



Journal of Nonlinear Mathematical Physics

ISSN (Online): 1776-0852

ISSN (Print): 1402-9251

Journal Home Page: <https://www.atlantis-press.com/journals/jnmp>

Propagation of Twist Solitons in Fully Inhomogeneous DNA Chains

Mariano Cadoni, Roberto de Leo, Sergio Demelio, Giuseppe Gaeta

To cite this article: Mariano Cadoni, Roberto de Leo, Sergio Demelio, Giuseppe Gaeta (2010) Propagation of Twist Solitons in Fully Inhomogeneous DNA Chains, Journal of Nonlinear Mathematical Physics 17:4, 557–569, DOI: <https://doi.org/10.1142/S1402925110001069>

To link to this article: <https://doi.org/10.1142/S1402925110001069>

Published online: 04 January 2021

Journal of Nonlinear Mathematical Physics, Vol. 17, No. 4 (2010) 557–569

© M. Cadoni *et al.*

DOI: [10.1142/S1402925110001069](https://doi.org/10.1142/S1402925110001069)

PROPAGATION OF TWIST SOLITONS IN FULLY INHOMOGENEOUS DNA CHAINS

MARIANO CADONI*, ROBERTO DE LEO[†] and SERGIO DEMELIO[‡]

*Dipartimento di Fisica, Università di Cagliari
and INFN, Sezione di Cagliari
Cittadella Universitaria, 09042 Monserrato, Italy*

^{*}*mariano.cadoni@ca.infn.it*

[†]*roberto.deleo@ca.infn.it*

[‡]*sergio.demelio@ca.infn.it*

GIUSEPPE GAETA

*Dipartimento di Matematica, Università degli Studi di Milano
via Saldini 50, 20133 Milano, Italy
giuseppe.gaeta@unimi.it*

Received 22 January 2010

Accepted 25 March 2010

In the framework of a recently introduced model of DNA torsional dynamics, we argued — on the basis of perturbative considerations — that an inhomogeneous DNA chain could support long-lived soliton-type excitations due to the peculiar geometric structure of DNA and the effect of this on nonlinear torsional dynamics. Here we consider an inhomogeneous version of this model of DNA torsional dynamics, and investigate numerically the propagation of solitons in a DNA chain with a real base sequence (corresponding to the Human Adenovirus 2); this implies inhomogeneities of up to 50% in the base masses and inter-pair interactions. We find that twist solitons propagate for considerable distances (2–10 times their diameters) before stopping due to phonon emission. Our results show that twist solitons may exist in realistic DNA chain models, and on a more general level that solitonic propagation can take place in highly inhomogeneous media. The most relevant feature for general nonlinear dynamics is that we identify the physical mechanisms allowing this behavior and thus the class of models candidate to support long-lived soliton-type excitations in the presence of significant inhomogeneities.

Keywords: Nonlinear models of DNA; soliton propagation in DNA chains.

1. Introduction

Solitonic excitations are a central feature in nonlinear dynamical systems [1, 2], and Nonlinear Science literature is to a considerable extent devoted to them.

Following a seminal paper, now thirty years old, by Engländer, Kallenbach, Heeger, Krumhansl and Litwin [3], it has been proposed that solitons might play an important role — not only dynamical but functional as well — also in processes undergone by the DNA double helix, such as transcription and denaturation [3–7].

Since the original proposal by Englander *et al.* [3], several simplified models for the nonlinear dynamics of the DNA chain have been proposed [8–21]; these models allow for solitonic solutions describing stretch and/or twist excitation of the DNA chain. Here, we will focus on twist excitations and thus torsional DNA dynamics models; in this framework, the soliton-supporting models of the family we consider [8, 16–21] (see e.g. [5] for a review of earlier models) describe the structural and dynamical features of DNA in a highly simplified way; in particular they model DNA either as a *homogeneous* polymer, or as an inhomogeneous one with rather special — and mathematically convenient — types of inhomogeneity.^a

It is not at all clear if solitonic excitations can exist and play a role in real DNA, i.e. when one takes into account inhomogeneities due to differences between the four nitrogen bases and their interactions; note that actually the genetic information is coded in the bases sequence and thus is intrinsically associated with inhomogeneities.

In fact, a strong objection against the existence of solitons in real DNA is precisely that solitons are expected to propagate in a significant way^b only in a *homogeneous* medium, hence only in the idealized homogeneous DNA.

Needless to say, this question is relevant not only in modeling DNA, but more generally in the nonlinear dynamics of inhomogeneous media; by this we mean both media which should in principle be homogeneous but are actually inhomogeneous due to defects or manufacturing imperfections (think e.g. of an optical fiber), and media which are intrinsically inhomogeneous (such as indeed DNA). While a completely general inhomogeneous media will not be able to support solitons, it appears that inhomogeneous media with a suitable structure^c can support solitons even when inhomogeneities are not small.

In the case of DNA, mass, inertia moment and intra-pair interaction strength differences between the four different bases are of the order of 50%, so that this definitely represents a significant test for the existence of solitons in inhomogeneous media beyond the perturbation regime.

In our analysis of the model introduced in [17], we argued that at the perturbation level one would expect the model to support solitons also in the inhomogeneous case; this was due to the interplay between geometrical and dynamical features of the model, ensuring that certain degrees of freedom — those related to the part of the molecule supporting inhomogeneities — became *auxiliary in perturbations* and were thus “slaved” to the degrees of freedom related to the fully homogeneous part; see [18–20] for details. As far as we know, the possibility of having solitons in fully inhomogeneous media via this mechanism was not noticed before.

