

## Искусственная нормировка при моделировании переноса заряда в молекулярных цепочках

Зайцев А.Ю.<sup>1</sup>, Ольшевец М.М.<sup>1</sup>, Фиалко Н.С.<sup>1</sup>, Яковлев В.В.<sup>1,2</sup>, Лахно В.Д.<sup>1</sup>

<sup>1</sup>*Институт математических проблем биологии РАН – филиал ИПМ РАН им. М.В. Келдыша*

<sup>2</sup>*Московский физико-технический институт (национальный исследовательский университет)*

[fialka@impb.ru](mailto:fialka@impb.ru)

Динамика заряда в цепочке молекулярных сайтов моделируется на основе гамильтониана Холстейна. Температурные флуктуации имитируются ланжевеновским термостатом. Мы рассматриваем модель переноса заряда вдоль фрагментов ДНК. Параметры модели определяются из физических характеристик рассматриваемой системы (масса сайта, энергия электрона на сайте и т.д.). Динамика заряда в ДНК из начального состояния «заряд локализован на сайте-доноре» до термодинамического равновесия может требовать больших расчетных времен. В системе есть сохраняющаяся величина – полная вероятность нахождения заряда в цепочке равна единице. При выполнении вычислительных экспериментов по расчету динамики системы это условие нарушается. В работе обсуждаются результаты тестовых расчетов с искусственной нормировкой, которая проводится при отклонении полной вероятности на величину больше заданной, проведено сравнение для разных значений максимального отклонения.

*Ключевые слова: вычислительный эксперимент, модель Холстейна, уравнение Ланжевена, искусственная нормировка.*

## Artificial normalization in the modeling of charge transfer in molecular chains

Zaytsev A.Y.<sup>1</sup>, Olshevets M.M.<sup>1</sup>, Fialko N.S.<sup>1</sup>, Yakovlev V.V.<sup>1,2</sup>, Lakhno V.D.<sup>1</sup>

<sup>1</sup>*Institute of Mathematical Problems of Biology RAS – the Branch of Keldysh Institute of Applied Mathematics RAS*

<sup>2</sup>*The Moscow Institute of Physics and Technology*

Based on Holstein Hamiltonian, the processes of charge transfer in the chain of molecular sites are modeling. Temperature fluctuations are simulated by the Langevin term in the equations of motion. We consider a charge transfer along DNA fragments. The values of model parameters are determined from the physical characteristics of the system (site mass, electron energy on the site, etc.). The simulations of the charge dynamics in DNA from the initial state «the charge is localized on the donor site» to equilibrium state can be very long, and questions arise about the correctness of the computations. The system has a conserved quantity: the total probability of a charge in the chain is equal to 1. When performing computational experiments on the dynamics of the system, this condition is violated. The paper discusses the results of test calculations with artificial normalization, which is carried out when the deviation of the total probability is greater than the specified value, a comparison is made for different values of the maximum deviation.

*Key words: computational experiment, Holstein model, Langevin thermostat, normalization procedure.*

### 1. Введение

Исследование механизмов переноса заряда в ДНК является актуальной задачей и представляет интерес для биофизики и биохимии [1–4], а также для нанобиоэлектроники [5, 6]. Ряд экспериментов по переносу заряда (электрона или дырки) в ДНК показал разброс относительной скорости переноса

на несколько порядков, в зависимости от типа нуклеотидной последовательности, длины фрагмента, разных молекулярных комплексов, служащих донорами и акцепторами заряда и др. [7–12], см. также обзоры [13, 14] и ссылки в них.

Вычислительные эксперименты, моделирующие процессы переноса заряда в квази-одномерных молекулярных цепочках, являются одним из инструментов изучения механизмов переноса.

Мы исследуем модель переноса заряда в цепочке из  $N$  сайтов (в случае ДНК сайтом считается комплементарная пара оснований), основанную на гамильтониане Холстейна в полуклассическом приближении. Уравнения движения в безразмерной форме имеют вид

$$i\dot{b}_n = \eta_{n,n-1}b_{n-1} + \eta_{n,n}b_n + \eta_{n,n+1}b_{n+1} + \chi u_n b_n, \quad (1)$$

$$\ddot{u}_n = -\omega^2 u_n - \chi |b_n|^2 + \gamma \dot{u}_n + \xi Z_n(t), \quad n = 1, \dots, N. \quad (2)$$

Здесь переменные  $b_n$  (амплитуда вероятности нахождения заряда на  $n$ -ом сайте) и  $u_n$  (смещение сайта из равновесного положения). Параметры  $\eta_{nk}$ ,  $\chi$ ,  $\omega$ ,  $\gamma$ ,  $\xi$  – постоянные, оценки их значений взяты из литературных данных [15, 16].  $Z_n$  – случайная величина со специальными свойствами  $\langle Z_n(t) \rangle = 0$ ,  $\langle Z_n(t)Z_m(t+t') \rangle = \delta_{nm} \delta(t')$ . Множитель  $\xi$  при случайной силе учитывает температуру термостата.

