
Access from the University of Nottingham repository:
http://eprints.nottingham.ac.uk/938/1/Pap.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the University of Nottingham End User licence and may be reused according to the conditions of the licence. For more details see:
http://eprints.nottingham.ac.uk/end_user_agreement.pdf

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk
Recent molecular dynamics (MD) simulations of Cubero et al. (1999) of a DNA duplex containing the ‘rogue’ base difluorotoluene (F) in place of a thymine (T) base show that breathing events can occur on the nanosecond timescale, whereas breathing events in a normal DNA duplex take place on the microsecond timescale. The main aim of this paper is to analyse a nonlinear Klein-Gordon lattice model of the DNA duplex including both nonlinear interactions between opposing bases and a defect in the interaction at one lattice site; each of which can cause localisation of energy. Solutions for a breather mode either side of the defect are derived using multiple-scales asymptotics and are pieced together across the defect to form a solution which includes the effects of the nonlinearity and the defect. We consider defects in the inter-chain interactions and in the along chain interactions. In most cases we find in-phase breather modes and/or out-of-phase breather modes, with one case displaying a shifted mode.

Keywords: Localisation, defect, lattice model, breathers

1. Introduction

In the context of the internal dynamics of DNA chains, ‘breathing’ refers to the opening of base-pairs, that is, a motion in which the hydrogen bonds between complimentary bases are temporarily broken. This allows the bases to pivot around the backbone and become exposed to the solvent. Such motion was demonstrated and analysed in hydrogen exchange studies on polynucleotides and transfer RNA by Englander et al. This analysis together with that of Mandal et al. suggested that on averages, bases spent 5% of the time in an open state, and that there was a surprisingly slow timescale of a second for such motions. Mandal et al. speculated that the reason for this was an enthalpic and possibly an entropic barrier. We shall be concerned with the low energy phenomenon whereby such motion is localised in space, only occurring to one or a few base pairs at a time.

With the aim of studying denaturation in particular, Peyrard and Bishop (1989) and (1990) have applied the discrete sine-Gordon equation to the motion of DNA chains; however, such models lack realism, since the interchain bond strength varies along DNA—AT base pairs having only two hydrogen bonds whilst CG pairs have three. Salerno addresses this limitation in his account (1991) of numerical simulations of kinks in a modified DSG system. Peyrard & Farago (2000) use the Kullback-Leibler divergence deduced from the Shannon entropy to analyse the localisation in a Klein-Gordon lattice. Their numerical simulations show that at finite temperature the lattice contains ‘hotspots’, where breathers exist and are mobile,
and ‘cold regions’ where the energy density is small. Cold regions act as barriers which confine the breathers to hot spots.

Olson et al. (1993) and (2000), Matsumoto & Olson (2002) and Chen et al. (2000) analyse models valid at a scale larger than atomistic yet with potential functions which include sequence-dependent structural information. This is crucial if realistic models of DNA chains are to be constructed and used for simulations at the mesoscopic level. However, much of this work is concerned with the flexibility and curvature of the double helix as a whole (due to twist, roll, tilt, rise, slide and shift in the terminology of Yakushevich) and less concerned with the motion of individual bases relative to their complimentary bases on the other strand, such as occur in breathing events (and in the stagger, stretch, shear, buckle, opening and propellor twist motions).

The need for mesoscopic models which treat DNA at the level of bases is vital since, even with current computing power, all-atom molecular dynamics simulations are limited to a few nanoseconds and the order of 10 base pairs, whereas there are many interesting dynamical phenomena involving thousands of base pairs and occurring on timescales of microseconds up to milliseconds. Over these larger scales solvent friction becomes relevant, thus useful models at this scale will need to include stochastic forcing terms due to the presence of a heat bath.

Breathing events caused by point defects have been observed in DNA simulations by Cubero et al. (1999), and modelled by Wattis et al. (2001) using a linear model. The aim of the current paper is to consider more general types of defect, and extend the model to include the nonlinear form of the inter-strand interactions. The derivation of breather solutions of nonlinear Klein-Gordon systems involves asymptotic expansions which lead to the nonlinear Schrodinger equation (NLS) equation: thus the work of Sukhorukov et al. (2001) is relevant to the current study. They analysed the effect of point defects in NLS equations of the form \(i\psi_t + \psi_{xx} + \psi[F(|\psi|^2) + \delta(x)G(|\psi|^2)] = 0\), in which \(\delta(x)\) is the Dirac delta (generalised) function satisfying \(\delta(x) = 0\) for all \(x \neq 0\) and \(\int \delta(x) f(x) dx = f(0)\).