^aIt should be noted that detailed numerical analysis of solitary wave excitations has been performed in the case of stretch DNA dynamics models, relevant for DNA denaturation [7]; here however, as already stated, we are concerned with torsional DNA dynamics.

^bIn DNA transcription, “significant” would mean over a full gene, i.e. from a few tens to some hundreds of bases. For other, more general, physical systems the relevance criterion should be based on relative propagation length with respect to soliton’s size.

^cIt is worth stressing that the general geometrical structure of DNA — i.e. a two-component medium, with a homogeneous component and an inhomogeneous one and a suitable geometry — is not unique to it, but is met in other media (not necessarily biological) as well; see e.g. [20] for an application of the mechanisms behind our model [17–19] to polyethylene.

As mentioned above, this result holds only at the perturbation level; unfortunately the inhomogeneous model is far too involved (see below for some explicit formulas) for an analytic study outside the perturbative region. On the other hand, differences between the four nitrogen bases are substantial and call for an analysis — necessarily, numerical — of the inhomogeneous case beyond the perturbation regime.

The purpose of this note is to present a first report of our numerical findings in this direction; they suggest that solitons are viable over rather long distances for a fully realistic inhomogeneous version of our model, in which the base sequence does not obey any given mathematical pattern, but is just the one of a real organism (the Human Adenovirus 2).^d

It is rather obvious — but nevertheless we would like to stress it again — that the relevance of the issue about the existence and stability of solitons in inhomogeneous media obviously goes far beyond DNA modeling: it concerns solitonic propagation in realistic — hence necessarily not perfectly homogeneous — nonlinear media. Any progress in this direction, in particular if *the physical mechanism allowing a significant soliton propagation in a given inhomogeneous medium is clearly identified*, as will be the case in the present study, is obviously relevant for a wide class of physical and technological applications.

Indeed, existence and stability of solitons, in particular topological solitons (kinks) as those considered here, is well established in a broad class of nonlinear phenomena (nonlinear optics, molecular chains, ferromagnetic waves, Josephson effect, etc. [2, 4, 22–25]), but these involve an idealized homogeneous medium. Some examples of solitons propagation in inhomogeneous media have been considered both in the framework of nonlinear DNA dynamics [16, 26–28] and of generic solitonic propagation [29]. However, these investigations either refer to localized inhomogeneities — in the form of e.g. discontinuities, potential barriers, delta potentials, etc.) — or else the inhomogeneities are parametrized by particular *ad hoc* rules.

The existence of solitonic perturbations in real DNA chains and in general inhomogeneous nonlinear media remains an involved and open question; we hope some aspects of it will be clarified by the present contribution. Starting from a homogenous system for which solitonic solutions exist, and introducing inhomogeneities in the system, one naturally expects that: (a) Solitons still propagate along the chain as long as the size of the soliton is substantially larger than the typical scale of the inhomogeneities; (b) The effect of the inhomogeneities on the soliton propagation will be the production of linear excitations (phonons) which will dissipate the kinetic energy of the soliton and bring the soliton to rest (or disappear for non-topological solitons) after a finite time.

The relevant question to be asked is the following: can there be fully inhomogeneous nonlinear systems in which solitons — of size comparable with the typical scale of the inhomogeneity — propagate for distances which are long enough for the soliton to be relevant for the physical (or biological) process it is assumed to describe?

^dWe stress these statements refer, as in general all of our work, to a Hamiltonian model of DNA *per se*. In actual biological functioning DNA is not isolated but immersed in the cell's fluid and submitted to thermal noise (and interacting with other biological molecules in its functioning, e.g. RNA-polymerase as far as transcription is concerned); thus a “realistic” DNA model should consider a randomly excited system in a highly dissipative medium. Our model — as all the literature concerned with Hamiltonian isolated DNA models — should therefore be seen as preliminary work studying isolated DNA before a more complete study of DNA in interaction with its environment.

As mentioned above, we will answer to this question (in the positive) for the case of soliton propagation in an inhomogeneous DNA chain with a real base sequence (we will consider the Human Adenovirus 2 (HA2) [34] base sequence) using a realistic — albeit simplified — model of DNA torsional dynamics. More precisely, we will observe solitons whose size is about 60 base pairs (*bp*), travelling along the inhomogeneous DNA chain for lengths up to 10 times their size. Further and more detailed numerical work in this direction is under way and will be described elsewhere [30]; it confirms the findings to be shown in this note.

2. Composite Inhomogeneous Model

Our starting point is (the inhomogeneous version of) the composite model for DNA torsional dynamics introduced and described in detail in [17]; this is a natural generalization of the model by Yakushevich [5, 8], taking into account more detail of the DNA molecule’s geometry.

In our model the state of each nucleotide at site n on chain $i = \pm 1$ is described by two angles, describing one the rotation of the sugar-phosphate group around the backbone chain (rotation angle $\theta_n^{(i)}$), and the other the rotation of the nitrogen base around the sugar atom to which this is attached (rotation angle $\varphi_n^{(i)}$); see Fig. 1. Thus a nucleotide is represented mechanically by a double pendulum, see Fig. 2, and the model can also be described as a double pendulums chain. See [17–19] for details.

Note that the inhomogeneity in real DNA is entirely contained in the structure and interactions of the bases, while the sugar-phosphate backbone is fully homogeneous. Thus our model treats separately the homogeneous and inhomogeneous components of the chain, and this will be a substantial feature of our treatment.

