A minimal model of DNA dynamics in interaction with RNA-Polymerase

Gianne Derks
Department of Mathematics, University of Surrey
Guildford, Surrey GU2 7XH (UK)

Giuseppe Gaeta
Dipartimento di Matematica, Università di Milano
via Saldini 50, 20133 Milano (Italy)

August 25, 2011

Abstract

We introduce a minimal mesoscopic model for the DNA/RNAP complex; this is obtained as an extension of the familiar Yakushevich model for DNA dynamics. We study in particular the existence and stability of topological solitary waves for our model, motivated by the literature on would-be solitonic excitations in DNA.

Keywords: Solitary waves; DNA/RNAP interaction; Perturbed sine-Gordon-type equations.

AMS subject classifications: 34D35, 35Q53, 37K50.

1 Introduction

The possibility that solitons are present – and play a functional role – in the DNA chain was first discussed by Englander et al. in a seminal paper [1] over thirty years ago. We will refer to this as the “solitons in DNA” (SDNA) hypothesis. This fascinating suggestion called the attention of researchers, in particular but not only theoretical physicists and nonlinear dynamics people, and triggered the formulation of several simple models of nonlinear DNA dynamics. Here by “simple” we mean models in which the state of each nucleotide is described by few – often just one or two – degrees of freedom, i.e. mesoscopic models. These models can be roughly divided in two families: those discussing “radial” deformation of the DNA double helix, related to DNA denaturation [2, 3, 4, 5, 6, 7, 8, 9, 10]; and those discussing “torsional” deformation of DNA, more closely in the spirit of the first suggestion by Englander et al., related to DNA transcription [11, 12, 13, 14, 15, 16, 17]. Here we are interested in this second class of models, and in DNA transcription.

Several serious arguments could be, and were, raised against the SDNA hypothesis. Let us mention what we feel are the strongest ones:

(i) The substantial inhomogeneity of the DNA molecule – inhomogeneity due to the sequence of nitrogen bases, i.e. to the very coding of genetic information in DNA – would prevent solitons to travel over appreciable distances before stopping due to phonon emission.

(ii) The models of DNA usually considered so far and supporting solitons are Hamiltonian; but it is well known that DNA evolves in a highly dissipative medium (the cell fluid), and is moreover subject to thermal motion.

(iii) All studies conducted so far to support the SDNA hypothesis analyze models of DNA (able to carry solitons) which are indeed models of DNA alone. This is sufficient when we consider DNA denaturation, as in the Peyrard-Bishop or Poland-Scheraga models [18, 19]; or even when

---

1Provided of course environment characteristics (e.g. salinity) are taken into account in the model parameters.

---

Provided of course environment characteristics (e.g. salinity) are taken into account in the model parameters.
we consider single-molecule laboratory experiments \[20, 21, 22\]. But in DNA transcription, a second main actor is present, i.e. RNA-Polymerase. It is thus quite fair to state that no model of DNA alone can really claim to shed light on the mechanism of transcription\(^2\).

The first objection (i) was substantially removed by considering models which take into account in greater detail the DNA geometry and structure \[23, 24\], and in particular which take into account the fact that the sugar-phosphate backbone (SPB) is completely homogeneous. By describing the state of the SPB unit and of the attached nitrogen base by separate (torsional) degrees of freedom \[25\] (the model is hence said to be a “composite” one), and taking into account the steric hindrances due to the geometry of the DNA molecule, it turns out that solitary wave excitations are substantially carried by the degrees of freedom related to the SPB, i.e. to the completely homogeneous part of the molecule. This suggested, on the basis of analytical and perturbative considerations \[25, 26, 27\], that such models could carry solitons over considerable distances even when taking fully into account the inhomogeneities of the base sequence, i.e. even in their “realistic” versions. This was recently confirmed by careful numerical simulations of the dynamics of a fully realistic such model – i.e. a model in which the base sequence is that of a really existing organism, and the relevant physical characteristics of bases at each site are considered \[28, 29\]. We can thus claim that the problem of phonon emission is not any more a strong argument against the SDNA proposal – at least at the mesoscopic level.

As for objection (ii), preliminary work on a version of the composite model in which forcing and dissipation have been introduced \[29\], following a method proposed by Yakushevich \[30\], showed that again the solitons of the composite model are present and able to travel over long distances when one takes into account dissipation.

Now that the problems (i) and (ii) related to inhomogeneity of the base sequence can be considered as substantially solved, it is in our opinion time to face objection (iii), i.e. to tackle the elaboration of models for the DNA/RNAP interaction, with a particular attention to the question if these can support solitons.

The purpose of this paper is precisely to propose a “minimal” model for the DNA/RNAP interaction (able also to deal with the dissipative nature of the motion of DNA in its environment). We intend to focus on this problem, and hence we will resort to a very simplified model of DNA – which is at the same time the basis for the more complex models mentioned above \[25\] – i.e. the model by Yakushevich\(^3\) \[11, 12\], in which the state of each nucleotide is described by a single angular variable. Moreover, our discussion will be conducted in terms of a continuum versions of the model; this is quite customary in discussing DNA models and is justified by the fact the structures we are studying extend over space scales much larger than the discreteness scale (set by the inter-base distance), and also by the successes obtained in this way in previous studies \[4, 12\]. The consideration of more refined models for DNA on the one hand, and of more complete description of the forces responsible for the movement of RNAP along the DNA chain on the other hand, will not be pursued here.

The plan of the paper is as follows. First of all, in section 2, we discuss the Y (Yakushevich) model for DNA (and a slight extension of it, embodying a phenomenological friction term): we will focus in particular on travelling solitary wave solutions and their stability. In section 3 we introduce our model, which extends the modelling approach leading to the Y model to the case where one is considering the DNA/RNAP complex. We will first discuss the simplest framework (model A), in which the RNAP motion is given – and corresponds to a translation along the DNA chain with constant speed – and then pass to consider the realistic situation where RNAP motion is due to pulling along the DNA and powered by ATP depletion (model B). This second model will admit constant speed RNAP motions as attracting stationary states, so that model A will actually describe the asymptotic dynamics of model B. We will also consider how these models are modified by taking into account dissipative effects due to the fluid environment in which the DNA/RNAP

\(^2\)We would like to stress that we are in no way criticizing authors (including one of us) studying these model so far: the first step in understanding if solitons can be present in the DNA/RNAP complex is of course to study the DNA nonlinear dynamics, and it has thus been entirely natural to first focus on this.

\(^3\)As recalled below, more refined versions of this model would give substantially equivalent results for what concerns solitonic excitations; we will thus stay with its original version.
complex is embedded. It will be convenient to reformulate and study the model in a moving frame, sliding along DNA with the RNAP, which we present in section 3 as well. The next section 4 is devoted to the (analytical) study of existence of solitary travelling wave solutions in our models; we study both the case where friction is disregarded and the one where it is taken into account. We are able to give existence results based on transversality arguments; as such, the proofs are robust but do not give a direct information about the explicit form of solutions. In section 5 these travelling solitary wave solutions are investigated numerically via a direct simulation of the evolution equations of our models. Finally, section 6 is devoted to conclusions, and to a discussion of our results.

2 DNA modelling

As mentioned above, our description of DNA will be at the level of the Yakushevich model \[11, 12\], called Y model in the following. This is a homogeneous model (i.e. all bases are in this description equal, as are the interactions between bases at different sites), in which only base rotations about the sugar-phosphate backbone are considered (see \[12\] for a review of this and similar models). Thus, the state of each base will be described by a rotation angle, and we denote by \(I\) the inertia moment of the bases for this rotation and by \(\phi_n^{(i)}\) the angle referring to the base at site \(n \in \mathbb{Z}\) on the chain \(i = \pm\). The native DNA configuration (B-DNA) will correspond to \(\phi_n^{(+)} = 0\) for all \(n\) and \(i\), i.e. the angles \(\phi_n^{(\pm)}\) measure displacement from the equilibrium configuration.

Figure 1: A sketch of the DNA in the Y model: on the left, the sugar-phosphate backbone is denoted by the bold green line, the (blue) bases rotate about the backbone. The coordinates \(\phi^{(+)}\) and \(\phi^{(-)}\) for a pair of bases at a fixed site are defined on the right.

It is convenient to change variables, passing to
\[
\psi_n = (\phi_n^{(+) + \phi_n^{(-)})/2}, \quad \chi_n = (\phi_n^{(+) - \phi_n^{(-)})/2}, \tag{1}
\]
hence \(\chi = 0\) corresponds to symmetric configurations and \(\psi = 0\) to anti-symmetric ones.

