Linear Random Knots and Their Scaling Behavior

Kenneth Millett,[†] Akos Dobay,^{‡,§} and Andrzej Stasiak^{*,‡}

Department of Mathematics, University of California, Santa Barbara, California 93106, Laboratoire d'Analyse Ultrastructurale, Bâtiment de Biologie, Université de Lausanne, 1015 Dorigny, Switzerland, and Theoretische Physik, Fakultät für Physik, Ludwig-Maximilians-Universität, 80333 München, Germany

Received June 20, 2004; Revised Manuscript Received October 21, 2004

ABSTRACT: We present here a nonbiased probabilistic method that allows us to consistently analyze knottedness of linear random walks with up to several hundred noncorrelated steps. The method consists of analyzing the spectrum of knots formed by multiple closures of the same open walk through random points on a sphere enclosing the walk. Knottedness of individual "frozen" configurations of linear chains is therefore defined by a characteristic spectrum of realizable knots. We show that in the great majority of cases this method clearly defines the dominant knot type of a walk, i.e., the strongest component of the spectrum. In such cases, direct end-to-end closure creates a knot that usually coincides with the knot type that dominates the random closure spectrum. Interestingly, in a very small proportion of linear random walks, the knot type is not clearly defined. Such walks can be considered as residing in a border zone of the configuration space of two or more knot types. We also characterize the scaling behavior of linear random knots.

Knots form on linear polymers and affect their physical behavior.¹ However, the determination of the knot type tied on an open polymer is not straightforward since in a strict sense the topology of open strings is not defined as every open string can be always untied by a simple continuous transformation.² For this reason various methods of determining and characterizing knots on open strings were proposed.³⁻⁷ An interesting approach was applied to determine knots formed on polypeptide chains in properly folded protein structures⁸ that therefore have the topological state of the chain "fixed".⁹ The latter method, based on a variant of progressive smoothing of analyzed polygonal trajectories,¹⁰ is very attractive, but even this approach can classify the same starting trajectory as knotted or unknotted depending simply on the order of smoothing operations (see Figure 1).

Therefore, one needs to accept the concept of uncertainty of knotted state of open strings and operate with probabilistic definition of knottedness.⁷ Accepting this concept, we systematically investigate the knottedness of ideal random walks of increasing size. Figure 2 shows the principle of our method of analyzing the knottedness of random walks. The center of gravity of analyzed open walk is placed in the center of a large sphere that encloses the walk. Subsequently, a random point chosen on the sphere is connected with straight segments to both ends of the analyzed random walk. (For practical reasons we have set the radius of the sphere to be twice larger than the total length of the analyzed walk.) This operation closes the knot and makes it amenable to standard topological analysis of the formed knot such as identification of the knot type based on calculation of HOMFLY polynomial.¹¹ For each analyzed configuration we have tried 10 000 random closure points on the surface of the sphere. Such an analysis satisfies the Copernican principle, as we do not favor any direction

[‡] Université de Lausanne.

Figure 1. Chain-smoothing algorithm proposed by Taylor⁸ can result in classifying the same initial polygonal trajectory as knotted or unknotted depending on the order of smoothing operations. When smoothing proceeds from vertex 1 to 8, the vertex 5 finds itself "above" the edge 7–8, producing unknot. When smoothing proceeds from vertex 8 to 1, the vertex 5 finds itself "below" the edge 7–8, resulting in a trefoil knot. The coordinates of the shown configuration are available on request.

of closure. Walks studied by us are so-called ideal random walks where all segments are of the same length and have no thickness. To generate a random walk, we first created a set of unit vectors randomly sampling the surface of a unit sphere. Subsequently, we use this set to construct a chain by sequential joining of randomly chosen vectors.¹² We repeated the procedure for each of the analyzed chains. This elementary method of chain generation samples the configuration space without any bias, and the analyzed configurations are not related to each other. Results of the closure analysis for an individual random walk of 300 segments are presented in Figure 3. 10 000 directions of closures produced a spectrum of knots which is presented in a form of a histogram. The analyzed walk can be therefore described as being a superposition state composed in 91.7% unknot, 4.1% trefoil, and 4.2% other knots where a given direction of closure "collapses" the superposition state and determines the resulting knot type.

[†] University of California, Santa Barbara.