From the mechanical point of view our DNA model is a conservative system described by $4N$ angular coordinates $(\theta_n^{(i)}, \varphi_n^{(i)})$, $n = 1, \dots, N$, $i = 1, 2$. But, the bases cannot make a complete rotation about the sugar-phosphate group because of steric hindrances. We model this fact through an effective confining potential V_c whose energy wall is high enough to restrict the range of the bases’ angles $\varphi_n^{(i)}$ to some segment I shorter than 2π ; thus the coordinate space of our system is $(S^1)^{2N} \times I^{2N}$.

The dynamical evolution of the mechanical system is governed by the Lagrangian $\mathcal{L} = T - V$. Here the kinetic energy T has contributions from the sugar-phosphate group (T_t) and from the bases (T_s); thus $T = T_t + T_s$.

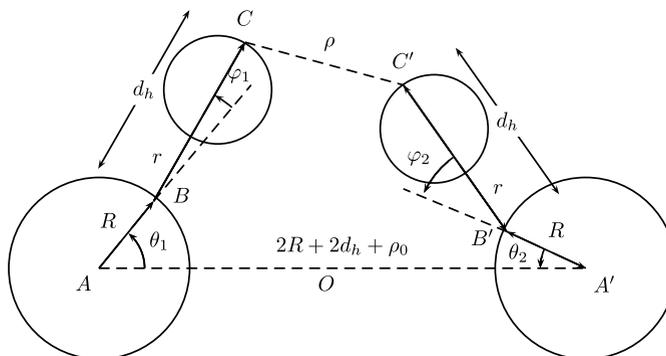


Fig. 1. The geometry of a nucleotide pair in our torsional model.

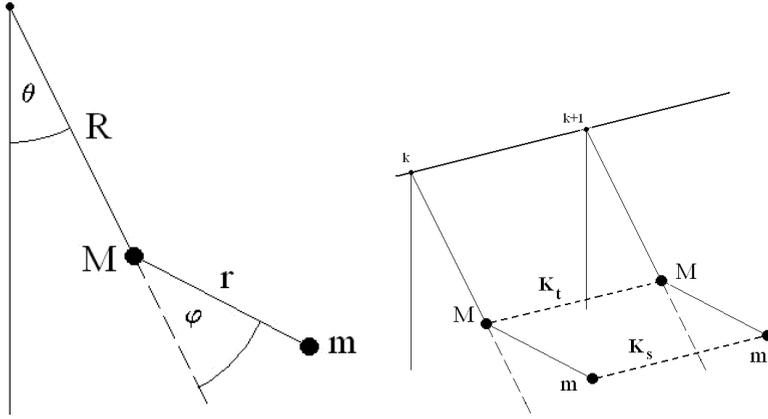


Fig. 2. The model can also be described as a chain of interacting double pendulums; here $M = I/R^2$.

With standard computations, as described in [17], the total kinetic energy for the double chain is

$$\begin{aligned}
 T = \frac{1}{2} \sum_n \sum_i & [m_n^{(i)} (r_n^{(i)})^2 (\dot{\varphi}_n^{(i)})^2 + 2m_n^{(i)} r_n^{(i)} (r_n^{(i)} + R \cos(\varphi_n^{(i)})) \dot{\theta}_n^{(i)} \dot{\varphi}_n^{(i)} \\
 & + (I + m_n^{(i)} (R^2 + (r_n^{(i)})^2) + 2m_n^{(i)} r_n^{(i)} R \cos(\varphi_n^{(i)})) (\dot{\theta}_n^{(i)})^2]. \quad (2.1)
 \end{aligned}$$

Here I and R are the moment of inertia (around its rotation axis) and the radius of the sugar-phosphate group, $m_n^{(i)}$ and $r_n^{(i)}$ are the mass and size of the nitrogen base (see Fig. 1 again) at site n on chain i . Note that now the parameters m and r (base mass and size) will also have chain and site indices, as they are different for different bases; see Table 1 for their values.

The potential energy has several contributions: $V = V_t + V_s + V_p + V_h + V_c$. Again the computations are exactly the same as in [17], except for the need to take into account the inhomogeneities of the chain. We will give only the final results, referring the reader to [17, 30] for detail.

The *torsional* potential energy V_t models the interaction between nearest neighbor sugar-phosphate groups and is described by a physical pendulum periodic potential,

$$V_t = K_t \sum_i \sum_n [1 - \cos(\theta_{n+1}^{(i)} - \theta_n^{(i)})]. \quad (2.2)$$

Table 1. Order of magnitude for the basic geometrical parameters of the DNA. Units of measure are: atomic unit for masses m , $1.67 \times 10^{-47} \text{kg} \cdot \text{m}^2$ for the inertia momenta I , Angstrom for l (the longitudinal width of bases and sugar), d_{bs} (the distances sugar-base) and d_{eq} (the distance at the equilibrium for the pairs AT and GC).

