

Solitary waves in helicoidal models of DNA dynamics

Giuseppe GAETA and Laura VENIER

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

E-mail: gaeta@mat.unimi.it and laura_venier@infinito.it

Received November 6, 2007; Accepted January 3, 2008

Abstract

We analyze travelling solitary wave solutions in the Barbi-Cocco-Peyrard and in a simplified version of the Cocco-Monasson models of nonlinear DNA dynamics. We identify conditions to be satisfied by parameters for such solutions to exist, and provide first order ODEs whose solutions give the required solitary waves; these are not solvable in analytical terms, but are easily integrated numerically.

Introduction

It is conjectured since a long time [13] that soliton-like excitations could be present, and play a functional role, in DNA. They could be relevant both for denaturation (here the relevant excitations would be breathers [17, 20]) and for DNA transcription (here the relevant excitations would be travelling kink solitons [15, 24]).

When trying to model DNA dynamics, one is faced with an extremely complex molecule; the nearly-regular structure (a regular backbone with attached bases, themselves of four possible types with similar properties) suggests to start from modelling a regular (i.e. homogeneous) chain.

In any case, the number of classical degrees of freedom per nucleotide is extremely high. On the other hand, the dynamics of molecules is governed by quantum mechanics, so that a number of degrees of freedom could (and actually will) be frozen at body temperature; one is thus led to identify the softer degrees of freedom and focus attention on these. Such an analysis leads to consider two degrees of freedom, whose activation energies are similar and lower than those of other ones [15].

One of these corresponds essentially to a radial opening of the double helix (that is, the bases in the pairs and hence the two helices separate moving away from the axis of the double helix); the other one corresponds to rotations of the bases (around the sugar-phosphate backbone) in a plane roughly orthogonal to the double helix axis.

Actually, it is quite lucky that in DNA denaturation it is essentially the first of these which is at play, while in DNA transcription the opening of the double chain (to let the RNA Polymerase access the genetic sequence) happens essentially via rotation of the bases so not to disrupt the double helix. It was thus entirely natural to elaborate models taking

into account either one of these two degrees of freedom, depending on the process one was aiming at modelling.

In fact, models for DNA denaturation – in particular, the Peyrard-Bishop (PB) model [18] and improvements thereof [19, 20] – consider a radial degree of freedom; while model dealing with the DNA structure modification met in the transcription process¹ – like the Yakushevich (Y) model [22] and improvements thereof [23, 24] – consider angular degrees of freedom.

As mentioned above, depending on the biological process at hand one is also interested in different nonlinear excitations. In the first case (PB model), one is mainly interested in breathers; while in the second one (Y model) one looks for travelling (topological [12]) solitons.²

Simple DNA models were able to support relevant nonlinear excitations, and quite successful in describing several experimentally measurable quantities associated with the dynamics of the DNA molecule [17, 20, 24]. By the end of the nineties, however, one was able to perform more refined experiments – able to carefully study the dynamics of a single DNA molecule subject to exactly controlled external forces. This called for more detailed models, and an obvious improvement was to consider both the softer degrees of freedom, i.e. the “radial” and the “torsional” one, at the same time.

The first of these models was proposed by Barbi, Cocco and Peyrard [2, 3, 4], and it is at the basis of most of the present-day DNA modelling [17, 20], either in its original form or in the modified version elaborated soon afterwards by Cocco and Monasson [10, 11]. These models were built as a development of the PB model, mainly with the aim of studying DNA denaturation; it is then quite natural that the analysis of nonlinear excitations focused, for the reasons mentioned above, on breathers (see [17, 20] for the results of this analysis). On the other hand, as these models also take into account torsional degrees of freedom, one could hope they are also relevant – as improvements on the Y model – in the context of DNA transcription; this calls for an analysis of travelling solitary waves solutions.

As far as we know, this is not present in the literature; our aim in this paper (which is a development of [21]) is indeed to study if – and under which conditions on the parameters of the model – the BCP and the CM models support travelling solitons. (We will actually consider a simplified version of the CM model, retaining its geometrical features but avoiding some analytical difficulties, see below.)

We would like to mention that a “composite” Yakushevich (cY) model, using two rotation angles per nucleotide, was recently introduced and analyzed [5]. In that case, travelling (topological) solitons are present; see [5] for further details on the model and its solitonic solutions, and [6, 7, 8, 9] for related matters. In this cY model one of the two angles is *de facto* limited to a small range, i.e. cannot go round a complete circle, so that one could think there is a correspondence with the situation described by the BCP or CM models (this was another motivation to us for the present work). It turns out this is not

¹These models are not properly speaking dealing with the transcription process: in the latter, RNA Polymerase and the environment play also a key role. We will however, for ease of writing, also simply say that these model deal with the transcription process.

²Or more precisely, travelling solitary waves: the interaction properties of solitons are not relevant in this context, as solitons would be associated with RNA Polymerase units, which cannot pass one through the other! We will however often say simply “solitons”, rather than “travelling solitary waves”, for ease of writing.

actually the case: our analysis will show the properties of the cY and those of the BCP or CM models are rather different for what concerns solitonic solutions; in particular, in the BCP or CM model the speed of the soliton is a free parameter within a certain range (explicitly determined below), while the cY model presents a remarkable mechanism of selection of the soliton speed [7].

1 The Barbi-Cocco-Peyrard (BCP) model

Let us briefly recall the main features of the Barbi-Cocco-Peyrard (BCP) model, referring to [2, 4, 17] for more detail.

The BCP model describes the DNA double chain as a sequence of units – the nucleotides at sites $n \in \mathbf{Z}$ on each chain $i = \pm 1$ – characterized by two coordinates, which describe their position (say, the position of the center of mass of the nucleotide) in the fixed plane orthogonal to the axis of the double helix in which they are constrained to move. These (polar) coordinates are the distance $r_n^{(i)}$ from the double helix axis and an angle $\varphi_n^{(i)}$ describing the orientation of the base with a given spatial direction. The equilibrium configuration for the chain is a regular double helix.³

If one restricts – as suggested by Barbi, Cocco and Peyrard – to symmetric motions (note the equilibrium configuration is symmetric), this implies $r_n^{(-1)} = r_n^{(1)}$, $\varphi_n^{(-1)} = \varphi_n^{(1)}$. We will from now on adopt this reduction, and write

$$r_n = r_n^{(1)} = r_n^{(-1)}, \quad \varphi_n = \varphi_n^{(1)} = \varphi_n^{(-1)}. \quad (1.1)$$

We will also denote by m the mass of the nucleotides, by h the distance between planes of successive base pairs, by ℓ_n the length of the phosphodiester chain segment linking basis at sites n and $n + 1$ on the same chain, by L the value of this length in the equilibrium configuration; moreover we denote by D , α the parameters entering in the Morse potential

$$V(r_n) := D (\exp[-\alpha(r_n - R_0)] - 1)^2 \quad (1.2)$$

modelling the intrapair interactions mediated by Hydrogen bonding, with R_0 the equilibrium distance between such bases, and finally by K the coupling constant for the harmonic elastic interaction between successive bases (see [2, 4] for details).