[§] Ludwig-Maximilians-Universität.



Figure 2. A nonbiased method of closure of random linear walks. A randomly chosen point on a sphere that encloses a given walk is connected with straight segments to both ends of the analyzed linear walk (the connecting segments are much longer that the segments constituting the analyzed random walk). The chain closure forms a knot which type is determined by calculation of HOMFLY polynomial. The whole procedure is repeated for many other randomly chosen points on a sphere (10 000 of random closures for each analyzed trajectory were applied in this work). In this schematic drawing, the diameter of the enclosing sphere was greatly reduced for pratical reasons and therefore does not reflect the real dimensions of the analyzed simulated chain as compared to the enclosing sphere.

Figure 4a presents the analysis of closure for a statistical sample composed of 100 different random walks with 48 segments each. The individual histograms of knotting spectra are presented in the orientation shown in Figure 3, placed side-by-side and ordered according to increasing level of knottedness revealed by a given random walk. For 100 random walks with 48 edges (see Figure 4a) only three walks showed always the same knot type irrespectively of the closing point (all were unknots). Other analyzed configurations can change their knot type depending on the direction of closure; however, 96 configurations (out of 100) showed a high certainty of knot type as over 90% of closing directions resulted in the same knot type. It is visible that the great majority of configurations show high certainty of the revealed knot type (more than 90% of closing directions reveal the same knot type), and this applies not only to configurations classified as trivial knots but also to trefoils and more complex knots. Only one configuration showed a lower certainty than 50%. Interestingly, the uncertainty of the knot type resulted mainly in appearance of one secondary knot type with significant occurrence, but in a few cases two or more secondary knot types were observed. Figure 4b,c shows the analysis of knotting certainty for 150 and 300 segments long random walks. It is visible that even in the case of these long chains the great majority of walks (90 and 82, respectively) show a high certainty of the knot type as determined by the fact that over 90% of

closing directions resulted in formation of the same knot type. As the chain size of analyzed walks increases, the proportion of unknots decreases as expected from wellknown Delbrück's conjecture¹³ and analytical proofs.^{3,14,15} However, this observation hardly affects the fact that great majority of random linear walks with this chain size still have a well-defined knot type. Nevertheless, it is to be expected that for much longer random walks the dominant knot type will show a lower dominance over the secondary and tertiary knot types produced through different directions of closure. We expect however that even for very long chains a spectrum of knots resulting from random closure of a given trajectory will provide a good signature characterizing the knottedness of a given configuration. Analysis of the relation between dominant and secondary knot types revealed that they are neighbors in configuration spaces of knots, i.e., represent knots that can be converted into each other by one strand passage reaction.¹⁶ Thus, for example, secondary knots associated with unknots consisted mainly of 31, 41, and 52 knots, and these can be obtained by one strand passage from unknots (see Figure 6 for schematic drawings of these knots). 51 knots were very rarely observed as secondary knots when unknots were dominant since there are two passages required to form a 51 knot from an unknot.¹⁶ Similarly secondary knots to dominant trefoils very rarely included 41 knots as there are two strand passages required to convert these knots into each other. In general, we can say that



Figure 3. Statistics of closure for a 300 segments random walk. The knotting spectrum of the walk reveals the dominant knot type (unknot) and the secondary knot types. The nomenclature of knots follows this in standard tables of knots where the main number indicates the minimal number of crossings of a given knot and the subscript number indicates the knot tabular position among the knots with the same number of crossings.² Under notation 7+ we have grouped prime knots are grouped as "composite". Notice that each panel in Figure 4 consists of 100 individual 3-D knotting histograms presented in the orientation corresponding to this shown here.

configurations that show significant level of secondary knot types seem to occupy a border zone between configuration spaces of two or more knot types.