2.1 Discrete Y model

Interactions among bases will be modelled by interaction potentials. These interactions amount to pairing between bases at corresponding sites on different chains, modelled by a potential \(\hat{V}_p[\phi_n^{(+) + \phi_n^{(-)}]}\); and by stacking between bases at successive sites on the same chain, modelled by a potential \(\hat{V}_s[\phi_n^{(i)} + \phi_n^{(i)}]\). The model is described by the Lagrangian
\[
L = \hat{T} - \hat{V}_p - \hat{V}_s. \tag{2}
\]
The kinetic energy \(\hat{T}\) is elementary to compute (we just have to pass to the new variables \(\psi\) and \(\chi\)); the stacking potential \(\hat{V}_s\) will be harmonic; as for the pairing potential \(\hat{V}_p\), this will be anharmonic and characterizes the model. The choice by Yakushevich gives
\[
\hat{T} = \sum_n (I/2) \left( \dot{\psi}_n + \dot{\chi}_n \right); \quad \hat{V}_s = \sum_n (K_s/2) \left[ (\psi_{n+1} - \psi_n)^2 + (\chi_{n+1} - \chi_n)^2 \right]; \quad \hat{V}_p = (K_p/2) r^2 \sum_n (1 + \cos^2 \chi_n - 2 \cos \psi_n \cos \chi_n). \tag{3}
\]
The corresponding (Euler-Lagrange) equations of motion are

\[ \begin{align*}
I \ddot{\psi}_n &= K_s (\psi_{n+1} - 2\psi_n + \psi_{n-1}) - K_p r^2 \sin \psi_n \cos \chi_n ; \\
I \ddot{\chi}_n &= K_s (\chi_{n+1} - 2\chi_n + \chi_{n-1}) - K_p r^2 \sin \chi_n (\cos \psi_n - \cos \chi_n) .
\end{align*} \tag{4} \]

Note that if we wish to consider also friction forces, due to interaction with the cellular fluid, this would amount in the simplest framework to add terms \(-\lambda \dot{\psi}_n\) and respectively \(-\lambda \dot{\chi}_n\) to the equations in (4); here \(\lambda\) is the friction coefficient for each nucleotide.

2.2 Continuum Y model

We will now pass to the continuum description, promoting the arrays \(\psi_n(t), \chi_n(t)\) to interpolating fields \(\psi(x,t), \chi(x,t)\); the relation between arrays and fields is given by

\[ \psi(n\delta, t) = \psi_n(t) , \quad \chi(n\delta, t) = \chi_n(t) . \tag{5} \]

Here \(\delta\) is the distance between successive sites along the axis of the double helix (this is \(\delta \approx 3.4\) Å in B-DNA), and we choose the origin of \(x\) to coincide with site \(n = 0.\)

In this way, the ODE system (4) is replaced – omitting terms of order \(\delta^4\) – by two field equations. We write, for ease of notation,

\[ \kappa_s := (K_s \delta^2 / I) , \quad \kappa_p := (K_p r^2 / I) . \tag{6} \]

With these, the Euler-Lagrange equations read

\[ \begin{align*}
\psi_{tt} &= \kappa_s \psi_{xx} - \kappa_p \sin \psi \cos \chi ; \\
\chi_{tt} &= \kappa_s \chi_{xx} - \kappa_p \sin \chi (\cos \psi - \cos \chi) .
\end{align*} \tag{7} \]

Consideration of dissipative forces arising from interaction with the surrounding cellular fluid, with dissipation coefficient \(\lambda\), would amount to adding terms \(-\lambda \psi_t\) and \(-\lambda \chi_t\) respectively to the right hand side of the equations above, where \(\lambda = \lambda / I\).

The space \(\chi = 0\) is invariant under the (7) – as well as under (4) – so that we can restrict to symmetric configurations \(\phi^{(+)}_n = \phi^{(-)}_n\) (and hence \(\chi = 0\)). In this case (7) reduce to the classical sine-Gordon equation [4, 31, 32]

\[ \psi_{tt} = \kappa_s \psi_{xx} - \kappa_p \sin \psi . \tag{8} \]

The space \(\psi = 0\) (antisymmetric configurations \(\phi^{(+)}_n = -\phi^{(-)}_n\)) is also invariant under (7); in this case we are reduced to the double sine-Gordon equation [4, 33, 34]

\[ \chi_{tt} = \kappa_s \chi_{xx} - \kappa_p \sin \chi (1 - \cos \chi) . \tag{9} \]

2.3 Limit conditions and kink solutions

The equations (7) can as well be obtained from the Lagrangian

\[ L = \int_{-\infty}^{+\infty} \mathcal{L} \, dx , \]

with Lagrangian density

\[ \mathcal{L} = \frac{1}{2} \left[ (\psi_t^2 - \kappa_s \psi_x^2) + (\chi_t^2 - \kappa_s \chi_x^2) \right] + \frac{1}{2} \kappa_p (1 + \cos^2 \chi - 2 \cos \psi \cos \chi) . \]

The total energy \(\mathcal{H}(t)\) at time \(t\) is then given by

\[ \mathcal{H} = \int_{-\infty}^{+\infty} \frac{1}{2} \left[ (\psi_t^2 + \kappa_s \psi_x^2) + (\chi_t^2 + \kappa_s \chi_x^2) \right] - \frac{1}{2} \kappa_p (1 + \cos^2 \chi - 2 \cos \psi \cos \chi) . \]

\(^{4}\)The same operation can of course be done on the arrays \(\phi^{(+)}_n(t)\), promoting them to fields \(\phi^{(+)}(x,t)\).
In order to have a finite total energy, it is thus needed that for \( x \to \pm \infty \) the \( t \) and \( x \) derivatives of both \( \psi \) and \( \chi \) go to zero, while the fields themselves should go to a minimum of the local potential \( \hat{V}_p \), i.e. to multiples of \( \pi \) with their sum being a multiple of \( 2\pi \) (this corresponds to \( \phi \pm (x,t) \) going to multiples of \( 2\pi \)). Note that when we restrict to the symmetric (or antisymmetric) fields configurations only, then the fields \( \psi \) and \( \chi \) are required to go to multiples of \( 2\pi \) themselves, as obvious from the previous remark and (1).

Thus, summarizing, we are restricted to consider field configurations which satisfy in all cases

\[
\lim_{x \to \pm \infty} \psi_t = \lim_{x \to \pm \infty} \psi_x = \lim_{x \to \pm \infty} \chi_t = \lim_{x \to \pm \infty} \chi_x = 0 ;
\]

and, with \( m_\pm \) integers such that \( m_+ + m_- \) is even,

\[
\lim_{x \to \pm \infty} \psi = m_+ \pi , \quad \lim_{x \to \pm \infty} \chi = m_- \pi .
\]

In the case of symmetric or antisymmetric field configurations, the latter is replaced (with \( n_\pm \) integers)

\[
\lim_{x \to \pm \infty} \psi = 2n_+ \pi , \quad \lim_{x \to \pm \infty} \chi = 2n_- \pi .
\]

2.4 Travelling wave solutions and friction

It is well known that the standard Y model (i.e. without friction) supports kink soliton solutions. In particular, in the symmetric sector these are the standard kink solutions of the sine-Gordon equation \([4, 31, 32]\) and in the anti-symmetric sector these are the standard kink solutions of the double sine-Gordon equation \([4, 33, 34]\). Here we briefly discuss, for the convenience of the reader, travelling wave solutions for the continuum Yakushevich model, taking into account the physical limit conditions discussed above.

If we require

\[
\psi(x,t) = \psi(x - ct) , \quad \chi(x,t) = \chi(x - ct)
\]

and write \( z := (x - ct) \), the equations (7) – adding also the friction terms, for the sake of generality – become

\[
\begin{align*}
\mu \psi_{zz} &= - \kappa_p \sin \psi \cos \chi + \lambda c \psi_z ; \\
\mu \chi_{zz} &= - \kappa_p \sin \chi (\cos \psi - \cos \chi) + \lambda c \chi_z .
\end{align*}
\]

We have written here

\[
\mu := c^2 - \kappa_s ;
\]

this quantity may be positive or negative depending on the speed \( c \) of the travelling wave. The equations (13) may be reinterpreted as the two-dimensional motion of a particle of unit mass in the effective potential

\[
W(\psi, \chi) = \frac{\kappa_s}{\mu} (\cos^2 \chi - 2 \cos \psi \cos \chi)
\]

and with effective friction coefficient

\[
\sigma(c) = - \frac{\lambda c}{\mu} .
\]

For \( \mu > 0 \) there are two obstacles to the existence of travelling wave solutions complying with the physical (i.e. finite energy) limit conditions: \( a \) friction is negative; \( b \) the rest points are minima and thus different ones cannot be joined by a trajectory admitting them as limit points. Note that albeit we can get rid of \( a \) if we are able to disregard friction, the obstacle \( b \) still makes that only trivial wave solutions – i.e. constant ones, taking values in one of the minima – are possible.

On the other hand, for \( \mu < 0 \) both problems disappear. It should be stressed that in the presence of friction we can still not have trajectories joining two different rest points (as these all correspond to the same energy), but this problem can be solved if we have some part of the system providing energy from outside. This will be the role of RNAP, providing energy through ATP depletion.
2.5 Stability of frictionless travelling waves

As said before, if no friction is present (i.e. $\lambda = 0$, hence also $\sigma = 0$), the existence of symmetric and anti-symmetric travelling wave solutions is known. Explicitly, in the sine-Gordon equation, for $c^2 < \kappa_s$, there are symmetric travelling wave solutions given by

$$\psi(z) = 4 \arctan \left[ \exp \left( \frac{z \sqrt{\kappa_p}}{\sqrt{\kappa_s - c^2}} \right) \right]$$

and $\chi(z) = 0$.