With these notation and variables, the BCP model is described by the Lagrangian

$$\begin{aligned} \mathcal{L} = & \sum_n (m \dot{r}_n^2 + m r_n^2 \dot{\varphi}_n^2) - \sum_n D (e^{-\alpha(r_n - R_0)} - 1)^2 + \\ & - \sum_n K (\ell_n - L)^2 - \sum_n G_0 (\varphi_{n+1} + \varphi_{n-1} - 2\varphi_n)^2. \end{aligned} \quad (1.3)$$

It should be stressed that the DNA structure requires that $L > h$; this will be relevant in the following.

The term with coupling constant G_0 is a curvature one, needed to prevent some singular behavior allowed by the discrete nature of the BCP model [2, 4]; as in the following we will study the continuum version of the BCP model, it will be inessential.

³It should be noted that the equilibrium configuration is of course defined up to a global rotation; this feature will show up again later on and be central to our discussion.

We refer to [2, 4, 17] for a discussion of the physical values of the parameters appearing in the model; the values proposed in [2], which we will also use below, are:

$$R_0 = 2\text{\AA}, \quad h = 3.4\text{\AA}, \quad \alpha = 4.45\text{\AA}^{-1}, \quad D = 0.04\text{eV}, \quad K = 1\text{eV}\text{\AA}^{-2}. \quad (1.4)$$

Note that while the interpair distance (the distance between planes in which base pairs at sites n and $(n+1)$ rotate, measured along the double helix axis) $h_n = h$ is a constant in this model (in a variant of this model, due to Cocco and Monasson and considered below, this is not be the case), and so is the equilibrium length L , the lengths ℓ_n are variable. These can be expressed in terms of the coordinates $(r_n, r_{n+1}, \varphi_n, \varphi_{n+1})$ as $\ell_n = \sqrt{h^2 + r_{n-1}^2 + r_n^2 - 2r_{n-1}r_n \cos(\varphi_n - \varphi_{n-1})}$; using this, the Lagrangian \mathcal{L} of (1.3) is rewritten as

$$\begin{aligned} \mathcal{L} = & \sum_n (m \dot{r}_n^2 + m r_n^2 \dot{\varphi}_n^2) - D(e^{-\alpha(r_n - R_0)} - 1)^2 + \\ & - \sum_n K(\sqrt{h^2 + r_{n-1}^2 + r_n^2 - 2r_{n-1}r_n \cos(\varphi_n - \varphi_{n-1})} - L)^2 + \\ & - \sum_n G_0(\varphi_{n+1} + \varphi_{n-1} - 2\varphi_n)^2. \end{aligned} \quad (1.5)$$

2 Travelling wave reduction for the BCP model

It is convenient to pass to the continuum approximation. That is, the arrays $r_n(t)$, $\varphi_n(t)$ will be replaced by (interpolating) fields $R(x, t)$ and $\Phi(x, t)$ such that, with δ the distance between neighboring nucleotides planes, $R(n\delta, t) \approx r_n(t)$, $\Phi(n\delta, t) \approx \varphi_n(t)$. In this way

$$\begin{aligned} r_n(t) &\approx R(n\delta, t), \quad \varphi_n(t) \approx \Phi(n\delta, t); \\ r_{n\pm 1}(t) &\approx R((n\pm 1)\delta, t) \approx R(n\delta, t) \pm \delta R_x(n\delta, t) + (\delta^2/2)R_{xx}(n\delta, t) + O(\delta^3), \\ \varphi_{n\pm 1}(t) &\approx \Phi((n\pm 1)\delta, t) \approx \Phi(n\delta, t) \pm \delta \Phi_x(n\delta, t) + (\delta^2/2)\Phi_{xx}(n\delta, t) + O(\delta^3); \end{aligned} \quad (2.1)$$

we will use this second order approximation (the first order approximation is trivial due to a cancellation) in our discussion.

Using (2.1) and omitting the curvature term (this is legitimate as we are now dealing with a continuum version of the model; see the discussion in [2, 4] for details), the BCP Lagrangian (1.5) is written, to second order in δ and omitting an inessential constant term,

$$\mathcal{L} = m(R_t^2 + R^2\Phi_t^2) - V(R) - K(1-s)(R_x^2 + R^2\Phi_x^2)\delta^2. \quad (2.2)$$

Here we have written $s := L/h$; thus $s > 1$ for the DNA BCP model.

The Euler-Lagrange equations for (2.2) read

$$\begin{aligned} mR_{tt} &= mR\Phi_t^2 - (1-s)K\delta^2(RP_x^2 - R_{xx}) - (1/2)V'(R), \\ mR^2\Phi_{tt} &= -2mRR_t\Phi_t + (1-s)K\delta^2(R^2\Phi_{xx} + 2RR_x\Phi_x). \end{aligned} \quad (2.3)$$

It will be convenient to introduce

$$\gamma := K(s-1)\delta^2; \quad (2.4)$$

note that for $L > h$ and hence $s > 1$ we have $\gamma > 0$. With this notation, the (2.3) read

$$\begin{aligned} mR_{tt} &= mR\Phi_t^2 - \gamma(R\Phi_x^2 - R_{xx}) - (1/2)V'(R), \\ mR^2\Phi_{tt} &= -2mRR_t\Phi_t + \gamma(R^2\Phi_{xx} + 2RR_x\Phi_x). \end{aligned} \quad (2.5)$$

These are the Euler-Lagrange equations describing general solutions to the field theoretic Lagrangian (1.5) at order δ^2 . It can be easily checked [21] that these are also obtained by first deriving the Euler-Lagrange equations for the discrete BCP Lagrangian (1.5) and then passing to the same continuum approximation via (2.1).

2.1 Boundary conditions

The equations (2.5) are PDEs, and as such they should be supplemented with suitable boundary conditions identifying the function space in which they are set – i.e. physically acceptable solutions.

The natural physical requirement is that solutions are C^∞ and go to an equilibrium configuration for large $|x|$ and for all t ; this will also imply *finite energy*. In terms of R , this means both R_x and R_t should go to zero for $x \rightarrow \pm\infty$. As for Φ , note that \mathcal{L} does not depend on Φ (but only on its derivatives), and the equilibrium configuration (an helix) corresponds to Φ_x a nonzero constant.

In other words, we require (with χ_\pm constants)

$$\begin{aligned} \lim_{x \rightarrow \pm\infty} R_x(x, t) &= 0, & \lim_{x \rightarrow \pm\infty} R_t(x, t) &= 0; \\ \lim_{x \rightarrow \pm\infty} \Phi_x(x, t) &= \chi_\pm, & \lim_{x \rightarrow \pm\infty} \Phi_t(x, t) &= 0. \end{aligned} \quad (2.6)$$

2.2 Travelling waves

We are specially interested in travelling wave solutions, i.e. in solution depending on x and t only through $z = (x - vt)$, i.e. satisfying

$$R(x, t) = R(x - vt), \quad \Phi(x, t) = \Phi(x - vt). \quad (2.7)$$

With this *travelling wave ansatz*, (2.5) reduce to

$$\begin{aligned} \mu f'' &= \mu f g'^2 - (1/2)V'(f) \\ \mu f^2 g'' &= -2\mu f f' g', \end{aligned} \quad (2.8)$$

where we have defined the constant, quite relevant in the following,

$$\mu = mv^2 + \gamma. \quad (2.9)$$

Note that as for $L > h$ we have $\gamma > 0$, in this case we always have $\mu > 0$.