From the analyses of the crystal structure of DNA sequences containing a long run of A bases, Nelson et al. (1987) found ‘bifurcated hydrogen bonds’ that is bonds which form diagonally across the major groove in addition to the normal Watson-Crick complimentary bonds. Such diagonal interactions can be incorporated into ladder-type such as those used here, and have been described in earlier work of Wattis (1998). Heinemann & Alings (1989) note that the relative orientation of bases (for example, their helical twist and propellor twist) depends on the context of the bases; it is not simply dependent on the base itself, but depends on the neighbouring bases. Timsit (1999) analyses the fidelity of DNA replication, and its dependence on the DNA sequence, deducing that ‘the structural properties of a DNA sequence ... can contribute to replication errors’ and hence ‘may be connected to both genetic functions and dysfunctions’. One example of this is the presence of mismatches in a \((CA)_n\) tract which can be seen in an examination of the major groove; however, enzymes scan the minor groove and this shows no evidence of mismatches.

In the remainder of this section we introduce the model, and review two simpler models both of which show energy localisation, but caused by entirely different mechanisms. In Sections two and three we show how to analyse a model and determine the localised modes when both mechanisms are present. Section two covers
Figure 1. Illustration of the ladder model of DNA.

(Horizontal lines denote along-chain interactions modelled by linear springs with constant $k$; vertical lines marked $\gamma$ denote the inter-chain interactions modelled by nonlinear force-displacement relationship. Displacements of the upper chain are denoted by $u_n(t)$ and of the lower chain by $v_n(t)$.)

the case of inhomogeneities in the inter-chain interactions, in Section three we analyse defects in the along-chain interactions, and the combined defect is analysed in Section four. Section five shows the non-existence of stationary localised modes in a case where there is a step-change in the interaction parameters, and the paper concludes with a discussion of the results in Section six.

(a) Model

We model each base in the DNA as a separate point mass attached to three other bases – one in each direction along the same chain, and one on the complementary chain. This last bond is modelled by nonlinear force-displacement relationship with coefficients $\gamma, \eta, \nu$; the along-chain interactions are modelled by linear springs with spring constant $k$. The Hamiltonian is then

$$
H = \sum_n \frac{1}{2}m_n \dot{u}_n^2 + \frac{1}{2}m_n \dot{v}_n^2 + \frac{1}{2}k_n^{(u)} (u_{n+1} - u_n)^2 + \frac{1}{2}k_n^{(v)} (v_{n+1} - v_n)^2 + \frac{1}{2}\gamma_n (u_n - v_n)^2 - \frac{1}{3}\eta_n (u_n - v_n)^3 - \frac{1}{4}\nu_n (u_n - v_n)^4,
$$

(1.1)

where $u_n$ denotes the transverse displacement from equilibrium of one chain and $v_n$ the other chain. We assume that there are no longitudinal displacements. From the Hamiltonian we obtain the equations of motion

$$
m_n \ddot{u}_n = k_n^{(u)} (u_{n+1} - u_n) - k_n^{(u)} (u_n - u_{n-1}) - \gamma_n (u_n - v_n) + \eta_n (u_n - v_n)^2 + \nu_n (u_n - v_n)^3,
$$

$$
m_n \ddot{v}_n = k_n^{(v)} (v_{n+1} - v_n) - k_n^{(v)} (v_n - v_{n-1}) + \gamma_n (u_n - v_n) - \eta_n (u_n - v_n)^2 - \nu_n (u_n - v_n)^3,
$$