In the double sine-Gordon equation, the anti-symmetric travelling waves are

$$\psi(z) = 0$$

and

$$\chi(z) = \pi + 2 \arctan \left( \frac{z \sqrt{\kappa_p}}{\sqrt{\kappa_s - c^2}} \right).$$

The study of sine-Gordon and, respectively, double sine-Gordon equations per se guarantees their soliton solutions are stable [4, 31, 32, 33, 34]. In the context of the present model, this means that the soliton solutions obtained when restricting to the symmetric, respectively antisymmetric, sector are stable against perturbations in the same sector – i.e. with the same symmetry properties.

We should also consider the stability of those waves in the full system (7), i.e. also test stability of solutions with a given symmetry also against perturbations with a different symmetry.

First we look at the linear stability of the fixed points $(2n\pi, 0)$, which also gives information about the continuous spectrum. We write

$$\left( \begin{array}{c} \psi \\ \chi \end{array} \right)(x,t) = \left( \begin{array}{c} 2n\pi \\ 0 \end{array} \right) + \left( \begin{array}{c} \delta \psi \\ \delta \chi \end{array} \right)e^{ikx + \alpha t}$$

and linearize. This gives

$$\begin{align*}
\alpha^2 \delta \psi &= -k_s k^2 \delta \psi - \kappa_p \delta \psi \\
\alpha^2 \delta \chi &= -k_s k^2 \delta \chi
\end{align*}$$

hence

$$\begin{align*}
0 &= (\alpha^2 + k_s k^2 + \kappa_p) \delta \psi \\
0 &= (\alpha^2 + k_s k^2) \delta \chi
\end{align*}$$

So we can conclude that the eigenvalues are $\alpha = \pm i \sqrt{\kappa_p + k_s k^2}$ and $\alpha = \pm i |k| \sqrt{\kappa_s}$. Thus the continuous spectrum is the full imaginary axis, with double branches in $(-i \sqrt{\kappa_p}, i \sqrt{\kappa_p})$ and fourfold ones outside this interval. The continuous spectrum associated with the anti-symmetric fixed points $(0, 2m\pi)$ is exactly the same, as can be shown in a similar way.

To analyse the linear stability of the travelling fluxons, first we write

$$\tilde{z} = \frac{z \sqrt{\kappa_p}}{\sqrt{\kappa_s - c^2}}$$

and drop the tilde for the remainder of this section. With this new spatial coordinate, the system (7) becomes

$$\begin{align*}
\psi_{tt}/\kappa_p &= \psi_{zz} - \sin \psi \cos \chi; \\
\chi_{tt}/\kappa_p &= \chi_{zz} - \sin (\cos \psi - \cos \chi).
\end{align*}$$

2.5.1 Symmetric travelling waves

First we consider the symmetric travelling wave. Denoting the sine-Gordon fluxon by $\psi_\text{fl}$, i.e., $\psi_{\text{fl}}(z) = 4 \arctan(e^z)$, using

$$\left( \begin{array}{c} \psi \\ \chi \end{array} \right)(z,t) = \left( \begin{array}{c} \psi_{\text{fl}}(z) \\ 0 \end{array} \right) + \left( \begin{array}{c} \delta \psi(z) \\ \delta \chi(z) \end{array} \right)e^{\alpha t},$$

and linearizing around $\psi_{\text{fl}}$, we get

$$\mathcal{M}_1 \Phi = \frac{\alpha^2}{\kappa_p} \Phi,$$  

\(\text{(17)}\)

\(5\) It seems this point is generally overlooked when analysing Y-type DNA models.
with
\[ \mathcal{M}_1 = \begin{pmatrix} L_1 & 0 \\ 0 & L_1 + 1 \end{pmatrix}, \quad \Phi = \begin{pmatrix} \delta \phi \\ \delta \chi \end{pmatrix}, \quad L_1 = D_{zz} - \cos(\psi_H(z)). \]

The operator \( L_1 \) is the linear operator associated with the linearization about the fluxon \( \psi_H \) in the sine-Gordon equation. It has only one eigenvalue, which is the eigenvalue 0 with the eigenfunction \( \psi_H'(z) \), and its continuous spectrum is \((-\infty, -1]\), see e.g. [35]. The eigenvalue problem (17) can be written as two uncoupled eigenvalue problems for \( L_1 \):

\[ \begin{align*}
0 &= \left[ L_1 - \alpha^2 / \kappa_p \right] \delta \psi \\
0 &= \left[ L_1 - (\alpha^2 / \kappa_p - 1) \right] \delta \chi.
\end{align*} \]

Using that the eigenvalue 0 is the only eigenvalue for \( L_1 \), it follows immediately that the only eigenvalues for \( \mathcal{M}_1 \) are \( \alpha = 0 \) and \( \alpha = \pm \sqrt{\kappa_p} \). The eigenvalue \( \alpha = 0 \) is embedded in the continuous spectrum (see above) and has the eigenfunction \( (\psi_H'(z), 0)^T \). The eigenvalues \( \alpha = \pm \sqrt{\kappa_p} \) have eigenfunctions \( (0, \psi_H'(z))^T \).

Hence the symmetric travelling waves are linearly stable against symmetric perturbations, but unstable against anti-symmetric perturbations. The initial effect of the instability is that in the original coordinates \( \phi_+ \) and \( \phi_- \), the soliton splits and the \( \phi_+ \) travelling wave starts travelling faster and the \( \phi_- \) slower (or the other way around). This is confirmed by a numerical simulation, see section 5. In other words, the symmetric solitons are linearly stable against symmetric perturbations, but present a direction of linear instability in the antisymmetric sector, and are hence unstable against perturbations having any component in this direction.

At the nonlinear level, numerical simulations show that this instability leads, if excited, to a speeding up of the \( \phi^{(+)} \) soliton accompanied by the creation of a small “wiggle” lying in the \( \phi^{(-)} \) component and a slowing down of the \( \phi^{(-)} \) soliton accompanied by the creation of a small “wiggle” lying in the \( \phi^{(+)} \) component (see Figure 9).

### 2.5.2 Anti-symmetric travelling waves

Next we consider the anti-symmetric travelling wave. Denoting the double sine-Gordon fluxon by \( \chi_H \), i.e., \( \chi_H(z) = \pi + 2 \arctan(z) \), using

\[
\begin{pmatrix} \psi \\ \chi \end{pmatrix} (z,t) = \begin{pmatrix} 0 \\ \chi_H(z) \end{pmatrix} + \begin{pmatrix} \delta \psi(z) \\ \delta \chi(z) \end{pmatrix} e^{\alpha t},
\]

and linearizing, we get
\[
\mathcal{M}_2 \Phi = \frac{\alpha^2}{\kappa_p} \Phi, \tag{18}
\]

with
\[
\mathcal{M}_2 = \begin{pmatrix} L_2 - 1 & 0 \\ 0 & L_3 \end{pmatrix}, \quad \Phi = \begin{pmatrix} \delta \phi \\ \delta \chi \end{pmatrix}, \quad L_2 = D_{zz} + \frac{2}{1 + z^2}, \quad L_3 = D_{zz} - \frac{2(3z^2 - 1)}{(1 + z^2)^2}.
\]

The operator \( L_3 \) is the linear operator associated with the linearization about the double sine-Gordon fluxon in the double sine-Gordon equation. As the double sine-Gordon fluxon is invariant under spatial translations, it follows immediate the operator \( L_3 \) has an eigenvalue zero with eigenfunction \( \chi_H' \). Since the eigenfunction \( \chi_H' \) has no zeroes, we can conclude that zero is the largest eigenvalue for \( L_3 \) as \( L_3 \) is a Sturm-Liouville operator. It is easy to see that the continuous spectrum for both \( L_2 \) and \( L_3 \) is \((-\infty, 0]\). Hence the eigenvalue zero for \( L_3 \) is its only eigenvalue. The operator \( L_2 \) doesn’t have a direct association with the double sine-Gordon equation and is less well-known.

The eigenvalue problem (18) can be written as two uncoupled eigenvalue problems for \( L_2 \) and \( L_3 \):

\[ \begin{align*}
0 &= \left[ L_2 - (1 + \alpha^2 / \kappa_p) \right] \delta \psi \\
0 &= \left[ L_3 - \alpha^2 / \kappa_p \right] \delta \chi.
\end{align*} \]
The eigenvalue zero for $L_3$ corresponds to an eigenvalue $\alpha = 0$ for $M_2$ with eigenfunction $\Phi = (0, x_0(z))$. A numerical analysis of the eigenvalues of $L_2$ shows an eigenvalue for $L_2$ at approximately 1.084 with an eigenfunction with no zeroes (so it is the largest eigenvalue). Hence this implies that $M_2$ has an eigenvalue $\alpha = \pm 0.29\sqrt{\kappa\rho}$ and the anti-symmetric travelling waves are linearly unstable against symmetric perturbations. As the unstable eigenvalue is significantly smaller than in the symmetric case, it can be expected that the instability takes longer to manifest itself. This is confirmed by the numerics in section 5, see Figure 9.