The (2.8) are the ordinary differential equations describing travelling wave solutions for (the continuum approximation of) the BCP model. Equivalently, these can be also seen as the Euler-Lagrange equations for the *travelling wave Lagrangian*

$$\mathcal{L}_{tw} = (1/2)\mu [(f')^2 - (fg')^2] - V(f). \quad (2.10)$$

(We would also arrive at the same equations also making the travelling wave reduction directly at the level of the continuum Lagrangian mentioned above. That is, the steps required to pass from the original Lagrangian to the (2.8) – Euler-Lagrange map, continuum limit, and travelling wave ansatz – can be performed in any order [21].)

It should be noted that the second of (2.8) is also rewritten as

$$\mu \frac{d(f^2 g')}{dz} = 0, \quad (2.11)$$

stating thus that the quantity in square bracket is conserved along the flow, i.e. we have the **integral of motion**

$$C := f^2 g' . \quad (2.12)$$

This also follows from the fact that \mathcal{L}_{tw} does not depend on g , and $J := \mu C$ is just the momentum conjugate to g , $J = (\partial \mathcal{L}_{\text{tw}} / \partial g')$. Recall, in this respect, that g is nothing else than the rotation angle φ (under the travelling wave ansatz); and, as remarked in section 1, the Lagrangian is invariant under a global rotation⁴. Thus this conservation law stems from a global symmetry of the Lagrangian, and could also be obtained by means of the Noether theorem [1, 16].

Finally, we note that due to (2.6) we are specially interested in solutions satisfying $f'(z) \rightarrow 0$, $g'(z) \rightarrow \chi_{\pm}$ for $z \rightarrow \pm\infty$.

3 Travelling wave solutions in the BCP model

We will now focus on the task of determining the travelling wave solutions to the BCP model; these are solutions to the Euler-Lagrange equations generated by the travelling wave Lagrangian \mathcal{L}_{tw} .

3.1 Ignorable coordinate and symmetry reduction

Making use of the conservation law (2.12), we can write

$$g' = C/f^2 , \quad \text{with } C := J/\mu = J/(mv^2 + \gamma) , \quad (3.1)$$

and insert this directly in the first of (2.8). Note that the equation for f will now not depend any more on g and its derivatives, i.e. we have obtained an autonomous equation for f (via a reduction “à la Routh” [16]); once this is solved, and f is hence a known function of time (depending parametrically on the value of the constant C , determined by initial conditions), this can be used to obtain, via (3.1), the explicit expression for $g(t)$.

Inserting (3.1) in the first of (2.8), we get

$$\mu f'' = \mu(C^2/f^3) - (1/2)V'(f) . \quad (3.2)$$

Assuming from now on $\mu \neq 0$ (if $\mu = 0$ we are left with stationary solutions corresponding to minima of the Morse potential, i.e. equilibrium configurations), we have

$$f'' = -\frac{V'(f)}{2\mu} + \left(\frac{C^2}{f^3}\right) . \quad (3.3)$$

This is the equation for a point particle of unit mass in the *effective potential*

$$W(f) := \frac{1}{2\mu} \left[V(f) + \frac{\mu C^2}{f^2} \right] . \quad (3.4)$$

The total energy for the motion in this effective potential is of course given by

$$H(f, f') = \frac{1}{2}(f')^2 + W(f) . \quad (3.5)$$

⁴That is, looking back at the original Lagrangian (1.5), we see that the invariance under a global angle shift is built in the BCP model and is not a consequence of the travelling waves reduction.

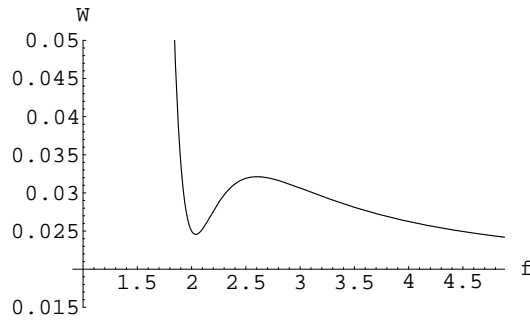


Figure 1. Plot of W , as given by (3.4), for $\lambda = \mu C^2 = 0.1$. With these values, the local maximum is reached for $f = f_0 \approx 2.60277$ and its value is $W(f_0) = W_0 \approx 0.0321189$.

3.2 The effective potential

The effective potential W defined in (3.4) depends – beside the physical parameters appearing in the Morse potential, which we think as given – on the parameters C and μ ; note C is arbitrary and is identified by initial conditions, while μ depends on the speed of the travelling wave. The values of these parameters affects not only the quantitative features of W , but its qualitative behavior as well, as we discuss in the following.

It is apparent from (3.4), the properties of the Morse potential, and $\mu > 0$, that

$$\lim_{f \rightarrow 0} W(f) = \infty ; \quad \lim_{f \rightarrow \infty} W(f) = D/\mu . \quad (3.6)$$

The potential can have zero, one or two critical points depending on the values of the parameters. We are specially interested in the case where the potential admits nontrivial solutions $f(z)$ which go to stationary points for $z \rightarrow \pm\infty$ (these correspond to solutions for which $R(x, t)$ is asymptotically constant for large $|x|$).

In view of the properties of W , such large amplitude travelling waves can exist only if W has a maximum, to which the required solution $f(z)$ is doubly asymptotic for $z \rightarrow \pm\infty$; this solution represents a separatrix for the phase portrait of (3.5).

When this condition on W is satisfied, the plot of $W(f)$ is as sketched in Figure 1 (note this has two critical points); correspondingly the level sets for $H(f, f')$ are as sketched in Figure 2 (note the presence of a separatrix).

We are thus led to investigate conditions under which W has a local maximum (recall we always have $\mu > 0$); when these are satisfied, we should moreover investigate the properties of the resulting solution doubly asymptotic to this local maximum.

3.3 Existence of travelling wave solutions

In the following, for ease of discussion, we will assume $v \geq 0$ (waves with negative velocity would be exactly symmetric to those with positive velocity). It is clear from (3.4) and $\mu > 0$ that we can as well consider the rescaled potential

$$\widehat{W}(f) := V(f) + (\lambda/f^2) \quad (3.7)$$

with $\lambda = \mu C^2 > 0$.

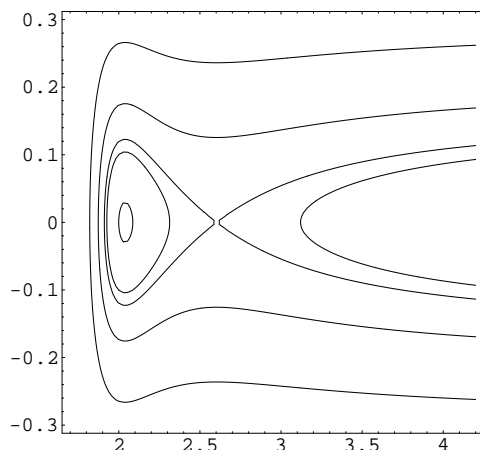


Figure 2. Contour plot for the $H(f, f')$ as given by (3.5), with $\lambda = 0.1$. We have plotted level manifolds for $H = 0.025$ and $H = 0.03$, corresponding to closed curves; for $H = W_0 \simeq 0.03212$ (the maximum of the effective potential); and for $H = 0.04$ and $H = 0.06$, corresponding to open curves.