The data presented in Figure 4 when averaged over 100 individual configurations analyzed for each chain size can be conveniently used to measure the average knottedness of random walks of a given chain size. This measure expresses the time averaged knottedness of freely fluctuating linear polymers composed of corresponding number of statistical segments and analyzed under so-called θ conditions in which independent segments neither attract nor repulse each other.¹⁷ Figure 5 presents the averaged probabilities with which linear random walks of 48, 150, and 300 segments form various knot types upon a random closure. Notice that the measured probabilities show similar distributions to these determined earlier in the case of random knotting of circular random chains.^{18,19} So for example in the case of 300 segment long circular random walks the trefoils and 4_1 knots formed with the probabilities of ca. 23% and 5%, respectively.¹⁸ In the case of linear random walks of 300 segments their random closure resulted in generation of trefoils with ca. 24% while 4_1 knots formed with a probability of ca. 5% (see Figure 5). The agreement between the classical knotting probability of circular random walks with the knotting probability for linear random walks demonstrates that linear random walks are just as likely to be knotted as the circular random walks provided that the knottedness of linear random walks is adequately assessed.

Since the analysis of 10 000 random closures for each trajectory is computationally demanding, we decided to test whether the straight segment closure connecting both ends of linear random walks captures the dominant knot type detected by our method of multiple random closures. Direct end-to-end closures were applied to the



Figure 4. Random closure analysis of random linear walks of increasing size. Each panel presents statistics of closure for a sample of 100 random walks of 48 (a), 150 (b), and 300 (c) segments, where statistics of each analyzed configuration is presented as individual histogram placed along the *x*-coordinate. The random configurations are ordered according to the dominant knot type revealed in them and then within each dominant knot type according to the observed certainty level (vertical axis).

300 configurations analyzed in Figure 4, and we observed that, with exception of 2 cases of 48-segment walks, 3 cases of 150-segment walks, and 10 cases of 300-segment walks, the knot type detected by direct end-to-end closure coincided with the dominant knot type detected by the random closure method. As expected, the exceptions occurred more frequently among these configurations that by the random closure method showed a low certainty level of their knot type. Therefore, we conclude that, for most practical purposes in chains with up to several hundred segments, the direct closure method is very likely to capture the dominant knot type in a given open random walk. We decided, therefore, to apply the direct closure method to distinguish individual knot types from a large statistical ensemble (100 000 random configurations for each of the 42 chain sizes analyzed in Figure 6) of linear random walks with up to 1008 segments. We investigated how overall dimensions of random walks, such as meansquare radius of gyration, are affected by the knot type detected in individual configurations of open random walks.

Figure 6 shows how the mean-square radius of gyration of open walks that revealed itself as different



Figure 5. Averaged probabilities with which linear random walks composed of 48 (a), 150 (b), and 300 (c) segments reveal itself as various knots upon a random closure.

knot types scales with the number of segments in analyzed random walks. Notice that the process of closure does not affect the radius of gyration as the calculation is performed for configurations without the joining segment. Walks identified as trivial knots (unknots) showed highest overall dimensions while open walks forming different knot types decreased their overall dimensions with the complexity of the knot. This result is analogous to the effect of different knots on circular random walks where it was observed that when tied on chains with the same lengths more complex knots are more compact than simpler knots.^{12,20,21} Previous analytical and numerical studies of knotted random walks revealed that when non-self-avoiding random walks are divided into individual knot types, each individual knot type should show the scaling behavior expected for self-avoiding walks.²²⁻²⁶ Therefore, while overall dimensions averaged over all of nonself-avoiding random walks (as these generated in this study) scaled with the number of segments N as N^{ν} with v = 0.5, the overall dimensions of walks representing a given knot type scaled with n exponent tending to 0.588.²² As shown in Figure 5, the mean-square radius of gyration of all random open walks shows the expected linear growth rate $(2\nu = 1)$. The individual knot types show, however, a higher growth rate. To determine the ν exponent of scaling behavior of open walks classified according to their individual knot types, we fitted the

experimental points using a formula

$$\langle R_{g}^{2}(N) \rangle = AN^{2\nu}(1 + BN^{-\Delta} + CN^{-1} + o(N-1))$$
 (1)