3 DNA dynamics in the presence of RNAP

We want to consider a model of the coupled dynamics of DNA (within the framework set by the Yakushevich model) and of the RNAP, supposed to be already binding to the DNA double helix. That is, we do not discuss the process by which RNAP binds to DNA, but focus on the dynamics of the coupled system once the RNAP has bound. In order to do this, we need to introduce in our framework a description of two physical phenomena:

(A) the presence of RNAP affects the internal motion of DNA in the binding region and causes its opening;
(B) the motion of RNAP takes place by pulling on the DNA molecule and is resisted by friction forces.

The first phenomenon will be described by a local potential at the given (time-dependent) location of RNAP binding along the double helix. The description of the second phenomenon requires to consider an equation describing the RNAP location and its motion, and the balance between friction – due to the cellular fluid in which RNAP is moving and to dissipative effects in the RNAP internal motions – and the energy available to power the motion via ATP depletion. In the simple model we are going to consider, RNAP pulls on DNA along the double helix axis, so that – given that here we consider DNA as non-deformable in such direction – the opening of DNA is only due to the local potential mentioned above and describing the local DNA/RNAP interaction.

In order to simplify our discussion and focus on the mechanism of DNA/RNAP interaction, in this section we will only introduce dissipation acting on the RNAP in our model at first, i.e. in this section, not dissipation action on the DNA. This will be discussed in the following section.

Moreover, we will consider for the sake of simplicity the case where a single RNAP is travelling along the double helix; it is quite clear that the model built below can as well account for the case where several RNAP are travelling at the same time along the same DNA double helix – provided these are sufficiently far apart, i.e. not interacting either directly or though DNA deformations.

3.1 Interaction with RNAP. Model A: DNA opening

In the simplest description of DNA/RNAP interaction, RNAP is binding to DNA and sliding along it with a given motion (the nature of this motion will be taken into account below). In this section, we focus on describing how DNA reacts to the presence – and motion – of RNAP.

When RNAP binds to DNA, it actually binds to a specific section of it, i.e. a certain number of consecutive bases. We describe the position of RNAP along the chain via the position of its center (or the center of the binding region) at time $t$; this would be an integer $h = h(t)$ in the discrete description of section 2.1, or a real variable $\xi(t)$ in the continuum description of section 2.2. The binding will be through bases which are at positions at a distance less than $d = k\delta$ from the center of the RNAP, i.e., such that $|n - h| \leq k$ (discrete description) or $|x - \xi| \leq d$ (continuum description). We will thus describe the DNA/RNAP interaction through a potential $W$ of the form $W = W_0 \cdot R$; here $W_0$ describes the interaction and $R$ is a “coupling factor” being one in the binding region and decreasing to zero outside it. The simplest specification, which we will adopt, for this is

$$R = \begin{cases} 
1, & \text{when inside the binding region} \\
0, & \text{when outside the binding region} 
\end{cases}$$
In terms of the discrete description, the position of RNAP is given by an integer \( h = h(t) \) and the binding region extends over \((2k + 1)\) bases; the coupling factor is

\[
R(h, n) = \Theta \left[ k^2 - (h - n)^2 \right],
\]

with \( \Theta \) the Heaviside function. As for the coupling potential \( W_0(\phi) \), this describes the effect caused on DNA by the binding to RNAP. This potential should keep the DNA open, and thus it should be a potential with a minimum in \( \phi = \pi \) (and in other angles \( \phi = (2k + 1)\pi \), of course). In the simplest case, this is given by

\[
W(\phi) = K_r \cos \phi.
\]

Thus our Lagrangian \( L \) is now

\[
L = \hat{T} - \hat{V}_p - \hat{V}_s - \hat{W},
\]

and the discrete Euler-Lagrange equations (4) are correspondingly

\[
\begin{align*}
I \ddot{\psi}_n &= K_s (\psi_{n+1} - 2\psi_n + \psi_{n-1}) - K_p r^2 \sin \psi_n \cos \chi_n \\
&\quad + K_r R(h, n) \sin \psi \cos \chi; \\
I \ddot{\chi}_n &= K_s (\chi_{n+1} - 2\chi_n + \chi_{n-1}) - K_p r^2 \sin \chi_n (\cos \psi_n - \cos \chi_n) \\
&\quad + K_r R(h, n) \sin \chi \cos \psi. 
\end{align*} \tag{19}
\]

In the continuum version, the position of RNAP is described by \( \xi \). For the time being, we consider \( \xi \) as a given function of \( t \), \( \xi = \xi(t) \). The coupling factor is

\[
R(\xi, x) = \Theta \left[ d^2 - (\xi - x)^2 \right],
\]

hence the interaction of the RNAP will be with the fields \( \psi \) and \( \chi \) in the interval \( D = [\xi - d, \xi + d] \). The equations (19) now yield

\[
\begin{align*}
\psi_{tt} &= \kappa_s \psi_{xx} - \kappa_p \sin \psi \cos \chi + \kappa_r R(\xi, x) \sin \psi \cos \chi; \\
\chi_{tt} &= \kappa_s \chi_{xx} - \kappa_p \sin \chi (\cos \psi - \cos \chi) + \kappa_r R(\xi, x) \sin \chi \cos \psi. 
\end{align*} \tag{20}
\]

where \( \kappa_r := (K_r / I) \).

Note that again the symmetric and antisymmetric sectors are invariant under this dynamics. In these sections, the equations (20) are reduced to respectively

\[
\begin{align*}
\psi_{tt} &= \kappa_s \psi_{xx} - [\kappa_p - \kappa_r R(\xi, x)] \sin \psi \tag{21} \\
\chi_{tt} &= \kappa_s \chi_{xx} - [\kappa_p (1 - \cos \chi) - \kappa_r R(\xi, x)] \sin \chi. \tag{22}
\end{align*}
\]

### 3.2 Interaction with RNAP. Model B: RNAP pulling

In the discussion above, RNAP was binding to DNA and sliding along it with a prescribed motion; we only aimed at describing how DNA reacts to the presence – and motion – of RNAP. In reality, RNAP moves by pulling on the DNA chain\(^6\), so that RNAP motion along the DNA double helix and the motion of the DNA double helix itself are part of the same process. In this section we will describe a “minimal” model describing this interaction and thus providing a dynamical description of these coupled motions; we will not enter into the detail of ATP supply and depletion, i.e. just suppose ATP is available in sufficient and constant quantity, and RNAP can move by exerting a constant force.

Moreover, the pulling of RNAP on DNA will not directly affect DNA (torsional) motions: the exerted pulling force will be in the direction of the double helix axis, and we assume DNA to be non deformable in that direction. Thus, again denoting by \( \xi(t) \) the \( x \) coordinate of the position of the center of the RNAP at time \( t \), the effect of RNAP motion on DNA torsional deformations will be due to an interaction potential as above.

\(^6\)The motion being permitted by “chemical motors” based on the depletion of ATP; RNAP motion and transcription stops if ATP is not available.
Leaving apart the friction forces acting against DNA torsional motions (to be introduced in the following section), but taking into account those acting against the motion of RNAP, our full model is described by

\[
\begin{align*}
\psi_t &= \kappa_s \psi_{xx} - \kappa_p \sin \psi \cos \chi + \kappa_r R(\xi, x) \sin \psi \cos \chi; \\
\chi_t &= \kappa_s \chi_{xx} - \kappa_p \sin \chi (\cos \psi - \cos \chi) + \kappa_r R(\xi, x) \sin \chi \cos \psi; \\
\xi_t &= P - \nu \xi_t.
\end{align*}
\]

(23)

Here \( P \) is the pulling force, which is determined by chemical mechanisms and depends on the ATP supply.\(^7\)

In this model, the dynamics of RNAP is effectively decoupled from the (torsional) one of DNA, and one can solve the last equation in (23) to get asymptotically

\[
\xi_t = P/\nu := c, \quad \text{hence} \quad \xi(t) = ct + \xi_0
\]

(24)

(at least when ATP supply is constant and large enough). Thus asymptotically, the RNAP moves with constant speed along the DNA. Substituting the relation for \( \xi(t) \) into the first two equations, gives effectively back model A described by (20), together with a relation for the function \( \xi(t) \).\(^8\) We will thus from now mainly work on model A as discussed in the previous section, i.e. (20) together with the relation \( \xi(t) = ct + \xi_0 \).

3.3 Travelling wave solutions in a moving frame

We are mainly interested in the case where the RNAP is moving at a constant speed \( c \). In this case it is quite natural to set our equations in a co-moving frame, i.e. in a frame in which the RNAP is stationary; it will be convenient to choose the one with origin coinciding with the center of the RNAP, so that \( \xi \equiv 0 \) in the new variables.