	A	T	G	C	Mean	Sugar
m	134	125	150	110	130	85
I	3.6×10^3	3.0×10^3	4.4×10^3	2.3×10^3	3.3×10^3	2.9×10^3
l	3.9	2.9	4.1	2.7	3.4	3.1
d_{bs}	1.0	1.0	1.0	1.0	1.0	—
d_{eq}	3.0	3.0	3.0	3.0	3.0	—

The *stacking* potential V_s models the $\pi - \pi$ bonds between the bases, and we use for it a simple harmonic potential in terms of the planar distance (distance between projections in a plane orthogonal to the molecule axis) between the centers of the base $P_n^{(i)}$. The stacking interaction will be treated as homogeneous along the chain^e; this choice is not completely justified physically — as stacking inhomogeneities due to different bases sequences are not always of lower order with respect to those present in the pairing interaction [31] — but corresponds to the desire of having a model as simple as possible (still maintaining inhomogeneities, to be introduced in the treatment of pairing, see below). In this way we get

$$V_s = \frac{1}{2}k_s \sum_i \sum_n d^2(P_{n+1}^{(i)}; P_n^{(i)}). \quad (2.3)$$

The *pairing* potential V_p models the bonds between bases in a given pair, i.e. at the same site on opposite chains; this interaction is described by a Morse-like potential in terms of the (planar, see above) distance between the extremal points $D_n^{(i)}$ of the discs modeling the bases on the opposite chains. The resulting potential term is

$$V_p = \sum_n k_{b_n} A [1 - e^{-a(d(D_n^{(i)}; D_n^{(-i)}) - d_0)}]^2. \quad (2.4)$$

Here D, a, d_0 are parameters characterizing the Morse potential; the dimensionless constant k_{b_n} takes into account the different potential energy in the AT (two hydrogen bonds) and GC (three hydrogen bonds) pairs. In our numerical simulations we will also use a harmonic approximation for the sake of comparison; this — albeit physically less significant than the Morse potential — was traditionally the simplest choice for modelling the pairing interaction [5, 8].

The *helicoidal* potential V_h models the forces between nucleotides in solution due to Bernal–Fowler filaments [32] and is taken to be

$$V_h = \sum_{n=1}^{N-5} \sum_{i=1} K_h [1 - \cos(\theta_{n+5}^{(i)} - \theta_n^{(i)})]. \quad (2.5)$$

The *confining* potential V_c , which models an “effective” interaction representing the steric constraint of the sugar-phosphate group on the bases, and is chosen as

$$V_c = \sum_n \sum_i K_c (\sin \varphi_n^{(i)})^{2M}, \quad (2.6)$$

where M is a large integer (we used $M = 50$ in our numerical calculations).

Our model is completely specified by also giving the various geometrical and dynamical parameters appearing in the Lagrangian. The evaluation of these is far from trivial due to the complexity of the molecular structure and the difficulty in making estimates of the related mechanical quantities. Here we use values in the range given in Table 2. They were determined in [21], to which we refer for a detailed discussion.

^eThus we will have a single stacking coupling constant K_s , rather than a $K_s^{(a)}$ depending on the site and base sequence. This simplifying assumption can be easily removed in numerical computations.

Table 2. Lower and upper bounds for the dynamical parameters of our model.

	K_t	K_s	A	a	K_h	K_c
lo bd	130 kJ/mol	—	30 meV	2 \AA^{-1}	$K_t/100$	2×10^4 kJ/mol
up bd	720 kJ/mol	16.6 N/m	50 meV	4 \AA^{-1}	$K_t/25$	2×10^4 kJ/mol