where N is the number of segments, while A, B, C, and ν are free parameters and where Δ is set to 0.5.²⁷ For unknot $A = 0.051 \pm 0.0003$, $B = 5.9 \pm 0.3$, $C = -7.2 \pm$ 2.9, and for trefoil $A = 0.052 \pm 0.0003$, $B = 4.4 \pm 0.3$, C $= -16.3 \pm 2.5$. We observed that for the data with best statistics, i.e., unknots and trefoils, the ν exponent adopted values of 0.599 \pm 0.019 and 0.586 \pm 0.013, respectively. Since these values were within the error range from the scaling exponent of self-avoiding walks, we repeated the fit but this time setting the ν exponent to 0.588. As shown in Figure 5, these fits were excellent in all cases with sufficient statistics. (We combined 5_1 and 52 knots together to decrease statistical fluctuations due to relatively small sample sizes of these five crossing knots.) This excellent fit supports our proposal that open knots can behave analogously to classical closed knots. Formula 1 serves to determine the value of scaling exponent ν for very long chains (not accessible to numerical simulations) on the basis of the data obtained for relatively short chains that could be analyzed by numerical simulations. In fact, the effective scaling exponent v analyzed over a small "sliding window of 200 segments" grows progressively from 0.5 to 0.55 as the analyzed chain size increases from 5 to 800 segments (analysis not shown) and as the frequency of random knotting increases. Since the parameter B is the main correcting term, the B^2 value for unknots can be interpreted as a characteristic chain size where the frequency of knots becomes significant and where the scaling profile changes its character from this typical for ideal random walks to this typical for self-avoiding walks. Chains with 36 segments are knotted with the frequency higher than 5%, which is a significant value.

The applicability of the fitting formula 1 was sometimes questioned²⁷ as it was believed that it may bias the fitting parameters in such a way that the ν exponent is "predetermined" to be close to 0.588, i.e., characteristic for self-avoiding walks. Concerned with this reservation, we have used the same fitting formula and the same value of the confluent exponent $\Delta = 0.5$ to find the ν exponent characterizing the scaling of all walks analyzed in Figure 5. The fitted value of the ν exponent for all walks grouped together irrespectively of their topology was very close to 0.5 (0.505 \pm 0.0002), and it was therefore characteristic for non-self-avoiding walks. This demonstrated that the fitting formula 1 does not impose any bias on the outcome of the ν exponent.

Looking at small differences between $\langle R_g^2 \rangle$ of unknotted chains and these with simple knots, one can conclude that knots are rather localized than being spread over the whole chain. Spreading of a knot would significantly diminish $\langle R_g^2 \rangle$ since segments that are separated by a long distance along a chain would need to approach each other. In localized knots only the segments that were separated by a short distance along the chain need to come close together, and this hardly affects overall dimensions of long chains. Localization of simple knots has been proposed earlier in the case of circular chains.^{4,25,28-30} More recently, however, the question became more precise: is the mean number of segments of the knotted domain independent of the total length of the chain or, rather, do knotted domains grow with the size of a chain as suggested by numerical



Figure 6. Scaling of $\langle R_g^2 \rangle$ of linear non-self-avoiding random knots with respect to the chain length. The scaling profile for all linear walks grouped together is drawn with the thicker lines, and it clearly follows a linear relation pointing to scaling exponent ν of 0.5 while scaling profiles of linear walks classified as trivial, 3_1 , 4_1 or 5_1 and 5_2 knots show higher growth rates characteristic for self-avoiding walks. Fitting the experimental data for open knots assuming the scaling exponent ν of 0.588 gave an excellent quality of fit in each analyzed case: R = 0.999 99 for unknots and 3_1 knots, R = 0.999 93 for 4_1 knots and R = 0.999 95 for 5_1 and 5_2 knots grouped together. Statistical data were obtained for 100 000 random open walks of each analyzed chain size. Schematic ν is set to 0.588 give 0.062 \pm 0.0002, 0.059 \pm 0.0001, and 0.056 \pm 0.0001 for trivial, 3_1 , 4_1 , $5_1 + 5_2$, respectively. Notice that for long walks simple knots constitute only a small fraction of the walks while remaining walks would reveal itself upon closure as complex knots with substantially decreased $\langle R_g^2 \rangle$ as compared to simple walks analyzed separatly in Figure 6.

simulations of knotted polymers?³¹ A closer look at Figure 6 and calculated prefactors of simple power law functions for individual knot types (see legend of Figure 6) reveals that the scaling profiles for linear walks with different simple knots diverge from each other. This result suggest that in fact mean number of segments of knotted domains increases with the length of the chain, and the more complex the knot the stronger is this tendency. The increase of the mean number of segments of the knotted domain with the length of the chain can be explained by the presence of spread knots. Even if such spread knots are rare, the "spread" of the knotted domain (in number of segments) can increase with increasing chain length. Thus, even if the most frequent number of segments observed for a knotted domain of a given type is independent of the size of the chain,⁴ the mean number of segments of knotted domains should increase with the length of the chain.