We will write, as before, \( z = (x - ct) \); we also require that \( \psi = \psi(z), \chi = \chi(z) \). In the new variables the coupling factor \( R \) is written as

\[
\rho(z; d) = \Theta (d^2 - z^2).
\]

(25)

In this way, and with the notation (14), the equations (20) of model A read

\[
\begin{align*}
\mu \psi_{zz} &= - \kappa_p \sin \psi \cos \chi + \kappa_r \rho(z; d) \sin \psi \cos \chi; \\
\mu \chi_{zz} &= - \kappa_p \sin \chi (\cos \psi - \cos \chi) + \kappa_r \rho(z; d) \sin \chi \cos \psi.
\end{align*}
\]

(26)

If we specifically look for travelling wave solution in the laboratory (DNA) frame, co-moving with RNAP, this amounts to looking for stationary solutions in the moving frame. In view of (25), the travelling wave problem (26) can be decomposed into three problems in three distinct regions of the real line,

\[
R_- = \{z < -d\}, \quad R_0 = \{-d \leq z \leq d\}, \quad \text{and} \quad R_+ = \{z > d\}.
\]

In the regions \( R_+ \) and \( R_- \) we have

\[
\begin{align*}
\mu \psi_{zz} &= - \kappa_p \sin \psi \cos \chi; \\
\mu \chi_{zz} &= - \kappa_p \sin \chi (\cos \psi - \cos \chi).
\end{align*}
\]

(27)

(boundary conditions will of course be different in \( R_+ \) and in \( R_- \)). In the region \( R_0 \) we have

\[
\begin{align*}
\mu \psi_{zz} &= -(\kappa_p - \kappa_r) \sin \psi \cos \chi; \\
\mu \chi_{zz} &= - \kappa_p \sin \chi (\cos \psi - \cos \chi) + \kappa_r \sin \chi \cos \psi.
\end{align*}
\]

(28)

We recall ATP is also needed for the synthesis of RNA-messenger; moreover, as the copying mechanism has an optimal speed, it has to be expected that \( P \) depends on the ATP supply \( \alpha \) so to yield a saturation of pulling force at some value \( P_0 \) when ATP is widely available. That is, one should expect \( P(\alpha) \approx P_0 \alpha^k/(1 + \alpha^k) \) for some constant \( k > 0 \).

\(^7\) Actually, the full model described by (23) would be of interest only when coupled to a (spatiotemporal) model of ATP supply to RNAP; this is however well beyond the scope of the present paper, and we pointed out this schematization only as a suggestion for further developments.
The limit behavior for \( z \to \pm \infty \) only depends on the equations in the \( R_{\pm} \) regions, which are the same as discussed in section 2. For this reason, we will from now on assume
\[
\mu < 0 . \tag{29}
\]

As before, the symmetric sector \( \chi \equiv 0 \) and anti-symmetric sector \( \psi \equiv 0 \) are still invariant. The dynamics in the symmetric sector reduces to
\[
\begin{align*}
\mu \psi_{zz} &= -\kappa_p \sin \psi \quad (z \in R_{\pm}) ; \\
\mu \psi_{zz} &= -(\kappa_p - \kappa_r) \sin \psi \quad (z \in R_0) . \tag{30}
\end{align*}
\]
Similarly, in the antisymmetric sector \( \psi \equiv 0 \) the equations read
\[
\begin{align*}
\mu \chi_{zz} &= -\kappa_p (1 - \cos \chi) \sin \psi \quad (z \in R_{\pm}) ; \\
\mu \chi_{zz} &= -[\kappa_p (1 - \cos \chi) - \kappa_r] \sin \chi \quad (z \in R_0) . \tag{31}
\end{align*}
\]

### 3.4 Models with dissipative forces acting on DNA

The discussion so far is conducted without considering friction forces acting on the DNA, i.e. without dissipative effects due to interaction of the DNA with the cellular fluid. We will now introduce these, proceeding as in section 2.

Introducing friction forces opposing the motion of the \( \phi^\pm \), passing to the \( \psi, \chi \) variables and going to the continuum limit, the equations (20) of model A become
\[
\begin{align*}
\psi_t &= \kappa_s \psi_{xx} - \kappa_p \sin \psi \cos \chi + \kappa_r R(\xi, x) \sin \psi \cos \chi - \lambda \psi_t ; \\
\chi_t &= \kappa_s \chi_{xx} - \kappa_p \sin \psi (\cos \psi - \cos \chi) + \kappa_r R(\xi, x) \sin \chi \cos \psi - \lambda \chi_t . \tag{32}
\end{align*}
\]
As for the travelling wave reduced equations (21) and (22), these read now
\[
\begin{align*}
\mu \psi_{zz} &= - (\kappa_p - \kappa_r) \sin \psi \cos \chi + \lambda c \psi ; \\
\mu \chi_{zz} &= - \kappa_p \sin \psi (\cos \psi - \cos \chi) + \rho \kappa_r \sin \chi \cos \psi + \lambda c \chi . \tag{33}
\end{align*}
\]
These equations can be interpreted as the two-dimensional motion of a particle of unit mass in an effective potential, subject to a friction force with friction coefficient \( \sigma = -(\lambda c/\mu) \) (positive as we assumed \( \mu < 0 \)); this is obtained rewriting the above as
\[
\begin{align*}
\psi_{zz} &= - \mu^{-1} (\kappa_p - \kappa_r) \sin \psi \cos \chi - \sigma \psi \\
\chi_{zz} &= - \mu^{-1} [\kappa_p \sin \chi (\cos \psi - \cos \chi) - \rho \kappa_r \sin \chi \sin \psi] - \sigma \chi . \tag{34}
\end{align*}
\]
In the symmetric sector, we get simply
\[
\psi_{zz} = - \mu^{-1} (\kappa_p - \kappa_r) \sin \psi - \sigma \psi , \tag{35}
\]
while in the anti-symmetric sector, we get
\[
\chi_{zz} = - \mu^{-1} [(\kappa_p - \kappa_r) \sin \chi - \kappa_p \sin \chi \cos \chi] - \sigma \psi . \tag{36}
\]
Note that \( \sigma \) depends on the speed \( c \) and the larger \( c \) gets, the larger the effective friction coefficient.

A similar analysis holds for the full model B. The equations (23) become, introducing friction terms for the DNA motion,
\[
\begin{align*}
\psi_t &= \kappa_s \psi_{xx} - \kappa_p \sin \psi \cos \chi + \kappa_r R(\xi, x) \sin \psi \cos \chi - \lambda \psi_t ; \\
\chi_t &= \kappa_s \chi_{xx} - \kappa_p \sin \chi (\cos \psi - \cos \chi) + \kappa_r R(\xi, x) \sin \chi \cos \psi - \lambda \chi_t ; \\
\xi_t &= P - \nu \xi_t . \tag{37}
\end{align*}
\]
The travelling wave reduction is
\[
\begin{align*}
\psi_{zz} &= \mu^{-1} [-\kappa_p \sin \psi \cos \chi + \kappa_r \rho \sin \psi \sin \chi] - \sigma \psi ; \\
\chi_{zz} &= \mu^{-1} [-\kappa_p \sin \chi (\cos \psi - \cos \chi) + \kappa_r R(\xi, x) \sin \chi \sin \psi] - \sigma \chi ; \\
\xi_t &= P - \nu \xi_t . \tag{38}
\end{align*}
\]
In the symmetric sector, these reduce to
\[
\begin{align*}
\psi_{zz} &= -\mu^{-1}(\kappa_p - \kappa_r \rho) \sin \psi - \sigma \psi_z; \\
\xi_{tt} &= P - \nu \xi_t.
\end{align*}
\]
In the antisymmetric sector, we get
\[
\begin{align*}
\chi_{zz} &= \mu^{-1} [-\kappa_p \sin \chi (1 - \cos \chi) + \kappa_r R(\xi, x) \sin \chi] - \sigma \chi_z; \\
\xi_{tt} &= P - \nu \xi_t.
\end{align*}
\]
As in the previous section, for constant ATP supply and hence constant pulling force \( P \), the equations for the full model reduce asymptotically – after a transient – to those for the simpler model A. In particular, equation (38) reduces to (34).