Вычислительный эксперимент при заданной температуре  $T$  сводится к численному интегрированию множества реализаций (траекторий) системы (1),(2) из разных начальных данных и с разными временными псевдо-случайными последовательностями) и расчету средних по выборке временных зависимостей.

При моделировании динамики в ряде задач необходимо интегрирование на больших временных интервалах. Для системы (1,2) на «идеально рассчитанной» траектории должно всегда выполняться условие:  $S = \sum_{n=1}^N b_n b_n^* \equiv 1$ . При расчетах это условие со временем нарушается. В расчетные программы для явного 2o2s1g-метода [17] добавлена искусственная нормировка: если для заданного значения  $\varepsilon$  на каком-то шаге

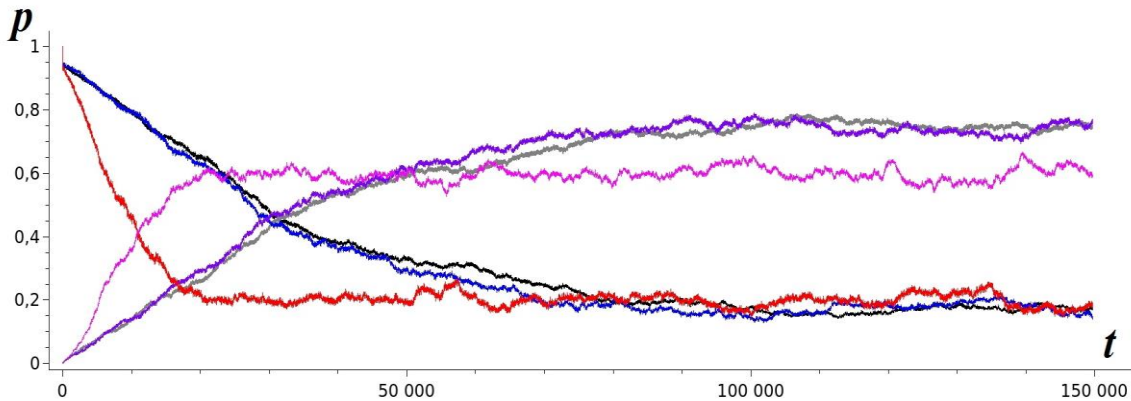
$$|S - 1| > \varepsilon, \quad (3)$$

то рассчитанные значения  $b_n$  заменяются на  $b_n / \sqrt{S}$ , и дальше расчет продолжается из этих новых данных. Ранее для этого метода был определен максимальный шаг интегрирования [18] при значениях параметров неоднородных фрагментов ДНК. В данной работе приведены результаты тестов для динамики системы (1,2) на больших расчетных временах при разной величине  $\varepsilon$ .