The condition to have two critical points – a minimum and a local maximum – is that dW/df has two distinct zeroes. We write

$$w_1 = \frac{d(V(f))}{df} ; \quad \frac{dW}{df} = w_1 + w_2 , \quad w_2 = \frac{d(\lambda/f^2)}{df} . \quad (3.8)$$

With this notation we have the requirement that the graphs of w_1 and w_2 have two intersections; see Figure 3.

The limit case where the graphs of w_1 and w_2 are tangent at $f = f_0$ with $\lambda = \lambda_0$, identified by $w_1 = w_2$ and $w'_1 = w'_2$, cannot be solved for analytically, as these are transcendent equations; we can however solve these numerically with any required precision, obtaining that $f_0 \simeq 2.2006$, and $\lambda_0 \simeq 0.229354$.

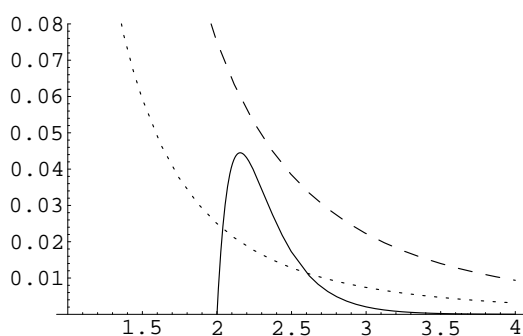


Figure 3. Plot of w_1 and w_2 as defined by (3.8); we show w_1 (continuous line) together with w_2 for $\lambda = 0.1$ (dotted) and $\lambda = 0.3$ (dashed).

We conclude that a solution with the required properties exists if and only if

$$0 < \lambda < \lambda_0 \simeq 0.229354 . \quad (3.9)$$

We also recall that $\lambda = \mu C^2$; hence this also reads $\mu < \lambda_0/C^2$. It follows from the definition of μ that there is a limit speed v_* (depending on the value of the constant of motion C) for travelling wave solutions, given explicitly by

$$v_* = \sqrt{\frac{\lambda_0/C^2 - \gamma}{m}} ; \quad (3.10)$$

travelling wave solutions exist only for $|v| < v_*$. It is worth stressing that γ is given, hence $v_* > 0$ sets a limit on C : we should have $|C| < \sqrt{\lambda_0/\gamma}$; we would arrive at the same conclusion recalling that $\mu > \gamma$ and that one should have $\mu > 0$.

We have thus identified conditions for travelling wave solutions with the required properties to exist; we will from now on assume they are satisfied, and investigate the corresponding solution.

3.4 Travelling wave solutions for the BCP model

We can finally set out to determine the travelling wave solutions to the BCP model. As discussed above, these will have an energy $H(f, f')$ exactly equal to W_0 , the maximum of the effective potential W . Recalling (3.5), we have by conservation of energy $f' = \pm\sqrt{2(E - W(f))}$; for the level set we are interested in, $E = W_0$ and we get

$$\frac{df}{dt} = \pm\sqrt{2[W_0 - W(f)]} . \quad (3.11)$$

We are thus reduced to a first order ODE; unfortunately this is – due to the exponential appearing in the Morse potential – a transcendental equation and cannot be solved analytically. On the other hand, the existence of solutions is guaranteed by general theorems; actually such solutions will depend on an arbitrary constant which can be put in relation with the value of $f(0)$, or equivalently to the center of the travelling wave, and will be unique up to these translations.

The equation (3.11) can be solved numerically; it is convenient to do this requiring that at $z = 0$ we get the minimum of $f(z)$, i.e. (calling f_0 the points for which $W(f)$ has the local maximum) the point particle reaches the point $f_1 \neq f_0$ at which $W(f_1) = W(f_0) = W_0$.

This numerical solution is shown in Figure 4(a), where we have used the values $C = 1$, $\mu = 0.1$ for the arbitrary constants. With these, the local maximum of $W(f)$ is reached at $f = f_0 \approx 2.60277$ and its value is $W_0 \approx 0.160594$. The same value is taken by W at $f_1 \approx 1.91092$, and we chose initial conditions so that for $z = 0$ we have $f(0) = f_1$.

The solution for $f(z)$ (i.e. for $R(x - vt)$) also allows to obtain immediately, via the conservation law (2.12), the solution for the function $g(z)$ (i.e. for $\Phi(x - vt)$). This is shown in Figure 4(b).

Detail of the solution around $z = 0$ (i.e near the inversion point f_1 for $f(z)$) is shown in Figure 5.

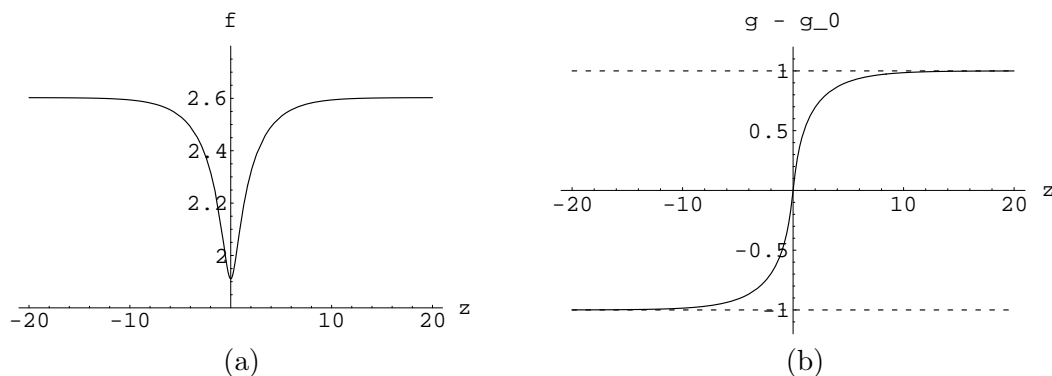


Figure 4. Numerical solution for the travelling wave reduction of the BCP model, for $C = 1$ and $\mu = 0.1$. (a): The plot for the radius $f(z)$ (i.e. $R(x - vt)$), measured in Angstroms, for $|z| < 20$. (b): The angle shift $g(z) - g_0(z)$ in units of π (the $g_0(z)$ corresponds to a regular helix).