4 Existence of travelling wave solutions and kinks

In analogue to the standard Y model which supports kink soliton solutions in both the symmetric and anti-symmetric sector, we will now show that our model for the dynamics of DNA in interaction with RNAP also supports soliton solutions\(^9\) in those sectors, and characterize these. First we pass to a moving frame and introduce the spatial coordinate
\[
\tilde{z} = \sqrt{\frac{\kappa_p}{\kappa_s - c^2}} (x - ct - \xi_0)
\]
and the functions \( \psi(x, t) = \tilde{\psi}(\tilde{z}, t) \) and \( \chi(x, t) = \tilde{\chi}(\tilde{z}, t) \); we drop the tilde from now on for ease of notation. With these new coordinates, equation (35) in the symmetric sector becomes
\[
\psi_{zz} = (1 - \rho(z, \Delta) q_r) \sin \psi - s \psi_z,
\]
while in the anti-symmetric sector, equation (36) becomes
\[
\begin{align*}
\chi_{zz} &= (1 - \rho(z, \Delta) q_r) \sin \chi - \sin \chi \cos \chi - s \psi_z, \\
\end{align*}
\]
where \( \Delta = \sqrt{\frac{\rho_c}{\rho_s - c^2}} \), \( q_r = (\rho / \rho_s) \), and \( s = (\lambda c / \sqrt{\kappa_p (\kappa_s - c^2)}) \). Using the definition of the step function \( \rho \), see (25), we can write the equations as
\[
\psi_{zz} = \begin{cases} 
\sin \psi - s \psi_z, & \text{for } z < -\Delta \\
(1 - q_r) \sin \psi - s \psi_z, & \text{for } -\Delta \leq z \leq \Delta \\
\sin \psi - s \psi_z, & \text{for } \Delta < z 
\end{cases} \quad (R_-) \quad (43)
\]
and
\[
\chi_{zz} = \begin{cases} 
\sin \chi - \sin \chi \cos \chi - s \psi_z, & \text{for } z < -\Delta \\
(1 - q_r) \sin \chi - \sin \chi \cos \chi - s \psi_z, & \text{for } -\Delta \leq z \leq \Delta \\
\sin \chi - \sin \chi \cos \chi - s \psi_z, & \text{for } \Delta < z 
\end{cases} \quad (R_+) \quad (44)
\]
We stress that these equations are defined in \( L^2(\mathbb{R}) \) if \( \psi_{xx}, \chi_{xx} \in H^2(\mathbb{R}) \), which implies that \( \psi, \chi \) are \( C^1(\mathbb{R}) \); we do not have to require any higher regularity.

Furthermore, (43) and (44) describe the motion of a particle of unit mass (with this description, \( z \) should be thought as an “effective time”), subject to a friction with coefficient \( s \), in the potential
\[
W_+(\psi) = \begin{cases} 
\cos \psi - 1 & \text{in regions } R_+ \\
(1 - q_r) (\cos \psi - 1) & \text{in region } R_0
\end{cases}
\]
and
\[
W_-(\chi) = \begin{cases} 
\cos \chi - \cos^2 \chi/2 - 1/2 & \text{in regions } R_+ \\
(1 - q_r) (\cos \chi - 1) - \cos^2 \chi/2 + 1/2 & \text{in region } R_0
\end{cases}
\]
in the symmetric case, and
\[
W_-(\chi) = \begin{cases} 
\cos \chi - \cos^2 \chi/2 - 1/2 & \text{in regions } R_+ \\
(1 - q_r) (\cos \chi - 1) - \cos^2 \chi/2 + 1/2 & \text{in region } R_0
\end{cases}
\]
in the anti-symmetric case.

\(^9\)Or more precisely, travelling solitary wave solutions.
4.1 The case with no DNA friction

Let us first consider the frictionless case, corresponding to \( s = 0 \). Now the motion is described by a Hamiltonian system with Hamiltonian \( H_\pm(p,u) = \frac{1}{2}p^2 + W_\pm(u) \), where “±” refers to the symmetric or antisymmetric case (and potential), and \( u \) stands for \( \psi \) in the symmetric case and \( \chi \) in the antisymmetric one.

In each region \( R_\pm, R_0 \), the Hamiltonian is constant. The kinks decay to a multiple of \( 2\pi \) in the regions \( R_\pm \), hence in those regions we have (with the form of potentials \( W_\pm \) given above) that \( H_\pm(p,u) = 0 \). The value of the Hamiltonian in the RNAP region \( R_0 \) is determined by \( \Delta \), representing the range of influence of the RNAP potential. If we denote the value of the Hamiltonian in \( R_0 \) by \( h \), then the relation between \( h \) and \( \Delta \) is

\[
\Delta = \int_{\gamma(h)} \frac{du}{p(u,h)} \quad \text{with} \quad p^2(u,h) = (h - W_\pm^0(u)),
\]

where \( \gamma(h) \) is the solution curve corresponding to the Hamiltonian value \( h \) and \( W_\pm^0 \) is the potential in \( R_0 \) in the symmetric/anti-symmetric case. The start and end points on the curve \( \gamma(h) \) are found by using that the solutions should be in \( H^2(\mathbb{R}) \), hence in \( C^1(\mathbb{R}) \), leading to matching conditions on the boundaries of the regions \( R_\pm \) and \( R_0 \). More details can be found in [36], where a similar problem is analysed in the case of Josephson junctions with defects.

To analyse this problem, we will vary the value of the Hamiltonian \( h \), and determine the value of \( \Delta \) associated with it. The problem, and the solution approach, can be visualized in the \((u,u_z)\)-phase plane. In the region \( R_- \), the kink decays to 0, hence in the phase plane it is on the unstable manifold of \((u,u_z) = (0,0)\). Similarly, in the region \( R_+ \), the kink decays to \( 2\pi \), hence in the phase plane it is on the stable manifold of \((u,u_z) = (2\pi,0)\). The stable and unstable manifolds of the fixed points \((2k\pi,0)\), will be denoted by red dashed curves in the phase portraits. In the symmetric section, there are two typical cases: \( 0 < q_r < 1 \ (\kappa_p > \kappa_r) \) and \( q_r > 1 \ (\kappa_p < \kappa_r) \), see Figures 2 and 3.

In the anti-symmetric section, there are three typical cases: \( 0 < q_r < 1 \ (\kappa_p > \kappa_r), \ 1 < q_r < 2 \ (\kappa_p < \kappa_r)\), and \( q_r > 2 \ (\kappa_p > 2\kappa_r) \), see Figure 4. In both sectors, the red dashed curve starting at \((0,0)\) has to be followed by the solution when \( z \in R_- \), and in \( R_+ \), the red dashed curve ending at \((2\pi,0)\) has to be followed. The blue dash-dotted curves are potential solution curves for when \( z \in R_0 \). In order to get a \( C^1 \) solution \( u(t) \), the blue and red curves have to intersect in the phase plane as this gives that \( u \) and \( u_z \) are continuous. The points of intersection are the endpoints of the regions.

The question of existence of a kink solution, which is a heteroclinic connection between the singular points in \((u,u_z) = (0,0)\) at \( z \to -\infty \) and \((u,u_z) = (2\pi,0)\) at \( z \to \infty \), can now be rephrased as the existence of a \( C^1 \) curve, made of part of the red dashed ones in the outer regions \( R_\pm \) and of part of a blue dash-dotted one in the RNAP region \( R_0 \) (recall that the kink has to be in \( H^2(\mathbb{R}) \), hence \( C^1 \)). It is quite clear from Figures 2, 3 and 4 that these exist for any length \( \Delta \) as the upper most blue curve corresponds to \( \Delta = 0 \), the lowest blue curve to \( \Delta \to \infty \) and there is a continuum...
we denote by \((\kappa_p = 1\text{ and }\kappa_r = 1.5)\). Again, the red dashed curve corresponds to the dynamics in the outer regions and the dash-dotted blue curves correspond to the dynamics in the inner region \(R_0\). As before, the solid black curve in the middle plot is a possible solution and its shape is given in the \(z\)-\(\psi\) plane on the right.

![Figure 3: Phase portrait in symmetric section for \(q_r > 1\) (\(\kappa_p = 1\text{ and }\kappa_r = 1.5\)).](image)

![Figure 4: Phase portrait in anti-symmetric section for \(q_r = 0.5\) (\(\kappa_p = 1\text{ and }\kappa_r = 0.5\), left), \(q_r = 1.5\) (\(\kappa_p = 1\text{ and }\kappa_r = 1.5\), middle), \(q_r = 3\) (\(\kappa_p = 1\text{ and }\kappa_r = 3\), right).](image)

of blue curves in between (see below for a formal proof). The length \(\Delta = 4d\sqrt{\kappa_p/(\kappa_s - c^2)}\) of the RNAP is given as function of the value of the Hamiltonian \(h\) in the RNAP region \(R_0\) by (47). Figure 5 visualizes this relation in the symmetric resp. anti-symmetric section for typical values of the coupling parameter \(q_r\) (the same values as used in Figures 2-4).