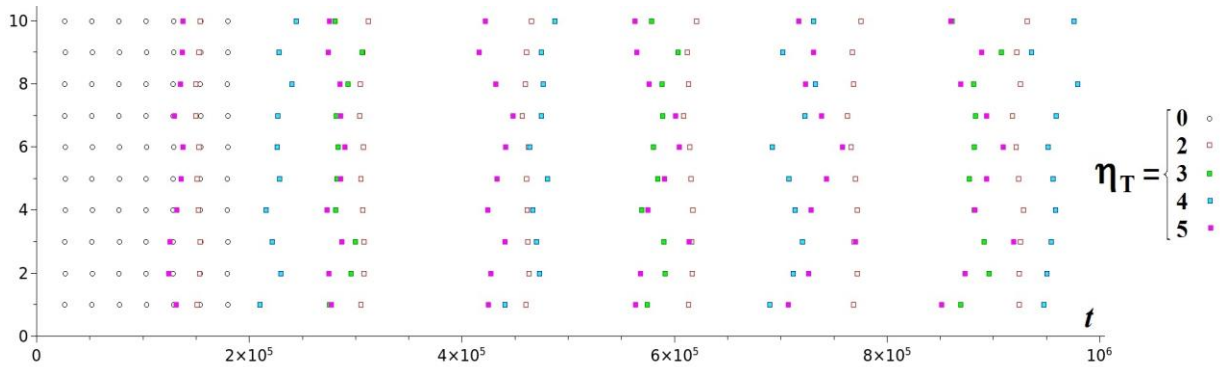
## 2. Нормировка на каждом шаге

Выполнение нормировки на каждом шаге (т.е. условие  $\varepsilon = 0$ ), хотя и сохраняет  $S = 1$ , но не позволяет значительно укрупнить шаг. На рисунке 1 приведены результаты расчетов 6-сайтовой модели GAAGGG, где в начальный момент заряд локализован на первом G:  $b_1(t=0) = 1$ . Интегрирование с большим шагом  $h = 0.005$ ,  $\varepsilon = 0$ , (красный график изображает вероятность  $p_1(t)$  на первом G) приводит к более быстрому времени перехода с G на GGG, чем результат для меньших шагов интегрирования (черная и синяя кривые), и к неправильному распределению вероятностей после перехода (когда графики колеблются вокруг среднего). Суммарная вероятность на сайтах GGG (розовая кривая) меньше, чем рассчитанная с шагом  $h = 0.0005$ ,  $\varepsilon = 0$  (фиолетовая кривая) и  $h = 0.0002$ ,  $\varepsilon = 0.0005$  (серая).

Данный пример иллюстрирует опасность искусственной нормировки на каждом шаге. Большое количество проведенных вычислений подсказывает, что для  $\varepsilon \neq 0$  в отдельных реализациях временные интервалы между нормировками примерно одинаковы; если интервалы сильно различаются – это повод для проверки рассчитанной траектории и, скорее всего, для уменьшения шага интегрирования.



**Рис. 1.** Динамика вероятностей нахождения заряда  $p_n = b_n b_n^*$  в цепочке GAAGGG при расчетах с шагом  $h = 0.005$ ,  $\varepsilon = 0$  (красная, розовая кривые),  $h = 0.0005$ ,  $\varepsilon = 0$  (синяя, фиолетовая кривые) и  $h = 0.0002$ ,  $\varepsilon = 0.0005$  (черная, серая). Приведены графики для первого G, на котором  $p_1(t=0) = 1$  (красный, синий, черный графики), и суммарная вероятность нахождения заряда на GGG  $p_4(t) + p_5(t) + p_6(t)$  (розовый, фиолетовый, серый).  $T = 300$  K, каждая кривая – среднее по 50 реализаций.



**Рис. 2.** Однородная цепочка 30 сайтов, параметры polyT-фрагмента ( $\eta_{m+1} = 2.4$ ,  $\chi = 1$ ,  $\omega = 0.5$ ,  $\gamma = 4\omega$ ), малая температура  $T = 10$ ,  $h = 0.0005$ ,  $\varepsilon = 0.0005$ . Расчеты для разных значений диагональных элементов  $\eta_{mm} = \eta_T$  (справа указаны значения  $\eta_{mm}$ ). Для 10 случайных реализаций приведены времена, на которых выполняется искусственная нормировка.

### 3. Нормировка в одной реализации

Сравнение двух траекторий системы (1,2) для разных шагов  $h$  неинформативно, т.к. случайная сила генерируется на каждом шаге интегрирования. Подходящая величина  $h$  определяется сравнением осредненных по реализациям траекторий [18]. При фиксированном  $h$  мы провели расчеты отдельной траектории (из одинаковых начальных данных и с одинаковым временным набором случайных величин) для разных значений  $\varepsilon$ . Тесты проводились для  $h = 5 \cdot 10^{-4}$  и  $2 \cdot 10^{-4}$ , для каждого шага по 4 реализации, время счета  $10^6$ , рассматривались разные  $\varepsilon$  от  $10^{-16}$  (решение в 6 цифрах совпадает с  $\varepsilon = 0$ ) до 0.01 (в этом случае условие (3) на всей длине траектории не выполнялось, т.е. искусственных нормировок не было). Начальные данные – полярон плюс малые возмущения по смещениям  $u_n$  и скоростям сайтов (которые соответствуют  $T = 10$ ). Для малых  $\varepsilon$  траектории не расходятся на сравнительно больших временах; при  $h = 2 \cdot 10^{-4}$  для решений с  $\varepsilon = 10^{-16}$  и  $\varepsilon = 10^{-15}$  разница в вероятностях  $p_n(t)$  не превышает  $2 \cdot 10^{-6}$  на временах до 10000. Однако требуемые времена счета гораздо больше. Для «больших»  $\varepsilon > 10^{-5}$  расхождение может быть значительным после первых двух нормировок. При этом средние по 100 реализациям для  $\varepsilon < 10^{-3}$  близки между собой (сравнивалось время разрушения начального полярона и максимум вероятности в равновесном состоянии).