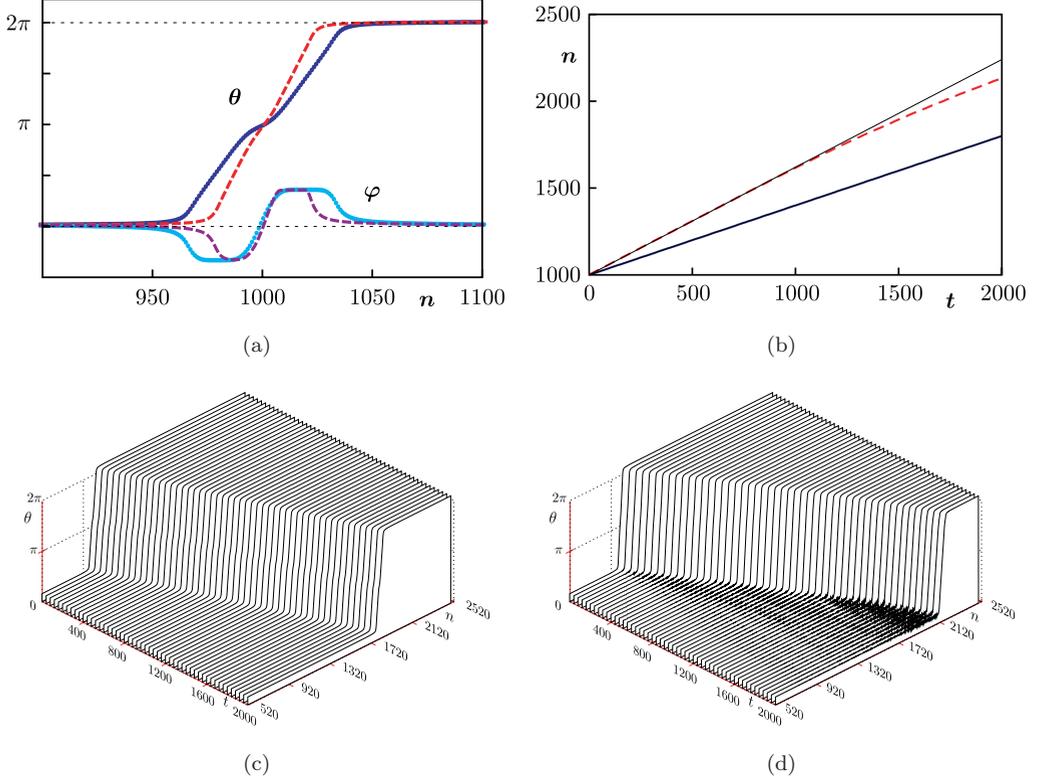


Fig. 3. (Color online) Profiles and time evolution for the composite, homogeneous, model with a pairing Morse potential, for solitons with topological numbers $(1, 1)$. (a) Initial profile with $v = 0.4$ (solid lines) and $v = 0.62 \simeq v_M$ (dashed lines). (b) Motion of the soliton center for $v = 0.4$ (thick solid line) and $v = 0.62$ (dashed line). The thin solid line represents the motion of the continuous soliton with speed $v = 0.62$. (c) Profiles of the soliton with $v = 0.4$ in the first 2000 TU (TU = 3.4×10^{-13} s) — no phonons emission is visible. (d) Profiles of the soliton with $v = 0.62$ in the first 2000 TU — phonons are emitted in backward direction as the soliton slows down.

We will not write explicitly the — quite involved — Euler–Lagrange equations corresponding to the Lagrangian $\mathcal{L} = T - V$.

3. Static and Travelling Solitons in Homogeneous Chains

Our task is to investigate the existence and the propagation of solitons along the DNA double chain within the model described so far. Investigations on this topic have been performed — for both harmonic and Morse pairing potential — at the analytic level (in the continuum limit) [20] and at the numerical one [17, 21]; but they have been almost exclusively concerned with the homogeneous approximation of the DNA chain. Moreover, most of the numerical investigations were focused on showing the existence of static solitonic

profiles of the chain and not on the propagation of the soliton in the chain. Despite some attempts [11, 16, 26–28] soliton propagation in real, hence inhomogeneous in a “generic” way, DNA chains has not been proved to occur, at least in a significant way.

We have now performed a systematic numerical analysis of the dynamics of the solitonic (kink) solutions of lowest topological charges $(0, 1)$, $(1, 0)$ and $(1, 1)$ (see below for their definition) of our system with $N = 2000$ and $N = 3000$ *bp* aimed at (a) Determine the initial profile of the soliton; (b) Let it evolve along the chain using that profile as initial condition. This analysis has been carried out for values of the parameters characterizing the DNA model within the physical range [21] and varying several features of the model:

- *Different bases sequences*: (a) A completely homogeneous chain; (b) A real DNA chain (the HA2); (c) A chain with random bases sequence.
- *Different pairing potential*: (a) Harmonic; (b) Morse.
- *Different DNA models*: (a) The composite model considered here; (b) The Yakushevich model, obtained by freezing to zero the angles $\varphi_{n,i}$.

The profile of the soliton is determined, as an extremum of the action, by two data: the boundary conditions and the position of its center. Since we are interested only in kinks in the topological angles θ , we use the boundary conditions $\varphi_1^{(i)} = \varphi_N^{(i)} = 0$ for the angles φ , while for the θ ones we use $\theta_1^{(i)} = 0$, $\theta_N^{(i)} = 2k_i\pi$ for some integers k_i ; these are the topological charges mentioned above. We use in the discrete action the analog of the continuous travelling wave ansatz $\dot{q}_n = -v\Delta_n q/\delta$, with $q_n(t) = (\theta_n^{(1)}(t), \theta_n^{(2)}(t), \varphi_n^{(1)}(t), \varphi_n^{(2)}(t))$ and δ the spatial separation between sites in the chain. The resulting discrete action is extremized using the “conjugate gradients” method in the independent implementations of NR [35] and GNU’s GSL [36]. The data obtained through the previous procedure provide the coordinates q_n at time $t = 0$. Their graph, as function of the discrete variable n , represents the profile of a kink with some speed given in input.

To study the evolution of the system from this initial data, we use the Hamiltonian formulation. To implement the Hamiltonian evolution of the system we have used two independent algorithms: the GSL version of the standard Runge–Kutta Prince–Dormand method and a Hamiltonian symplectic integrator kindly provided by Hairer [37].

4. Results

The main results of our numerical investigation are described below; a more extended and detailed version will be presented elsewhere [30]. For our values of the physical parameters static soliton profiles always exist for the (physically more realistic) Morse pairing potential but not for the harmonic one. Moreover, the diameter and the energy of the kink depend on its speed in the expected way: as the speed approaches the limiting solitons speed v_M , the energy increases whereas the diameter shrinks. This is perfectly consistent with the relativistic nature of the kinks.

The propagation of the soliton on the chain, differently from the determination of the profiles, is highly sensitive to different choices of models, form of pairing potential and kind of chain sequence. We will, therefore, discuss separately the propagation of the soliton in the homogeneous and inhomogeneous chain.

4.1. Homogeneous chain

In the continuous limit of the homogeneous DNA chain, kinks (if they exist and are stable) propagate at constant speed without losing energy. We expect this to be not completely true for the discrete homogeneous DNA chain, as invariance under continuous translation is lost. Owing to discreteness of the chain we expect the propagating solitons to lose kinetic energy through phonon emission, although at a low rate. Moreover, the phonon emission rate should increase with the soliton speed. These expectations have been confirmed by our numerical simulations, see Fig. 3. When the soliton speed is much smaller than the maximum soliton speed v_M the soliton profile is wider, travels at constant speed and no emission of phonons is visible. When the phonon speed gets closer to v_M , the soliton diameter becomes smaller and phonons are emitted.