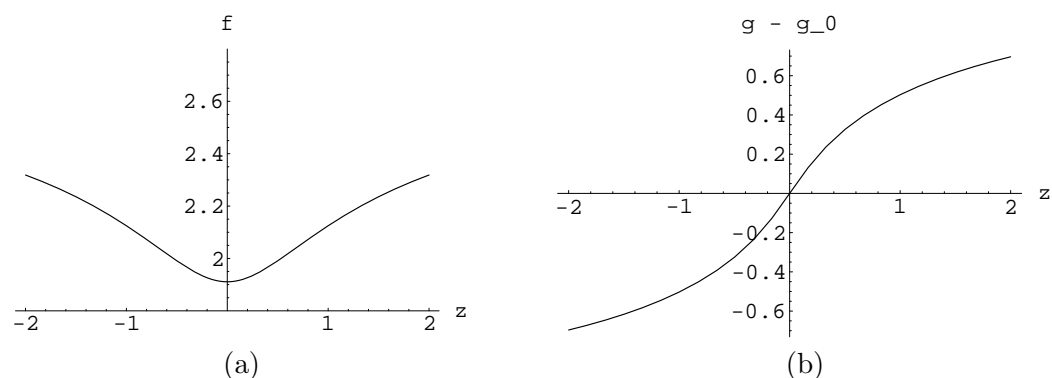


Figure 5. Detail of the numerical solution for the travelling wave reduction of the BCP model, for $C = 1$ and $\mu = 0.1$, in the region near $z = 0$. Units and notation as in Figure 4 above. (a): The plot for $f(z)$. (b): The plot for the angle shift $g(z) - g_0(z)$.

4 The simplified Cocco-Monasson model and its travelling wave reduction

We will now pass to consider the Cocco-Monasson (CM) model. This is rather similar to the BCP one, but while in BCP the distance h between planes corresponding to successive nucleotides is fixed and the length L of the phosphodiester chain unit can be stretched, in the CM the opposite applies: that is, L is fixed and h can vary.

Moreover, in this model the stacking interaction is modelled with an additional term (including an exponential prefactor) which allows to better reproduce experimental results. In our discussion we will just discuss the *simplified model* resulting from disregarding the later feature, which introduces substantial analytical difficulties; on the other hand, this simplified model implements a different geometry than the BCP model and is in this sense the exact analogue of the BCP model once one takes into account that the phosphodiester chain unit length is fixed, not the distance between base planes.

It should be stressed that in terms of the \mathbf{R}^3 ambient space, this means that the

nucleotides can also move along the double helix axis, and hence there is a corresponding contribution to kinetic energy; however, in the CM model one overlooks this contribution [10, 11], and we will keep to this approximation, which simplifies considerably several equations and computations.

The CM model is defined by the Lagrangian

$$\begin{aligned}
L = & \sum_n (m \dot{r}_n^2 + m r_n^2 \dot{\varphi}_n^2) - \sum_n D(e^{-\alpha(r_n - R_0)} - 1)^2 + \\
& - \sum_n Q e^{-b(r_{n+1} + r_n - 2R_0)} (r_{n+1} - r_n)^2 + \\
& - \sum_n K(\sqrt{L^2 - r_{n+1}^2 - r_n^2 + 2r_{n+1}r_n \cos(\varphi_{n+1} - \varphi_n)} - h)^2 + \\
& - \sum_n \nu(\varphi_{n+1} - \varphi_n) .
\end{aligned} \tag{4.1}$$

As for the term with ν , it can and will be omitted in the continuum approximation (its origin and physical meaning are analogous to those of the G term in the BCP model). As said before, we will just consider the case with $Q = 0$, but we will await a bit longer to make this simplification in order to show that some features resulting from our discussion do not depend on such an assumption (the Q terms will be put within square brackets in following computations so that the ‘‘simplified’’ version of these is easily readable).

The CM Lagrangian is rather complex, and in view of our goals it is convenient to pass directly to the continuum approximation: in other words, we use (2.1). In this way, (4.1) produces at order δ^2 a Lagrangian density (to be integrated in x ; we omit a global δ^{-1} factor due to normalization and an additive constant term $-K(L - h)^2$)

$$\mathcal{L} = (mR_t^2 + mR^2\Phi_t^2) - V(R) - \gamma(R_x^2 + R^2\Phi_x^2) - \left[q e^{-2b(R-R_0)} R_x^2 \right] . \tag{4.2}$$

Here we have written $q := Q\delta^2$, $\gamma := K\delta^2(h - L)/L$, and written simply $V(R) = D(e^{-\alpha(R-R_0)} - 1)^2$ for the Morse potential (1.2), for ease of notation. Note that $L > h$ implies $\gamma < 0$, while for $L < h$ we get $\gamma > 0$.

The (Euler-Lagrange) equations of motion corresponding to the Lagrangian (4.2) are

$$\begin{aligned}
m R_{tt} = & m R \Phi_t^2 - (1/2)V'(R) + \gamma (R_{xx} - R\Phi_x^2) + \\
& + [q e^{-2b(R-R_0)} (R_{xx} - b R_x^2)] , \\
m R^2 \Phi_{tt} = & -2m R R_t \Phi_t + \gamma R (R\Phi_{xx} + 2R_x\Phi_x) .
\end{aligned} \tag{4.3}$$

Inserting now the travelling wave *ansatz* (2.7) into these, we finally obtain⁵ the Euler-Lagrange equations for travelling wave solutions:

$$\begin{aligned}
(mv^2 - \gamma - e^{-2b(f-R_0)}) f'' = & -(1/2)V'(f) + (mv^2 - \gamma) f (g')^2 + \\
& - [b e^{-2b(f-R_0)} q (f')^2] , \\
(mv^2 - \gamma) (2f f' g' + f^2 g'') = & 0 .
\end{aligned} \tag{4.4}$$

⁵As for the BCP model, one could arrive to the same result performing the key operation (continuum interpolation, applying the Euler-Lagrange operator and inserting the travelling wave *ansatz*) in a different order [21]. In particular, we could write down a travelling wave Lagrangian \mathcal{L}_{tw} in this case as well.

5 Travelling wave solutions for the CM model

Having written down the equation for travelling wave solutions, we can now pass to solve them; we will make use – as in the BCP case – of a constant of motion they admit.

5.1 Ignorable coordinate and constant of motion

The CM model has, like the BCP model, an exact constant of motion; this is apparent from the invariance of the Lagrangian (4.1) under a global φ shift (correspondingly, \mathcal{L}_{tw} is independent of g and depends on its derivatives). Hence g is an ignorable coordinate, and the conjugate momentum $p_g = (\partial\mathcal{L}/\partial g')$ will be a first integral.

The explicit expression of this is obtained at once from the second of (4.4); in fact (for $mv^2 \neq \gamma$) this reads $(f^2 g')' = 0$. Thus, we have shown that *the quantity $J := f^2 g'$ is a constant of motion*. We stress that we reached this conclusion for the general CM model, without any simplifying assumption.

We can write therefore

$$g' = C/f^2 \tag{5.1}$$

on solutions to (4.4), with C a constant depending on initial conditions. Inserting (5.1) into the first of (4.4), we get

$$\left(mv^2 - \gamma - e^{-2b(f-R_0)} q \right) f'' = -\frac{1}{2} V'(f) + (mv^2 - \gamma) \frac{C^2}{f^3} - \left[b e^{-2b(f-R_0)} q (f')^2 \right]. \tag{5.2}$$

Note that this simplifies considerably – not only for what concerns the algebraic complexity but also in qualitative terms – if we pass to consider the simplified CM model, i.e. for $Q = 0$ (and hence $q = 0$); we will from now on consider only this. In this case we are left with

$$(mv^2 - \gamma) f'' = \frac{D\alpha}{2} e^{-\alpha(f-R_0)} \left(e^{-\alpha(f-R_0)} - 1 \right) + (mv^2 - \gamma) \frac{C^2}{f^3}. \tag{5.3}$$

5.2 The effective potential

Equation (5.3) can be written in the form $f'' = -\nabla[W(f)]$ provided we conveniently define the effective potential $W(f)$. Writing from now on $\mu := (mv^2 - \gamma)$, and assuming $\mu \neq 0$, we have $f'' = -\mu^{-1} V'(f) + C^2/f^3$. This is in the required form provided we define the effective potential $W(f)$ as

$$W(f) := \frac{1}{2\mu} V(f) + \frac{C^2}{2f^2} = \frac{C^2}{2} \left(\lambda V(f) + \frac{1}{f^2} \right), \tag{5.4}$$

where of course we defined $\lambda := [1/(C^2\mu)]$, assuming $C \neq 0$; in this case the factor $C^2/2$ is inessential for qualitative considerations. Needless to say, the qualitative behavior of $W(f)$ will depend on the sign of λ , and hence of μ . In particular, if we require that $W(f)$ diverges for $f \rightarrow 0^+$ and has a minimum near $f = R_0$, it is necessary that $\lambda > 0$ and hence $\mu > 0$.