### 4. Частота нормировок и диагональный матричный элемент $\eta_{mm}$

Диагональный матричный элемент  $\eta_{mm}$  определяется энергией заряда на сайте. С физической точки зрения, важна не сама величина  $\eta_{mm}$ , а разность между ними на разных сайтах. Поэтому, например, при моделировании переноса заряда в однородных цепочках полагают  $\eta_{mm} = 0$ .

С точки зрения вычислителя, эти значения можно менять (с сохранением разности  $\eta_{mm}$ ). Если задаться целью уменьшить при расчете количество нормировок (т. е. уходов фазовой точки  $(b_1, \dots, b_N)$  со сферы  $\sum_{n=1}^N b_n b_n^* = 1$  на расстояние больше фиксированного  $\varepsilon$ ), то тесты показывают, что  $\eta_{mm} = 0$  не всегда оптимально для однородных цепочек. На рисунке 2 отмечены моменты времени, в которые выполнялась искусственная нормировка, для одинаковых  $h = 0.0005$ ,  $\varepsilon = 0.0005$  и разных значений  $\eta_{mm}$ . Отметим, что для средних по 50 реализациям траекторий очевидных расхождений (как, например, на рис. 1) не найдено. Результаты для  $\eta_{mm} = 0$  «обрезаны» на времени  $t = 2 \cdot 10^5$ , для лучшей визуализации остальных результатов, частота нормировок примерно одинакова на всем времени счета  $t = 10^6$ .

Из рисунка 2 следует, что наименьшее количество искусственных нормировок – 3 раза за расчетный интервал – происходит при  $\eta_{mm} = 3$ . Для  $\eta_{mm} = 4$  выполнено 4 нормировки, для  $\eta_{mm} = 2$  и  $\eta_{mm} = 5$  – по 6. Для  $\eta_{mm} = 8$  количество нормировок (30) близко к числу нормировок (35) для  $\eta_{mm} = 0$ . Т. е. для количества искусственных нормировок при некотором  $\eta_{mm} \sim 3$  есть минимум.

Для неоднородных цепочек ситуация сходная. Мы тестировали цепочки вида GT...TGGG, в которых между сайтами разного вида  $\Delta\eta_{mm} = 10$ . Предположение, что наименьшее количество нормировок будет при  $\eta_G = -5$ ,  $\eta_T = +5$ , т. е. величины параметров модели наименьшие по модулю, не подтвердилось. По результатам тестов, наименьшее количество нормировок при  $\eta_G$  от  $-2$  до  $-1$  (соответственно,  $\eta_T$  от 8 до 9).

### 5. Заключение

В работе обсуждаются технические вопросы, связанные с искусственной нормировкой, т.е. с внешней корректировкой траектории при расчетах

явным методом на больших временных интервалах в системе со случайной правой частью.

Показано, что при большом шаге интегрирования процедура нормировки приводит к неправильной траектории.

Для «подходящего» шага нормировка с величиной  $\varepsilon$  около машинной точности сохраняет отдельную траекторию на временах  $\sim 10^7$  шагов. Однако расчеты динамики требуется проводить на два-три порядка дольше. Показано, что средние по 100 реализациям траектории для  $\varepsilon < 10^{-3}$  близки между собой.

Показано, что можно получить уменьшение количества таких корректировок на расчетных временах при фиксированном  $\varepsilon$  при изменении диагональных элементов  $\eta_{nn}$  с сохранением разности между ними в случае неоднородных цепочек.

## 6. Благодарности

Расчеты проводились на суперкомпьютерах k-60 и k-100 ИПМ им. М.В.Келдыша РАН.

Работа выполнена при частичной поддержке Российского научного фонда, проект 16-11-10163.