Random walks are good models of polymers behavior in solution such as polyethylene or DNA.32-37 Our results demonstrate, therefore, that populations of randomly fluctuating linear polymers can be conceptually divided into configurations that momentarily reside in a given knot space. However, knot spaces for linear chains do not have strict borders as is the case for classical closed chains.¹⁶ Until now a polymer closure (like circularization of double- and single-stranded DNA) was required to investigate how frequently momentarily configurations of polymers with a given size visit configuration spaces of different knot types.³⁸⁻⁴¹ Our approach that considers momentary configuration as "frozen" allows us to perform such an investigation without polymer closure. We showed here that it makes sense to characterize knots in momentary configurations of fluctuating polymer in solution provided that we are able to "photograph" such momentary configurations and analyze them as frozen embeddings. Of course, the "photographed" trajectory itself may rapidly change

knot type upon further fluctuation. Analyzing the knottedness of open trajectories, we need however to accept a certain degree of uncertainty of knot type. In some particular cases two or more knot types can be formed with almost equal probabilities by random closure. Such configurations are interesting as they live in a border zone between two or more configuration spaces of different knot types of linear chains. We demonstrated here that, for most of practical purposes, a simple closure with a straight segment connecting both ends captures the dominant knot type detected by the "Copernican" method of random closures. We showed also that open random knots share some similar characteristics with classical closed random knots. This includes relative ordering of overall sizes of different knot types and the same scaling exponent when one analyses how overall dimensions change with the chain length in open and closed knots of a given knot type.

Acknowledgment. A.D. and A.S. thank J. Dubochet for his interest and constant support. This work was supported in part by Swiss National Science Foundation, Grants 3152-068151 and 3100A0-103962. Correspondence and requests for materials should be addressed to A.S (Andrzej.Stasiak@unil.ch).

References and Notes

- (1) de Gennes, P.-G. Macromolecules 1984, 17, 703–704.
- (2) Adams, C. C. *The Knot Book*; Freeman and Co.: New York, 1994; p 306.
- (3) Janse van Rensburg, E. J.; Sumners, D. W.; Wasserman, E.; Whittington, S. G. J. Phys. A: Math. Gen. 1992, 25, 6557– 6656.
- (4) Katritch, V.; Olson, W. K.; Vologodskii, A. V.; Dubochet, J.; Stasiak, A. Phys. Rev. E 2000, 65, 5545–5549.
- (5) Belmonte, A.; Shelley, M. J.; Eldakar, S. T.; Wiggins, C. H. *Phys. Rev. Lett.* **2001**, 87, 114301–114304.
- (6) Pieranski, P.; Przybyl, S.; Stasiak, A. Eur. Phys. J. E 2001, 6, 123–128.