Apart from the overall sign, the qualitative behavior of W will also depend on the actual value of λ ; in particular, this will have consequences on the number of critical points – and thus on the existence of a local maximum – for W .

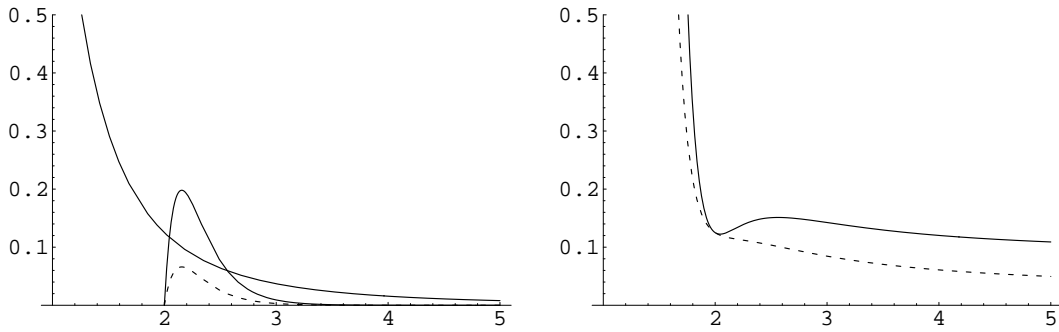


Figure 6. Left: Plot of $w_2(f) = 1/f^3$ together with $w_1(f)$ with physical values of constants and $c = 1$ for $\mu = 1$ (continuous line) and $\mu = 3$ (dotted). Right: plot of $W(f)$ with physical values of constants and $c = 1$ for $\mu = 1$ (continuous line) and $\mu = 3$ (dotted).

The required finite energy travelling wave solutions will be present, as in the BCP case, if and only if $W(f)$ has a global minimum and a local maximum.

In order to determine the conditions on the parameters of the model which ensure this is the case, we can proceed as in the BCP case. First of all we note that

$$W'(f) = C^2 [(\lambda/2) V'(f) - (1/f^3)] . \tag{5.5}$$

We thus write $W'(f) = C^2 [w_1(f) - w_2(f)]$, having defined

$$\begin{aligned} w_1(f) &:= (\lambda/2)V'(f) = -\lambda\alpha^2 D (e^{-\alpha(f-R_0)} - 1) e^{-\alpha(f-R_0)} ; \\ w_2(f) &:= 1/f^3 . \end{aligned} \tag{5.6}$$

Critical points for W correspond to solutions to $w_1 = w_2$. Note that $w_2(f) > 0$ for all f , while $w_1(f)$ changes sign at $f = R_0$; in particular, for $\lambda > 0$ and hence $\mu > 0$ (as required above), it is negative (and monotonically increasing) for $f < R_0$ and positive for $f > R_0$.

Thus if $\lambda < 0$ we have one and only extremal point for W ; for $\lambda > 0$ we could have zero or two extremal points. In this setting, a necessary and sufficient condition for the existence of two extrema (a local maximum and minimum) for W is that at the point f_* where w_1 has a maximum, the inequality

$$w_1(f_*) > w_2(f_*) \tag{5.7}$$

is satisfied. The point f_* is just $f = R_0 + \alpha^{-1} \log(2)$, as immediately seen by computing $w_1'(f)$. At this point we have

$$w_1(f_*) = \left(\frac{D\alpha^2}{4}\right) \lambda , \quad w_2(f_*) = (R_0 + \alpha^{-1} \log(2))^{-3} . \tag{5.8}$$

Thus the condition for the existence of a local maximum – and therefore of a finite energy solitary travelling wave – is simply

$$\lambda > \lambda_0 := \frac{4}{D\alpha^2 (R_0 + \alpha^{-1} \log(2))^3} . \tag{5.9}$$

With the physical values in (1.4), we get

$$\lambda_0 \approx 0.602283 . \quad (5.10)$$

These relation can also be expressed directly in terms of μ and C ; we get

$$\mu < \mu_0 := (C^2 \lambda_0)^{-1} , \quad (5.11)$$

and with the physical values $(\mu_0 C^2) \approx 1.66035$.

This is specially interesting in that μ depends on the speed of the travelling wave; in terms of the latter we get

$$\begin{aligned} v^2 < v_*^2 &:= \frac{[1/(C^2 \lambda_0)] + \gamma}{m} \\ &= \frac{1}{4mC^2L} \left[D\alpha^2L (R_0 + \alpha^{-1} \log(2))^3 + 4C^2K\delta^2(h-L) \right] . \end{aligned} \quad (5.12)$$

As stressed above, $L < h$ implies $\gamma > 0$; in this case we always have $v_*^2 > 0$, i.e. the speed v_* always exists and represents the limiting speed for solitary travelling waves. On the other hand, in the physical case $L > h$ we have $\gamma < 0$. Thus, we have a necessary and sufficient condition linking the physical parameters of the model for the existence of solitary travelling waves in the physical simplified CM model; this reads immediately from (5.12) and is

$$C^2 < \frac{D\alpha^2L^2}{4(L-h)K\delta^2} (R_0 + \alpha^{-1} \log(2))^3 . \quad (5.13)$$

5.3 Travelling wave solutions

Once again, like for the BCP model, we can use conservation of energy, and study the Hamiltonian system defined by

$$H(f, f') = \frac{1}{2} (f')^2 + W(f) \quad (5.14)$$

in order to determine travelling wave solutions (provided the conditions discussed in the previous section are satisfied); this allows to reduce our problem to a single first order ordinary differential equation, $f' = \sqrt{2[E - W(f)]}$ with E corresponding to the value attained by $W(f)$ at its local maximum. Once again, the resulting ODE is trascendent and cannot be solved exactly; it is however possible to solve it numerically.

In figure 7 we plot the numerical solution we obtained with $\mu = 0.5$ and $C = 1$ (and values of the physical parameters given by (1.4) above). In this case $E_0 \approx 0.231806$, and $f_* \approx 2.78769$; the other solution of $W(f) = E_0$ is given by $f = f_0 \approx 1.87919$. We have used this as initial condition, i.e. we have set $f(0) = f_0$ (this produces a symmetric solution; changing the initial datum with $f(0) \in [f_0, f_*]$ only shifts the symmetry center of the solution). The plot might give the impression there is a cusp at $z = 0$; this is not the case, as shown by the detail plot around $z = 0$, see Figure 8.