To summarize, we have shown:

*If there is no friction (\(s = 0\)), then for any RNAP range \(d\), any speed \(c\), with \(|c| < \kappa_s\) (\(\kappa_s\) is the constant in the stacking potential), and any coupling strength \(\kappa_r\), there are always a symmetric and a anti-symmetric kink soliton solution to the (frictionless) model \(A\) and hence to the full model \(B\) describing the dynamics of DNA in the presence of RNAP.*

As for the formal proof, with reference to Figures 2-4 we denote by \((u, u_z) = (\pi \pm u_0, p_0)\) the phase plane coordinates of points where the blue and the red separatrices cross. First we note that any blue curve intersecting the red separatrix at a point \((\pi - \hat{u}, p)\) in the region with \(0 < \hat{u} < u_0\), also intersects it at a point \((\pi + \hat{u}, p)\). And the intersection between the two sets of curves is always transversal (for \(\hat{u} = 0\)). If the Hamiltonian value \(h\) of the blue curve approaches the Hamiltonian value of the blue separatrix, the RNAP range \(\Delta(h)\) goes to infinity. And if the Hamiltonian value \(h\) of the blue curve approaches the value of the blue curve through the point on the red separatrix with \(\hat{u} = 0\), then the RNAP range \(\Delta(h)\) goes to zero. Note that transversality easily follows from the explicit expression of the Hamiltonians. In fact for the Hamiltonian in \(R\), we have

\[
\psi_z = \sqrt{2(1 - \cos \psi)} \text{ (symmetric) or } \chi_z = \sqrt{1 - 2\cos \chi + \cos^2 \chi} \text{ (anti-symmetric).}
\]

While in \(R_0\) we have

\[
\psi_z = \sqrt{2(h - q_r \cos \psi)} \text{ (symmetric) or } \chi_z = \sqrt{2h - 2q_r \cos \chi + \cos^2 \chi} \text{ (anti-symmetric).}
\]

Hence

\[
\frac{\partial \psi_z}{\partial \psi} = \begin{cases} \sin \psi/\psi_z , & \text{in } R_+ , \\ q_r \sin \psi/\psi_z , & \text{in } R_0 , \end{cases} \text{ (symmetric sector)}
\]
Figure 5: Sketch of friction-less $h$-$\Delta$ curves, illustrating that a kink exists for every positive RNAP range $\Delta$. The left plot shows two typical curves in the symmetric section: the blue solid line is the case $q_r = 0.5$ ($\kappa_s = 1$ and $\kappa_r = 0.5$) and the red dashed line is for $q_r = 1.5$ ($\kappa_s = 1$ and $\kappa_r = 1.5$). The right plot shows three typical curves in the anti-symmetric section: the blue solid line is the case $q_r = 0.5$ ($\kappa_s = 1$ and $\kappa_r = 0.5$); the red dashed line is for $q_r = 1.5$ ($\kappa_s = 1$ and $\kappa_r = 1.5$) and the green dash-dotted line is for $q_r = 3$ ($\kappa_s = 1$ and $\kappa_r = 3$).

\[
\frac{\partial \chi_z}{\partial \chi} = \begin{cases} \sin \chi (1 - \cos \chi)/\chi_z, & \text{in } R_0^+, \\
\sin \chi (q_r - \cos \chi)/\chi_z, & \text{in } R_0, \end{cases} \quad \text{(anti-symmetric sector)}.
\]

Thus as $q_r \neq 1$, two curves meeting at a point $(\psi, \psi_z)$ or $(\chi, \chi_z)$ are transversal, except at the points with either $\sin \psi = 0$ or $\psi_z = 0$ in the symmetric section or $\sin \chi = 0$ or $\chi_z = 0$ in the anti-symmetric section.

### 4.1.1 Other travelling waves

So far we have focused on travelling kink solutions, in analogue with the solutions present in the standard Y model. However, in the presence of the RNAP, the model also allows for small excitation travelling solitary wave solutions in the symmetric and anti-symmetric sector. The main characteristic of such wave is that it connects to the zero state, both for $z \to \pm \infty$. In the phase portraits, one has to connect the stable and unstable manifolds of the zero state with an orbit from the dynamics in $R_0$.

In the anti-symmetric sector, such waves can always exist if the RNAP is present. In the symmetric sector, they exist if $\kappa_r > \kappa_p$ (or $q_r > 1$). This follows readily from zooming into the areas around the zero state in the phase portraits in Figures 3 and 4, see Figure 6 for some illustrations.

Figure 6: Illustrations of small solitary waves, using the phase portrait representation. On the left is the phase portrait in the symmetric section, similar to the ones in Figure 3 with $q_r = 1.5$ ($\kappa_p = 1$, $\kappa_r = 1.5$). The solid black curve represents a solitary wave solution. Its shape is represented by the dashed line in the middle plot. On the right is the phase portrait in the anti-symmetric section, similar to the ones in the middle plot of 4 with $q_r = 1.5$ ($\kappa_p = 1$, $\kappa_r = 1.5$). As before, the solid black curve represents a solitary wave solution and its shape is represented in the middle plot by the solid black line.
4.2 The case with DNA friction

If friction is included ($s = \lambda c / \sqrt{\kappa_p (\kappa_s - c^2)} > 0$), the systems are not Hamiltonian anymore and the orbits are not restricted to level sets of a Hamiltonian. On the other hand, the phase plane can still be used. A travelling wave kink is a solution which starts on the unstable manifold of $(u, u_z) = (0, 0)$ and ends on the stable manifold of $(u, u_z) = (2\pi, 0)$. The connection between these manifolds is made by an orbit in $R_0$. This idea is also present in the case without friction. Since all intersections between the unstable manifold and the orbit in $R_0$ and the stable manifold and the orbit in $R_0$ are transversal in the case without friction, it follows immediately that for small friction, there still will be kink solutions. This is illustrated in Figures 7 and 8 for the symmetric sector with $s = 0.1$ and $q = 0.5$ ($\kappa_s = 1$ and $\kappa_r = 0.5$).

Figure 7: Phase portrait in the symmetric sector with $s = 0.1$ and $q = 0.5$ ($\kappa_s = 1$ and $\kappa_r = 0.5$). In the phase portrait on the left, the bold solid green curve is the unstable manifold of $(u, u_z) = (0, 0)$ and the bold dashed red curve the stable the stable manifold of $(u, u_z) = (2\pi, 0)$. The blue dash-dotted curves in the middle plot are solutions in $R_0$. The right plot combines the left and middle plot and shows the relevant parts of the stable and unstable manifolds with the band of dash-dotted blue orbits that will form kink solutions.

Figure 8: The bold solid black curve in the left plot of the phase plane is an example of a kink solution. The parameters are the same as in Figure 7; the unstable manifold of $(u, u_z) = (0, 0)$ is the green dotted bold curve, the red dashed and blue dash-dotted curves have the same meaning as in Figure 7. A plot of this solution in the $z$-$\psi$ plane is given on the right.

A comparison with Figure 2 shows the deformation of the stable and unstable manifolds and the orbits in band of solutions that can be used to connect the stable and unstable orbits. As the stable and unstable manifold do not coincide anymore, there will be a strictly positive minimal length $\Delta_{\text{min}}(s)$ for the range of influence of the RNAP that allows for the existence of kinks. Note that $\Delta_{\text{min}} \to 0$ if $s \to 0$. Similar arguments can be used to show the persistence of the travelling waves in all other cases.

5 Simulations and stability of travelling waves

First we look at the instability of the symmetric and anti-symmetric solutions in the Y-model without RNAP as described in section 2.5. If no damping is present, a simulation of those waves with any tiny symmetry-breaking perturbation (e.g. numerical dissipation; though this will not
work if the numerical code is symmetric) in the full system will show the instability. As predicted by the analysis in section 2.5, a symmetry breaking perturbation triggers the unstable eigenmode in the anti-symmetric sector and after while the $\phi^+$ and $\phi^-$ kinks start travelling with different speeds (though the anti-symmetric kink takes longer to destabilise than the symmetric one). This instability can be seen in Figure 9.

![Figure 9](image_url)

Figure 9: The symmetric and anti-symmetric kinks develop an instability when integrated in the full Y-model PDE dynamics with a small symmetry-breaking perturbation (the initial condition for $\phi^{(+)}$ is multiplied with $10^{-7}\sin(x - L/2)$, where $L$ is the length of the integration interval). The solutions are displayed in the moving frame coordinates of the symmetric kink. It can be observed that in both cases the dominant instability mode causes a difference in wave speeds for the $\phi^{(+)}$ and $\phi^{(-)}$ components and the waves start travelling apart leading to a kink in one $\phi$ component with a small excitation on the zero background for the other component. In the two left pictures, the time evolution of both $\phi^{(+)}$ and $\phi^{(-)}$ components of the symmetric kink are shown. The right pictures show the same for the anti-symmetric wave. Note the different timescales in the two cases. The anti-symmetric instability takes longer to develop than the symmetric one.

If damping is present in the Y-model, but no RNAP is present, the travelling kinks induced by the instability cannot sustain themselves and they will stop at a stable steady state which is not fully symmetric nor anti-symmetric. This behaviour can be seen from Figure 10 for the symmetric solution. Here the initial conditions are given by the stationary symmetric solution plus a very small symmetry breaking perturbation. The instability causes the $\phi^{(+)}$ and $\phi^{(-)}$ waves to travel apart, but after a while this motion is stopped by the damping and the waves converge to an a-symmetric stable steady state. We have numerically determined this stable steady state to which both the symmetric and anti-symmetric waves converge. This state is depicted in Figure 11. As we start a simulation with this wave shape in the system without RNAP nor dissipation, the wave shape does not change, even after integrating for a long time.