The behavior is qualitatively the same in the case of the Yakushevich model with a pairing Morse potential. On the other hand, using an harmonic pairing potential we have a qualitatively different behavior. In particular, in the case of the composite model with harmonic pairing potential the soliton does not propagate at all, implying that non-static solitons in this case do not exist [33].

4.2. Inhomogeneous chain

In Fig. 4 we show our results for the propagation of solitons in the DNA sequence of the HA2. The profiles are slightly more “wavy” than in the homogeneous case, stay almost identical to the initial one and emit phonons, as expected, even far from v_M ; moreover phonons in this case are clearly emitted in both directions and, again, much more when the velocity is close to the upper limit. For $v = 0.4$ km/s, the motion remains constant for about 700 TU ($\text{TU} = 3.4 \times 10^{-13}$ s) then slowly decelerates, whereas for $v = 0.45$ km/s it soon deviates from a constant speed.

Every kink along the chain has a behavior of this kind, and both diameter and energy do not change much (within 10^{-2}). On the other hand, the maximal distance d traveled by the soliton depends on the initial position of its center. We show in Fig. 5 a distribution of d performed along a complete sequence of about 30,000 sites (sampled with a step of ten sites) for the HA2 and for a random chain. In the case of HA2 (Fig. 5(a)) the mean diameter of the solitons is $\bar{\rho} \simeq 68$ bp with standard deviation $\sigma \simeq 2$ bp; the average length traveled by kinks is $\bar{d} \simeq 345$ bp with standard deviation $\sigma \simeq 60$ bp. In both cases of the HA2 and of the random chain d is always between 150 bp and 600 bp, namely between 2 and 10 times their diameter.

With reference to our criteria for the physical or biological relevance of the solitons (see the Introduction, in particular footnote *b* in there), we note that this means solitons travel for a distance significant both in relative terms (with respect to their size) and in absolute terms (i.e. for a significant number of bases).^f

^fIt should also be mentioned that the absolute size of the solitons is slightly too large — albeit of the right order of magnitude — compared with real “transcription bubbles” in DNA, which are of the order of 30 bp, i.e. half of the one observed here. We believe however that obtaining the right order of magnitude is quite remarkable considering how rough our model is, and moreover that in our model we do not take into account the interaction with RNA-polymerase, essential to transcription.

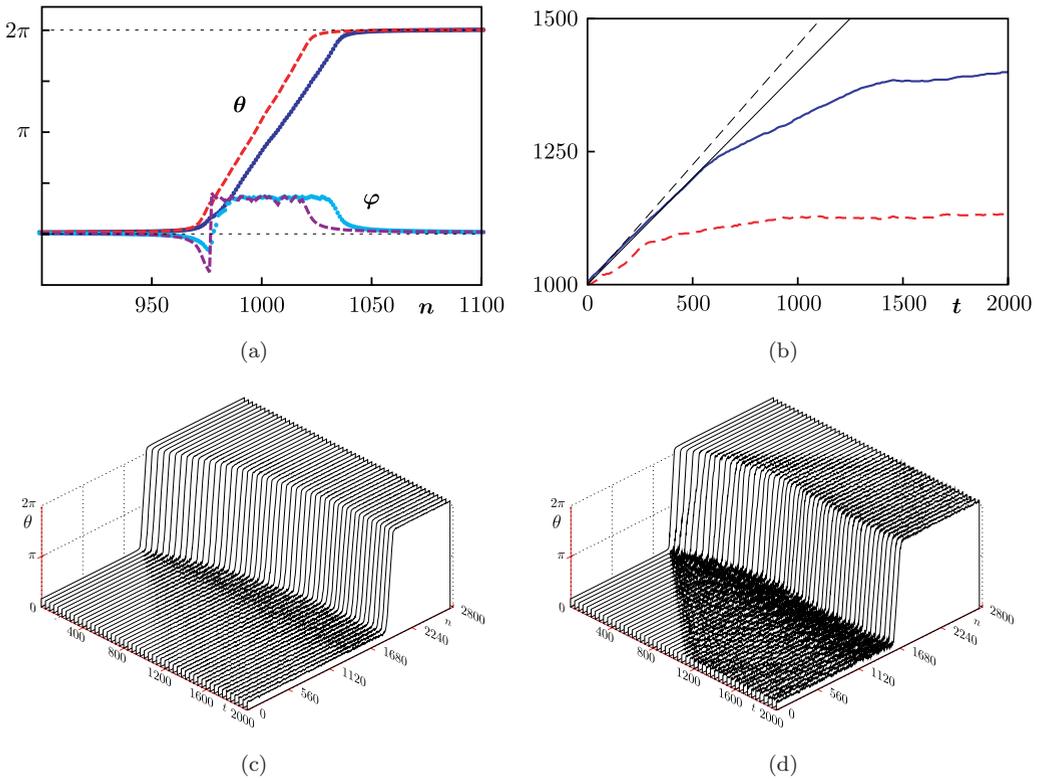


Fig. 4. (Color online) Profiles and time evolution of solitons for the composite model in the HA2 sequence with a Morse potential, topological numbers $(1, 1)$ — so the profiles on the two chains coincide. (a) Initial profile with $v = 0.4$ (solid lines) and $v = 0.45 \simeq v_M$ (dashed lines). (b) Motion of the soliton center for $v = 0.4$ (thick solid line) and $v = 0.45$ (dashed thick line). The solid and dashed thin lines represent the corresponding constant-speed motion of the continuous soliton. (c) Profiles of the soliton within the first 2000 TU — very little phonons emission is visible in both directions. (d) Profiles of the soliton with $v = 0.62$ in the first 2000 TU — phonons are clearly emitted in both directions but mainly backward.