- (7) Mansfield, M. L. Struct. Biol. 1994, 1, 213-214.
- (8) Taylor, W. R. Nature (London) 2000, 406, 916-919.
- (9) Lifshitz, I. M.; Grosberg, A. Y.; Khokhlov, A. R. Rev. Mod. Phys. 1978, 50, 683-713.
- (10) Koniaris, K.; Muthukumar, M. Phys. Rev. Lett. 1991, 66, 2211–2214.
- (11) Freyd, P.; Yetter, D.; Hoste, J.; Lickorish, W.; Millett, K.; Ocneau, A. Bull. Am. Math. Soc. 1985, 12, 239–246. Ewing, B.; Millett, K. C. In Progress in Knot Theory and Related Topics; Hermann: Paris, 1996; pp 51–68.
- Topics; Hermann: Paris, 1996; pp 51–68.
 (12) Dobay, A.; Sottas, P.-E.; Dubochet, J.; Stasiak, A. Lett. Math. Phys. 2001, 55, 239–247.
- (13) Delbrück, M. İn Mathematical Problems in the Biological Sciences; Bellman: Providence, 1962; pp 55-63.
- (14) Sumners, D. W.; Whittington, S. G. J. Phys. A: Math. Gen. 1998, 21, 1689–1694.
- (15) Diao, Y. J. Knot Theory Its Ramifications 1995, 4, 189-196.
- (16) Darcy, I.; Sumners, D. W. Math. Proc. Cambridge Philos. Soc. 2000, 128, 497–510. Flammini, A.; Maritan, A.; Stasiak, A. Biophys. J. 2004, 87, 2968–2975.
- (17) de Ĝennes, P. G. *Scaling Concepts in Polymer Physics*; Cornell University Press: Ithaca, NY, 1979.
- (18) Katritch, V.; Olson, W. K.; Vologodskii, A.; Dubochet, J.; Stasiak, A. Phys. Rev. E 2000, 61, 5545-5549.
- (19) Deguchi, T.; Tsurusaki, K. J. Knot Theory Its Ramifications **1994**, 3, 321–353.
- (20) Stasiak, A.; Katritch, V.; Bednar, J.; Michoud, D.; Dubochet, J. Nature (London) 1996, 384, 142-145.
- (21) Vologodskii, A.; Crisona, N.; Laurie, B.; Pieranski, P.; Katritch, V.; Dubochet, J.; Stasiak, A. J. Mol. Biol. 1998, 278, 1–3.
- (22) Dobay, A.; Dubochet, J.; Millett, K.; Sottas, P.-E.; Stasiak, A. Proc. Natl. Acad. Sci. U.S.A. 2003, 100, 5611–5615.
- (23) Des Cloizeaux, J. J. Phys., Lett. 1981 42, L443-L436.
- (24) Deutsch, J. M. Phys. Rev. E 1999, 59, R2539-R2541.
- (25) Grosberg, A. Y. Phys. Rev. Lett. 2000, 85, 3858-3861.

- (26) Diao, Y.; Dobay, A.; Kusner, R. B.; Millett, K.; Stasiak, A. J. Phys. A: Math. Gen. 2003, 36, 11561–11574.
- (27) Moore, N. T.; Lua, R.; Grosberg, A. Y. Proc. Natl. Acad. Sci. U.S.A. 2004, 101, 13431–13435.
- (28) Orlandini, E.; Tesi, M. C.; Janse Van Rensburg, E. J.; Whittington, S. G. J. Phys. A: Math. Gen. 1998, 31, 5935– 5967.
- (29) Ben-Naim, E.; Daya, Z. A.; Vorobieff, P.; Ecke, R. E. Phys. Rev. Lett. 2001, 86, 1414-1417.
- (30) Metzler, R.; Hanke, A.; Dommersnes, P. G.; Kantor, Y.; Kardar, M. Phys. Rev. Lett. 2002, 88, 188101.
- (31) Marcone, B.; Orlandini, E.; Stella, A. L.; Zonta, F. arXiv: cond-mat/0405253, 2004, v2.
- (32) Farago, O.; Kantor, Y.; Kardar, M. Europhys. Lett. **2002**, 60, 53–59.
- (33) Vologodskii, A. V.; Levene, S. D.; Klenin, K. V.; Frank-Kamenetskii, M.; Cozzarelli, N. R. J. Mol. Biol. 1992, 227, 1224–1243.
- (34) Smith, S. B.; Finzi, L.; Bustamante, C. Science 1992, 258, 1122-1126.
- (35) Katritch, V.; Bednar, J.; Michoud, D.; Scharein, R. G.; Dubochet, J.; Stasiak, A. Nature (London) **1996**, 384, 142– 145.
- (36) Arsuaga, J.; Vazquez, M.; Trigueros, S.; Sumners, D. W.; Roca, J. Proc. Natl. Acad. Sci. U.S.A. 2002, 99, 5373–5377.
- (37) Arsuaga, J.; Tan, R. K.-Z.; Vazquez, M.; Sumners, D. W.; Harvey, S. C. *Biophys. Chem.* 2002, 101–102, 475–484.
- (38) Rybenkov, V. V.; Cozzarelli, N. R.; Vologodskii, A. V. Proc. Natl. Acad. Sci. U.S.A. 1993, 90, 5307-5311.
- (39) Shaw, S. Y.; Wang, J. C. Science 1993, 260, 533-536.
- (40) Mueller, J. E.; Du, S. M.; Seeman, N. C. J. Am. Chem. Soc. 1991, 113, 6306-6308.
- (41) Bucka, A.; Stasiak, A. Nucleic Acids Res. 2002, 30, e24.

MA048779A