## 6. Список литературы

1. Dimple B., Harrison L. Repair of oxidative damage to DNA: Enzymology and biology. *Annu. Rev. Biochem.* 1994. V. 63. P. 915–948. doi: [10.1146/annurev.bi.63.070194.004411](https://doi.org/10.1146/annurev.bi.63.070194.004411).
2. Loft S., Poulsen H.E. Cancer risk and oxidative DNA damage in man. *J. Mol. Med.* 1996. V. 74. № 6. P. 297–312. doi: [10.1007/BF00207507](https://doi.org/10.1007/BF00207507).
3. David S.S., Williams S.D. Chemistry of glycosylases and endonucleases involved in base-excision repair. *Chem. Rev.* 1998. V. 98. № 3. P. 1221–1262. doi: [10.1021/cr980321h](https://doi.org/10.1021/cr980321h).
4. *Charge Migration in DNA. Perspectives from Physics, Chemistry, and Biology.* Ed. Chakraborty T. Berlin Heidelberg New York: Springer, 2007. 288 p. (Series: NanoScience and Technology. XVIII).
5. Lakhno V.D. DNA nanobioelectronics. *Int. J. Quant. Chem.* 2008. V. 108. № 11. P. 1970–1981. doi: [10.1002/qua.21717](https://doi.org/10.1002/qua.21717).
6. *Nanobioelectronics – for Electronics, Biology, and Medicine.* Eds. Offenhausser A., Rinaldi R. New York: Springer, 2009. 337 p.
7. Bran A.M., Harriman A.J. Dynamics of Electron transfer between intercalated polycyclic molecules: effects of interspersed bases. *J. Am. Chem. Soc.* 1992. V. 114. № 10. P. 3656–3660. doi: [10.1021/ja00036a013](https://doi.org/10.1021/ja00036a013).
8. Murphy C.J., Arkin M.R., Jenkins Y., Ghatlia N.D., Bossman S., Turro N.J., Barton J.K. Long-range photoinduced electron transfer through a DNA helix. *Science.* 1993. V. 262. № 5136. P. 1025–1029. doi: [10.1126/science.7802858](https://doi.org/10.1126/science.7802858).

9. Meade T.J., Kayyem J.F. Electron transfer through DNA site-specific modification of duplex DNA with ruthenium donors and acceptors. *Angew. Chem. Int. Ed. Engl.* 1995. V. 34. № 3. P. 352–354. doi: [10.1002/anie.199503521](https://doi.org/10.1002/anie.199503521).
10. Meggers E., Michel-Beyerle M.E., Giese B. Sequence Dependent Long Range Hole Transport in DNA. *J. Am. Chem. Soc.* 1998. V. 120. № 49. P. 12950–12955. doi: [10.1021/ja983092p](https://doi.org/10.1021/ja983092p).
11. Giese B., Amaudrut J., Koehler A.K., Spormann M., Wessely S. Direct observation of hole transfer through DNA by hopping between adenine bases and by tunnelling. *Nature.* 2001. V. 412. № 6844. P. 318–320. doi: [10.1038/35085542](https://doi.org/10.1038/35085542).
12. Lewis F.D., Wu Ya. Dynamics of superexchange photoinduced electron transfer in duplex DNA. *J. Photochem. Photobiol. C.* 2001. V. 2. №1. P. 1–16. doi: [10.1016/S1389-5567\(01\)00008-9](https://doi.org/10.1016/S1389-5567(01)00008-9).
13. Genereux J.C., Barton J.K. Mechanisms for DNA Charge Transport. *Chem. Rev.* 2010. V. 110. № 3. P. 1642–1662. doi: [10.1021/cr900228f](https://doi.org/10.1021/cr900228f).
14. Zhang Y.Q., Zhang W.B., Liu C.R., Zhang P., Balaeff A., Beratan D.N. DNA charge transport: Moving beyond 1D. *Surface Science.* 2016. V. 652. P. 33–38. doi: [10.1016/j.susc.2016.03.011](https://doi.org/10.1016/j.susc.2016.03.011).
15. Voityuk A.A., Rosch N., Bixon M., Jortner J. Electronic Coupling for Charge Transfer and Transport in DNA. *J. Phys. Chem. B.* 2000. V.104. № 41. P. 9740–9745. doi: [10.1021/jp001109w](https://doi.org/10.1021/jp001109w).
16. Jortner J., Bixon M., Voityuk A.A., Roesch N. Superexchange Mediated Charge Hopping in DNA. *J. Phys. Chem. A.* 2002. V. 106. № 33. P. 7599–7606. doi: [10.1021/jp014232b](https://doi.org/10.1021/jp014232b).
17. Greenside H.S., Helfand E. Numerical integration of stochastic differential equations - II. *Bell Syst. Tech. J.* 1981. V. 60. P. 1927–1940. doi: [10.1002/j.1538-7305.1981.tb00303.x](https://doi.org/10.1002/j.1538-7305.1981.tb00303.x).
18. Пятков М.И., Зайцев А.Ю., Ольшевец М.М., Фиалко Н.С. Поиск оптимального шага интегрирования в задаче моделирования динамики заряда в гетерогенных ДНК. Доклады Международной конференции “Математическая биология и биоинформатика”. Под ред. В.Д. Лахно. Том 7. Пушино: ИМПБ РАН, 2018. Статья No e11. doi: [10.17537/icmbb18.65](https://doi.org/10.17537/icmbb18.65).