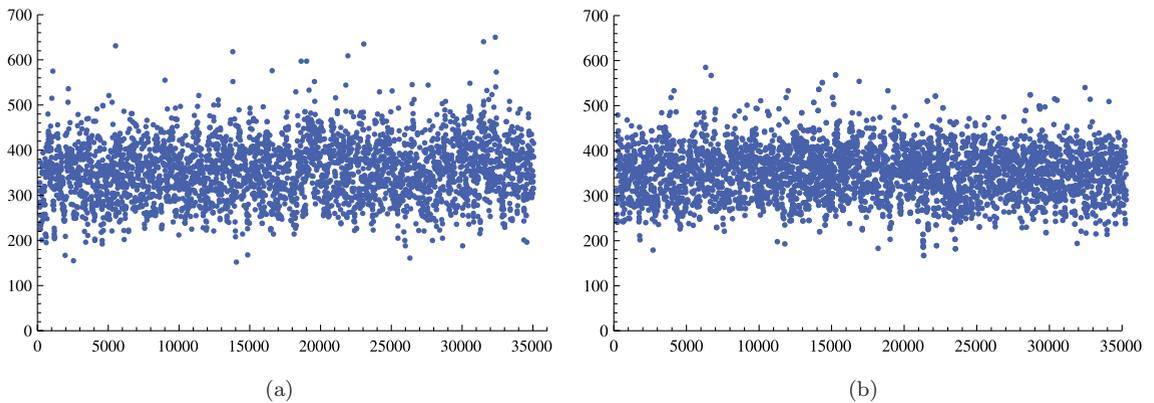


Fig. 5. Maximal distance reached by kinks in the composite-Morse model as function of the initial coordinate of its center in the HA2 (a) and in a random chain (b).

For the Yakushevich model with a Morse pairing potential, soliton propagation in the real DNA chain, relative to a similar distribution, is qualitatively the same but the average traveled distance drops to $\bar{d} \simeq 272 \text{ bp}$ ($\sigma \simeq 39 \text{ bp}$) with a mean diameter $\bar{\rho} \simeq 66 \text{ bp}$ ($\sigma \simeq 3 \text{ bp}$). All data for d lie now between 130 bp and 400 bp .

A key observation is that the average distance traveled by the soliton does not depend on the particular combinatorics of the DNA sequence. Indeed we evaluated the same quantities on a random sequence of same length and found the very same statistics (Fig. 5(b)).

5. Conclusions

In this note we have reported on our numerical investigation about the propagation of twist solitons in a real DNA chain. Summarizing, we obtained the following results for what concerns the DNA model concretely investigated.

(1) The simulations show that solitons of size of about 60 bp can propagate in the DNA chain till 10 times the soliton size. Since we use a simplified mechanical model of DNA, which does not take into account effects that may enhance the soliton performance (nor the presence of the DNA Polymerase in real transcription), our results give a strong indication that twist solitons may indeed be present in real DNA and play a role in its transcription, as first suggested by Englander *et al.* [3].

(2) A second relevant result is that, for soliton dynamics, a real DNA sequence is almost indistinguishable from a random one. This means that the DNA's sequence, which is of course fundamental for biological processes, does not play a significant role in the torsional dynamics of DNA; this enhances point (1) above, as the transcription mechanism is independent of the actual base sequence.

We believe our study is also of some significance for Nonlinear Dynamics of inhomogeneous media; in this respect, we believe the following conclusions can be drawn.

(3) In more general terms, our numerical investigation has shown that soliton propagation is also possible in fully inhomogeneous media with a suitable structure. This possibility can be traced back to two different features of DNA, taken into account by our model. The first is the presence in the molecular chain of both a homogeneous part that supports the topological soliton (the sugar-phosphate group) and an inhomogeneous part (the bases) that plays the role of a dissipative medium. The second is that the Morse potential localizes the interaction of the inhomogeneous part essentially near the potential minimum (away from this minimum the interaction becomes very weak); again, this weakens the soliton sensitivity to inhomogeneities in the chain.

(4) The statement above also implies that when solitons are possible in an inhomogeneous medium with a structure like the one we are discussing (see point 3 above), this should not depend on the actual sequence of inhomogeneities, i.e. solitons would be present also in a randomly inhomogeneous medium (with of course some limit on the range of variation of the randomly distributed characteristic of the medium).

In summary, we believe this numerical work shows that solitons can exist in substantially inhomogeneous media with the two-component structure embodied in our model [17–19] well beyond the perturbation regime; albeit we worked with a model of DNA, and actually with

a specific DNA sequence, it is clear that the actual sequence is inessential (see also point 2 above) and that our work shows in more general term the possibility of having long-lived solitonic type excitations in fully inhomogeneous models with a suitable geometric structure. As already mentioned, further and more extensive numerical work in this direction is under way [30]; preliminary results confirm the findings discussed in this note.

Acknowledgments

We thank E. Hairer for kindly sending us his Hamiltonian symplectic integrator, and G. Derks for useful discussions about simulation of solitons in nonlinear chains.