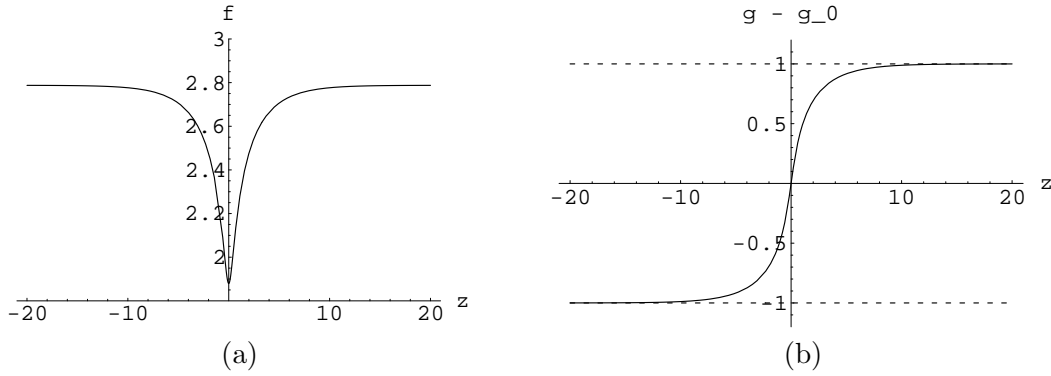


Figure 7. Travelling wave solution for the simplified CM model with $C = 1$ and $\mu = 0.5$; see text for values of the other parameters. (a): The plot for the radius $f(z)$ (i.e. $R(x - vt)$), measured in Angstroms, for $|z| < 20$. (b): The angle shift $g(z) - g_0(z)$ in units of π (the $g_0(z)$ corresponds to a regular helix).

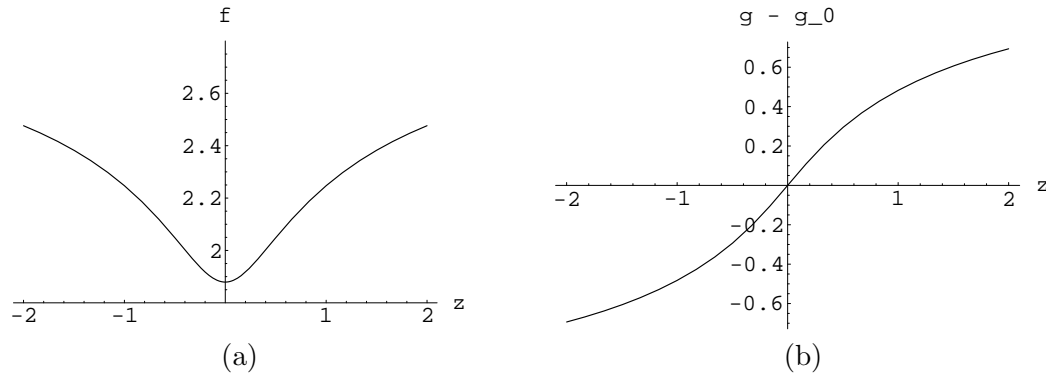


Figure 8. Detail of the travelling wave solution for the simplified CM model; parameters and units as in Figure 7. (a): Plot for $f(z)$. (b): Plot for the angle shift $g(z) - g_0(z)$.

6 Conclusions

We have discussed the existence of travelling solitary wave excitations of finite energy in the Barbi-Cocco-Peyrard (BCP) and in a simplified version of the Cocco-Monasson (sCM) models. In both cases, our analysis was conducted for general (real positive) values of the physical parameters, albeit the geometrical constraint $L > h$ is present in DNA modelling.

Our analysis was conducted in the continuum approximation, expanding up to second order in the parameter δ describing the discrete step in the nonlinear chains described by the models, see (2.1); that is, we analyzed the nonlinear wave equations (2.5) and (4.3) respectively for the two models.

We passed to the (Euler-Lagrange) equations for travelling waves via the travelling wave (TW) ansatz (2.7); we are thus reduced to a single independent variable $z = x - vt$. In both cases, the equation for the angular degree of freedom φ (denoted as g after the TW reduction; recall that f represents the radial degree of freedom after TW reduction) led to identify a constant of motion $J := f^2 g'$. The existence of this can also be immediately deduced from invariance properties of the BCP and CM Lagrangians, via Noether

theorem.

Using this, we are reduced to study a single second order equation for $f(z)$; in both cases it can be described as the Newton equation for the motion of a point particle of unit mass in an effective potential $W(f)$; this depends on a parameter μ whose sign depends on the speed v of the travelling wave.

The finite energy condition is translated, after the TW reduction, into limit conditions at $z \rightarrow \pm\infty$ for $f(z)$ and $g(z)$; these can be satisfied only if $W(f)$ admits two proper extrema, and this condition in turn puts limitation on the allowed range of μ .

For the BCP model, we observed that this conditions can be satisfied only for $L > h$, which is guaranteed by the geometry of the model. The condition is satisfied provided relation (3.9) is satisfied; this also yields a limit speed for travelling waves, see (3.10).

In the sCM model, it also happens that the conditions can be satisfied only for $L > h$ (again granted by geometry); travelling wave solutions exist provided μ satisfy condition (5.11), which again yields a limit speed for travelling waves, see (5.12).

It should be stressed that while our work on the BCP model was a complete one, for the CM model we analyzed a simplified version of the model (moreover – as traditional in analyzing this model – we overlooked the longitudinal degree of freedom, which introduces further complications in the analysis and a new component in the nonlinear wave equations to be analyzed); thus we cannot draw any rigorous conclusion for what concerns the existence of travelling waves in DNA as described by the CM model, albeit we are confident the situation will not be much different with respect to the one occurring in the simplified model.

Our results show that there are travelling waves in DNA as described by the BCP and sCM models; the speed of these waves should be less than a limiting value v_* , but within this range is a free parameter.

This should be compared with recent results [5] on a different model (cY model) with two degrees of freedom per base pair, where solitary wave excitations are present and were studied in some detail. For that model, it resulted that topologically nontrivial solitary wave excitations travel with a speed set (up to small perturbations) by a certain compatibility condition, i.e. there is a soliton speed selection. In this respect, we note that while in the BCP and sCM models the two degrees of freedom describe a radial motion and a rotation, in the cY model the two degrees of freedom describe two rotation angles (of different parts of the nucleotide). We refer to [5] for further details; see also [6, 7, 8, 9] for related matters.

Appendix. Stability of solutions

We have shown – under suitable conditions – the existence and numerically determined solutions for (the continuum version of) the general BCP and CM models in the form of travelling solitary waves. It is natural to wonder if these are stable.

Such an analysis is not possible analytically, as the solutions themselves were only determined numerically; their numerical study goes beyond the limits of the present work (and our numeric ability). We would however like to briefly sketch the scheme allowing to investigate, and possibly determine, stability of these solutions; see also [14, 20] for details.

We focus for the sake of concreteness on the BCP model. The same considerations will also apply for the CM model.

