество упорядоченных фаз, образующихся в адсорбционном слое с ростом химического потенциала (концентрации молекул на поверхности твердого тела), будет увеличиваться с понижением температуры. Как следствие, при температурах близких к абсолютному нулю, но отличных от него, в рассматриваемой системе будет наблюдаться явление "чёртовой лестницы" фазовых переходов. Таким образом, вопрос о существовании при конечной температуре "чёртовой лестницы" фазовых переходов в модели адсорбции димеров на сотовой решётке, по-видимому, имеет положительный ответ.

Литература

1. Fefelov V.F., Gorbunov V.A., Myshlyavtsev A.V., Myshlyavtseva M.D. Model of homonuclear dimer adsorption in terms of two possible molecule orientations with respect to surface: Square lattice // Phys. Rev. E. 2010. Vol. 82 P. 041602(1-5).
2. Fefelov V.F., Gorbunov V.A., Myshlyavtsev A.V., Myshlyavtseva M.D. Akimenko S.S. Devil's staircase behavior of a dimer adsorption model // Adsorption 2013 Vol. 19 P. 495.
3. Runnels, L.K., Combs, L.L. Exact Finite Method of Lattice Statistics. I. Square and Triangular Lattice Gases of Hard Molecules // J. Chem. Phys. 1966, Vol. 45, P. 2482- 2492.
4. Rikvold, P.A., Collins, J.B., Hansen, G.D., Gunton, J.D. Three-state lattice gas on a triangular lattice as a model for multicomponent adsorption // Surf. Sci. 1988, Vol. 203, P. 500-524.
5. Myshlyavtsev, A.V., Dongak, M.D. Statistics of adsorption on top and bridge sites of a square lattice: Transfer-matrix approach // J. Stat. Phys. 1997, Vol. 87, P. 593-606.

6. Myshlyavtsev, A.V., Myshlyavtseva, M.D. Modeling of adsorption and phase diagrams for stepped surfaces: Transfer matrix approach // *Appl. Surf. Sci.* 2007, Vol 253, P. 5591-5595.
7. Быков В.И., Мышлявцев А.В., Слинко М.Г. Применение метода трансфер-матрицы для описания процессов на поверхности катализатора // Доклады Академии наук. 2002. Т. 384. С. 650-654.
8. Мышлявцев А.В., Мышлявцева М.Д. Вычислительные аспекты метода трансфер-матрицы – Кызыл: ТувИКОПР СО РАН, 2000 – 101 с.
9. Жихалкина Н.Ф., Логинов К.В., Семин С.Л., Файзуллин Р.Т. Поиск оптимальных режимов работы больших гидросетей и нефтепроводов. – Омск: ОмГУ 1999, 96 с.
10. Логинов К. В., Мызников А. М., Файзуллин Р. Т. Расчет, оптимизация и управление режимами работы больших гидравлических сетей // Матем. моделирование 2006, Т. 18, №.9, сс 92–106.
11. Fog A. Pseudo random number generators.
URL: <http://www.agner.org/random/> (дата обращения: 20.01.2015)