As we have seen above, the symmetric solution is unstable in absence of the RNAP: the symmetry is broken and the kinks start travelling apart, one to the left, the other to the right. However, if the RNAP is present and sufficiently long and strong ($d = 1$, $\kappa_r = 0.5$ is too short, $d = 5$, $\kappa_r = 0.1$ is too weak but $d = 5$, $\kappa_r = 0.5$ works), the kinks start “bouncing” inside the RNAP. With damping turned on ($\lambda \neq 0$), this leads to a nearly symmetric stable solution formed by a kink in the symmetric plane and a (little) bump in the anti-symmetric one, see Figure 12. We
Figure 10: The symmetric solution develops an instability and the two kinks start travelling apart. Damping is present, so the travelling waves will be stopped and a stable steady state is shown. In the left picture, the time evolution of the $\phi^-$ wave is shown, in the right one the view is for $\phi^+$. This illustrates that the kinks in the $\phi^+$ and $\phi^-$ components are travelling apart and develop a little excitation on the zero background in the other component.

have started with the solution with $c = 0$ in a non-travelling frame, thus in the “moving frame” kink-equation, $\sigma = s = 0$, however in the full system there is damping as $\sigma = -\lambda c/\mu$ and $\lambda \neq 0$. This damping just helps to converge towards a stable steady solution, it doesn’t affect the solution itself.

If we start with the anti-symmetric solution, in first instance its shape doesn’t change much, it just adapts a little bit as predicted by the phase-plane analysis for the presence of RNAP. However, after quite a while, an instability sets in and the wave deforms slightly to a nearly anti-symmetric stable steady state, see Figure 13.

Finally, if we start with the stable a-symmetric stable state as found in the RNAP-less simulations, it undergoes a minor correction (as would be expected from a higher dimensional phase plane argument (intersection of 2D unstable manifold, 2D stable manifold and 4D solution manifold in the middle), see Figure 14. This configuration is very stable (iterated up to $t = 200$).

6 Conclusions and discussion

Back in 1980, Englander, Kallenbach, Heeger, Krumhansl and Litwin [1] conjectured that sine-Gordon solitonic excitations could be present and play a functional role – easing DNA opening and closing in front and behind the travelling RNAP – in DNA transcription. Their proposal led to the formulation and study of a number of mesoscopic Hamiltonian models for DNA dynamics, in which the state of each nucleotide is described by one or a few degrees of freedom. While these models have been rather successful in the context of DNA denaturation [4, 5, 10], they met greater obstacles in the frame of DNA transcription. Two major obstacles, related to DNA inhomogeneity and to the non-Hamiltonian nature of real DNA dynamics, where recently overcome in the frame of mesoscopic models [28, 29]; however, a major shortcoming of previous DNA models supporting topological (sine-Gordon type) solitons lies precisely in the fact these are models of DNA alone, while transcription is intimately linked to the presence of RNAP and to the interaction between this and the DNA molecule.

In this paper we have presented a first, “minimal” model for the DNA/RNAP dynamics. This deals with DNA in the framework of the classical Yakushevich model, and considers the modification in the local DNA dynamics caused by the presence of RNAP; at the same time, the dynamics of RNAP along the DNA chain is described. Thus, our first result was to show that a simple mesoscopic model of the DNA/RNAP complex can be formulated.

We have moreover shown that our model supports sine-Gordon type solitonic excitations travelling along the DNA chain and localized around the RNAP, even in the presence of dissipation.

We also studied the stability of these excitations; we found that they are stable, and actually that the presence of the RNAP improves their stability with respect to the “bare DNA” (original
Figure 11: The stable steady state solution in the full Y-model PDE. On the left the kink in one
component, in the middle the small “wiggle” on the zero background in the other component (note
the difference in scales). On the right the black line is the steady state in the ψ-χ plane, plotted
on top of contour lines of the potential \( W(\psi, \chi) = 2 \cos \psi \cos \chi - \cos \chi^2 \). Note that the stable
steady state is a connection between \((\psi, \chi) = (0, 0)\) and \((\psi, \chi) = (\pi, \pi)\), hence \(0-\pi\) kinks in the
symmetric and anti-symmetric planes. The steady states in the symmetric and anti-symmetric
sectors correspond to lines connecting 0 and \(2\pi\) on the \(\psi\)- resp. \(\chi\)-axis. Thus the stable solution is
a mountain pass connection, while the symmetric and anti-symmetric ones form a connection via
the minimum.

Figure 12: Simulation with the RNAP and dissipation present \((\lambda = 0.1)\), starting with the sym-
metric wave of the Y-model. The RNAP stops the instability from developing fully and the \(\phi^+(+)\)
and \(\phi^(-)(-)\) components stay together in a nearly symmetric fashion as illustrated by the right plot
showing the final steady state in \(\phi^\pm\), \(\psi\) and \(\chi\) coordinates. The left and middle plots show the
time evolution of the \(\phi^+(+)\) resp. \(\phi^(-)(-)\) components for \(\kappa_p = 1, \kappa_r = 0.5, \lambda = 0.1\) and \(c = 0\).

Y model) case. In particular, it was shown that for bare DNA Y model, the symmetric soliton is
linearly stable against symmetric perturbation, but presents a direction of linear instability against
antisymmetric perturbations; and conversely for the antisymmetric soliton. This instability leads
to two separate solitons: a \(\phi^+(+)\) soliton together with a small \(\phi^(-)(-)\) wiggle and a \(\phi^(-)(-)\) soliton together
with a small \(\phi^+(+)\) wiggle. In the moving frame, these solitons travel in opposite directions. Thus
in the physical frame one travels faster, the other slower than the original soliton.

In the case where RNAP is present, this instability get suppressed, and (nearly) symmetric or
(nearly) antisymmetric DNA solitons are globally stable. This new nearly one-component soliton is
stable under presence of the RNAP. In other words, the presence of RNAP stabilizes DNA solitons
and the class of stable DNA solitons is enlarged by the presence of RNAP.

As already mentioned, our model describes DNA in the same terms as the classical Y model and
we only deal with homogeneous DNA, disregarding the differences between the four types of bases.
In real DNA these are not negligible, and are ultimately responsible for the genetic coding. It was
recently shown [25] that a somewhat more detailed mesoscopic description of DNA (via “composite”
models) is able to provide better results and also supports solitonic excitations with more realistic
features – in particular for what concerns the speed of such excitations [26, 28, 29, 30]. Moreover, in
Figure 13: Simulation with the RNAP and dissipation present ($\lambda = 0.1$), starting with the anti-symmetric wave of the Y-model. After a short initial phase the solution seems to settle down to a purely anti-symmetric shape as predicted by the phase portrait analysis from section 4. As predicted, this shape is very close to $\pi + 2 \arctan x$. This is illustrated in the left plot, where the difference between $\chi$ at $t = 100$ and $\pi + 2 \arctan x$ is plotted as the shapes themselves are identical in a plot. However, after an intermediate phase, an instability starts to develop, but the RNAP stops the instability from developing fully and the $\phi^+$ and $\phi^-$ components stay together in a nearly anti-symmetric fashion as illustrated by the right plot showing the final steady state in $\phi(\pm)$, $\psi$ and $\chi$ coordinates. The middle plot shows the time evolution of the $\phi^+$ component for $\kappa_p = 1$, $\kappa_r = 0.5$, $\lambda = 0.1$ and $c = 0$.

Figure 14: Simulation with the RNAP and dissipation present ($\lambda = 0.1$), starting with the asymmetric stable steady wave of the Y-model, see Figure 11. The RNAP only causes a minor modification of this wave. The left and middle plots show the time evolution of the $\phi^+$ resp. $\phi^-$ components for $\kappa_p = 1$, $\kappa_r = 0.5$, $\lambda = 0.1$ and $c = 0$, illustrating the minor modification in the initial stage and the stable state afterwards. The right plot shows the final steady state in $\phi(\pm)$, $\psi$ and $\chi$ coordinates.

The frame of such composite models, solitonic excitations are able to travel long distances along the DNA chain also taking into account the real DNA inhomogeneities [29]. We expect the same would hold for our DNA/RNAP model; that is, a description at the basic level considered here would not be able to support solitons travelling over biologically significant distances in the inhomogeneous case, but going over to a description of the DNA/RNAP complex in which DNA is described as in the composite models [25, 26] and the DNA/RNAP interaction is modelled as in this paper, solitons would be able to travel over significant distances also in the presence of realistic inhomogeneities. In fact, we have shown that the presence of RNAP stabilizes the soliton – so we expect the situation would be even more favourable for solitonic excitations than in the bare DNA case.

Moreover, it is rather clear that generalizations of our simple framework and model should be considered; e.g. we should consider a more realistic framework, allowing also longitudinal DNA deformation. Even more relevant would be to take into account that the pulling will be not exactly along the DNA chain, but at an angle with it, so that it will directly affect DNA opening. Considering such more detailed models or inhomogeneous ones would however go beyond the limits of the present work, in which we wanted to focus on the possibility to model effectively the DNA/RNAP interaction – and their dynamics – in the frame of simple mesoscopic models.
References


[14] L.V. Yakushevich, “Is DNA a nonlinear dynamical system where solitary conformational waves are possible?”, *J. Biosci.* **26** (2001), 305-313


[34] F. Calogero and A. Degasperis, Spectral transform and solitons, North Holland (Amsterdam) 1982