We consider in general the nonlinear wave equation

$$m \Phi_{tt} = k \Phi_{xx} + F(\Phi) \quad (6.1)$$

with $F(\Phi) = -\partial V(\Phi)/\partial \Phi$. We are dealing with a travelling wave solution to (6.1), i.e. $\Phi = \varphi(x - vt)$; and we should consider small perturbations around this, i.e. solutions of the form

$$\Phi(x, t) = \varphi(x, t) + \varepsilon \psi(x, t), \quad (6.2)$$

where $\psi(x, t)$ satisfies $\lim_{x \rightarrow \pm\infty} \psi \rightarrow 0$.

As (6.1) is Lorentz invariant, we can pass to a reference frame travelling with the same speed v as the unperturbed solution (with independent variables (z, τ) in this new frame) and in which (6.1) keeps the same form, i.e. is written as

$$m \Phi_{\tau\tau} = k \Phi_{zz} + F(\Phi) . \quad (6.3)$$

In this frame $\varphi(x - vt)$ appears as a stationary solution, i.e. is written as $\Phi = \varphi(z)$. Equation (6.2) is hence written as

$$\Phi(z, \tau) = \varphi(z) + \varepsilon \psi(z, \tau) ; \quad (6.4)$$

note $(\partial\tau/\partial t) > 0$, so stability in τ is equivalent to stability in t .

Inserting this into (6.3), we get

$$m (\varphi_{\tau\tau} + \varepsilon \psi_{\tau\tau}) = k (\varphi_{zz} + \varepsilon \psi_{zz}) + F(\varphi + \varepsilon \psi) ; \quad (6.5)$$

the last term is expanded at first order in ε as $F(\varphi) + \varepsilon F'(\varphi) \psi + O(\varepsilon^2)$. Recalling that φ is a solution to (6.3), the above simplifies to

$$m \psi_{\tau\tau} = k \psi_{zz} + F'(\varphi) \psi . \quad (6.6)$$

Moreover, as φ depends only on z by construction, we have $F'(\varphi(z)) = -V''(\varphi(z)) := f(z)$ (note this is determined by the curvature of V at its minimum).

By Fourier transforming in τ , we have

$$\psi(z, \tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} A_\omega(z) e^{i\omega\tau} d\omega ; \quad (6.7)$$

the finite energy condition requires that $A_\omega \rightarrow 0$ for $\omega \rightarrow \pm\infty$.

Thus the function A satisfies the equation (depending on $f(z)$)

$$-m \omega^2 A_\omega(z) = k A''_\omega(z) + f(z) A_\omega(z) . \quad (6.8)$$

This holds $\forall \omega$, hence we have an eigenvalue problem,

$$-k \frac{d^2 A}{dz^2} + U(z) A = m \omega^2 A , \quad (6.9)$$

with boundary conditions $A_\omega \rightarrow 0$ for $|\omega| \rightarrow \pm\infty$ (see above). Here $U(z) = -f(z) = V''(\varphi(z))$ plays the role of the potential; in this way, we are considering a stationary Schroedinger with potential $U(z)$.

The solution $\varphi(z)$ is hence (linearly) stable if and only if $\omega^2 > 0$ (i.e. ω is real); in other words, if and only if the spectrum of the Schroedinger operator $[-k d^2/dz^2 + U(z)]$ is positive in the function space of allowed perturbations ψ . See [20] for further detail on this general approach.

References

- [1] V.I. Arnold, *Mathematical methods of classical mechanics*, Springer, Berlin 1983
- [2] M. Barbi, S. Cocco and M. Peyrard, “Helicoidal model for DNA opening”, *Phys. Lett. A* **253** (1999), 358-369
- [3] M. Barbi, S. Cocco and M. Peyrard, “Vector nonlinear KleinGordon lattices: general derivation of small amplitude envelope soliton solution”, *Phys. Lett. A* **253** (1999), 161-167
- [4] M. Barbi, S. Cocco, M. Peyrard and S. Ruffo, “A twist-opening model of DNA”, *J. Biol. Phys.* **24** (1999), 97-114
- [5] M. Cadoni, R. De Leo and G. Gaeta, “A composite model for DNA torsion dynamics”, *Phys. Rev. E* **75** (2007), 021919
- [6] 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
- [7] M. Cadoni, R. De Leo and G. Gaeta, “A symmetry breaking mechanism for selecting the speed of relativistic solitons”, *J. Phys. A* **40** (2007) 8517-8534
- [8] 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
- [9] M. Cadoni, R. De Leo, S. Demelio and G. Gaeta, “Twist solitons in complex macromolecules: from DNA to polyethylene”, preprint *arXiv:0710.4475 [q-bio]* (2007), to appear in *Int. J. Nonlin. Mech.* (special issue on Macromolecules); see also R. De Leo and S. Demelio, “Numerical analysis of solitons profiles in a composite model for DNA torsion dynamics”, preprint *arXiv:0711.1069 [q-bio]* (2007), to appear in *Int. J. Nonlin. Mech.* (special issue on Macromolecules)
- [10] S. Cocco and R. Monasson, “Statistical mechanics of torque induced denaturation of DNA”, *Phys. Rev. Lett.* **83** (1999), 5178-5181
- [11] S. Cocco and R. Monasson, “Theoretical study of collective modes in DNA at ambient temperature”, *J. Chem. Phys.* **112** (2000), 10017-10033
- [12] A. Dubrovin, S.P. Novikov and A. Fomenko, *Modern geometry*, Springer (Berlin) 1984
- [13] 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”, *PNAS USA* **77** (1980), 7222-7226
- [14] G. Gaeta, “Compact travelling waves and non-smooth potentials”, *EuroPhys. Lett.* **79** (2007), 20003
- [15] G. Gaeta, C. Reiss, M. Peyrard and Th. Dauxois, “Simple models of non-linear DNA dynamics”, *Rivista del Nuovo Cimento* **17** (1994) n.4, 1-48

-
- [16] L.D. Landau and I.M. Lifshitz, *Mechanics* (Theoretical Physics, vol.1), Pergamon Press, London 1955
- [17] M. Peyrard, “Nonlinear dynamics and statistical physics of DNA”, *Nonlinearity* **17** (2004) R1-R40
- [18] M. Peyrard and A.R. Bishop, “Statistical mechanics of a nonlinear model for DNA denaturation”, *Phys. Rev. Lett.* **62** (1989), 2755-2758
- [19] M. Peyrard, A.R. Bishop and Th. Dauxois, “Entropy-driven DNA denaturation”, *Phys. Rev. E* **47** (1993), R44-R47
- [20] M. Peyrard and Th. Dauxois, *Physique de solitons*, Editions CNRS, Paris 2004; *Physics of Solitons*, Academic Press, London 2006
- [21] L. Venier, “Onde solitarie nei modelli della catena del DNA”, *M.Sc. Thesis, Università di Milano* (2007)
- [22] L.V. Yakushevich, “Nonlinear DNA dynamics: a new model”, *Phys. Lett. A* **136** (1989), 413-417
- [23] L.V. Yakushevich, “Nonlinear DNA dynamics: hierarchy of the models”, *Physica D* **79** (1994), 77-86
- [24] L.V. Yakushevich, *Nonlinear Physics of DNA*, Wiley (Chichester) 1998; second edition 2004