References

- [1] Th. Dauxois and M. Peyrard, *Physics of Solitons* (Cambridge University Press, Cambridge, 2006).
- [2] R. K. Dodd, J. C. Eilbeck, J. D. Gibbon and H. C. Morris, *Solitons and Nonlinear Wave Equations* (Academic Press, London, New York, 1982).
- [3] S. W. Englander, N. R. Kallenbach, A. J. Heeger, J. A. Krumhansl and A. Litwin, Nature of the open state in long polynucleotide double helices: possibility of soliton excitations, *Proc. Nat. Acad. Sci. USA* **77** (1980) 7222–7226.
- [4] A. S. Davydov, *Solitons in Molecular Systems* (Kluwer, Dordrecht, 1981).
- [5] L. V. Yakushevich, *Nonlinear Physics of DNA* (Wiley, Chichester, 1998) second edition (2004).
- [6] G. Gaeta, C. Reiss, M. Peyrard and Th. Dauxois, Simple models of non-linear DNA dynamics, *Rivista del Nuovo Cimento* **17** (1994) n.4, 1.
- [7] M. Peyrard, Nonlinear dynamics and statistical physics of DNA, *Nonlinearity* **17** (2004) R1–R40.
- [8] L. V. Yakushevich, Nonlinear DNA dynamics: A new model, *Phys. Lett. A* **136** (1989) 413–417.
- [9] M. Peyrard and A. Bishop, Statistical mechanics of a nonlinear model for DNA denaturation, *Phys. Rev. Lett.* **62** (1989) 2755–2758.
- [10] Th. Dauxois, Dynamics of breather modes in a nonlinear helicoidal model of DNA, *Phys. Lett. A* **159** (1991) 390–395.
- [11] F. Zhang and M. A. Collins, Model simulations of DNA dynamics, *Phys. Rev. E* **52** (1995) 4217–4224.
- [12] M. Barbi, S. Cocco and M. Peyrard, Helicoidal model for DNA opening, *Phys. Lett. A* **253** (1999) 358–369.
- [13] M. Barbi, S. Cocco and M. Peyrard, Vector nonlinear Klein-Gordon lattices: General derivation of small amplitude envelope soliton solution, *Phys. Lett. A* **253** (1999) 161–167.
- [14] M. Barbi, S. Cocco, M. Peyrard and S. Ruffo, A twist-opening model of DNA, *J. Biol. Phys.* **24** (1999) 97–114.
- [15] S. Cocco and R. Monasson, Statistical mechanics of torque induced denaturation of DNA, *Phys. Rev. Lett.* **83** (1999) 5178–5181.
- [16] L. V. Yakushevich, A. V. Savin and L. I. Manevitch, Nonlinear dynamics of topological solitons in DNA, *Phys. Rev. E* **66** (2002) 016614.
- [17] M. Cadoni, R. De Leo and G. Gaeta, A composite model for DNA torsion dynamics, *Phys. Rev. E* **75** (2007) 021919.
- [18] M. Cadoni, R. De Leo and G. Gaeta, Sine-Gordon solitons, auxiliary fields, and singular limit of a double pendulums chain, *J. Phys. A* **40** (2007) 12917–12929.
- [19] M. Cadoni, R. De Leo and G. Gaeta, Solitons in a double pendulums chain model, and DNA roto-torsional dynamics, *J. Nonlin. Math. Phys.* **14** (2007) 128–146.
- [20] M. Cadoni, R. De Leo, S. Demelio and G. Gaeta, Twist solitons in complex macromolecules: From DNA to polyethylene, *Int. J. Non-Linear Mech.* **43** (2008) 1094–1107.

- [21] R. De Leo and S. Demelio, Numerical analysis of solitons profiles in a composite model for DNA torsion dynamics, *Int. J. Non-Linear Mech.* **43** (2008) 1029–1039.
- [22] A. Barone and G. Paternò, *Physics and Applications of the Josephson Effect* (Wiley, New York, 1982).
- [23] J. R. Taylor, *Optical Solitons* (Cambridge University Press, Cambridge, 1992).
- [24] C. Cattuto and F. Marchesoni, Unlocking of an elastic string from a periodic substrate, *Phys. Rev. Lett.* **79** (1997) 5070–5073.
- [25] N. Manton and P. Sutcliffe, *Topological Solitons* (Cambridge University Press, Cambridge, 2004).
- [26] M. Salerno, Discrete model for DNA-promoter dynamics, *Phys. Rev. A* **44** (1991) 5292–5297.
- [27] M. Daniel and V. Vasumathi, Nonlinear molecular excitations in a completely inhomogeneous DNA chain, *Phys. Lett. A* **372** (2008) 5144–5155.
- [28] L. Morales Molina and O. Pavel Fernandez, DNA torsional solitons in the presence of localized inhomogeneities, *Phys. Scripta* **66** (2002) 385–390.
- [29] L. Ferreira, B. Piette and W. J. Zakrzewski, Dynamics of the topological structures in inhomogeneous media, *J. Phys. Conf. Series* **128** (2008) 012027.
- [30] M. Cadoni, R. De Leo and S. Demelio, Soliton propagation in homogeneous and inhomogeneous models for DNA torsion dynamics, accepted for publication on *J. Nonlin. Math. Phys.*
- [31] W. Saenger, *Principles of Nucleic Acid Structure* (Springer, Berlin, 1984).
- [32] J. D. Bernal and R. H. Fowler, A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions, *J. Chem. Phys.* **1** (1933) 515–548.
- [33] R. De Leo and G. Derks, in preparation.
- [34] <http://www.ebi.ac.uk/cgi-bin/emblfetch?J01917>.
- [35] <http://www.nr.com/>.
- [36] <http://www.gnu.org/software/gsl>.
- [37] <http://www.unige.ch/~hairer/software.html>.