for the motion of the atoms on each chain of the double helix. This model can be simplified by using the substitution $u_n = \frac{1}{2}(x_n + y_n)$, $v_n = \frac{1}{2}(x_n - y_n)$ or, equivalently, $x_n = u_n + v_n$, $y_n = u_n - v_n$. To fully separate the equations we require $k_n^{(u)} = k_n^{(v)}$ for all $n$. This corresponds to assuming that the strength of the interactions along each chain is independent of the bases involved. We thus assign
We use small amplitude asymptotics to derive analytic approximations to solutions of this system. Clearly the types of oscillation observed will depend on the relative size of the wave to the inhomogeneity. For small amplitudes, there are coherent modes in which, to leading-order, the bases oscillate in a simple harmonic fashion due to the first three terms in the Hamiltonian, with the anharmonic terms \( (\eta \text{ and } \nu) \) determining the spatial shape of the mode. If \( \tilde{\gamma} \ll \varepsilon \), where \( \varepsilon \) is the amplitude of the mode then such a mode is expected to persist, however if \( \varepsilon \ll \tilde{\gamma} \) then one may expect the presence of the inhomogeneity to dominate the spatial structure of any coherent oscillation. When \( \tilde{\gamma} \sim \varepsilon \) then the two effects are of equal importance and it is this regime which we shall focus on in this paper. The second, and less studied inhomogeneity is that of a transition from one set of parameter values in \( n \leq 0 \) and another set of values in \( n > 0 \).
Nonlinear breathing modes due to a defect in a DNA chain

Table 1. Parameter values following the fitting shown in Figure 2.
(Eg 2 indicates the parameters used in examples of the out-of-phase modes)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>AT</th>
<th>AF</th>
<th>Eg 2 units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>0.038</td>
<td>0.042</td>
<td>J m$^{-2}$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.0515</td>
<td>0.035</td>
<td>J m$^{-2}$</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.047</td>
<td>0.034</td>
<td>$\times 10^{10}$ J m$^{-3}$</td>
</tr>
<tr>
<td>$\nu$</td>
<td>-0.010</td>
<td>-0.0080</td>
<td>$-0.1 \times 10^{20}$ J m$^{-4}$</td>
</tr>
<tr>
<td>$\nu_0^\text{in}$</td>
<td>0.22</td>
<td>0.18</td>
<td>$\times 10^{20}$ J m$^{-4}$</td>
</tr>
<tr>
<td>$\nu_0^\text{out}$</td>
<td>0.26</td>
<td>0.21</td>
<td>$-0.28 \times 10^{20}$ J m$^{-4}$</td>
</tr>
</tbody>
</table>

Before we analyse the interaction of breathers with defects, we first summarise the behaviour of a chain with all linear interactions and a single point defect, which can cause energy localisation. We then also summarise the methodology for finding breathers in a homogeneous system with nonlinearity, that is where the parameters are independent of spatial position ($n$).

(b) Parameterisation of model

We parameterise this model in a similar way to that explained in Wattis et al. (2001). We use the molecular dynamics package AMBER to measure the energy of the interchain-potential and the along-chain potential in a series of test configurations. These configurations have the form of a single base-pair being displaced by a given amount and all other base pairs left at equilibrium, i.e. $y_n = 0$ for all $n$ except $n = 0$ where $y_0 \neq 0$ and a range of values of $y_0$ are used. This corresponds to $u_0 = \frac{1}{2}y_0$ and $v_0 = -\frac{1}{2}y_0$ with all other $u_n = 0 = v_n$ in (1.1). The energy in the along-chain interactions is then $E_{\text{along}} = \frac{1}{2}ky_0^2$ and in the inter-chain interactions is $E_{\text{inter}} = \frac{1}{2}\gamma y_0^2 - \frac{1}{4}\eta y_0^3 - \frac{1}{4}\nu y_0^4$. Fitting this to the data generated from gnuplot yields the values given in Table 1, where the conversion factor 1 Kcal M$^{-1}$Å$^{-2}$ = 0.6955 $\times 10^{-3}$ J m$^{-2}$ is used and we assume that all bases have the same mass of $307/N_A$ g, where $N_A$ is Avogadro’s number, giving $m = 0.5 \times 10^{-24}$ kg.

The linear spring constants $\gamma$ for the AT and the AF base pair are significantly different from those derived in the earlier work (Wattis et al. 2001) due to the addition of non-harmonic terms in the inter-chain interactions which improve accuracy over an increased range of displacements. Parameters for the AF basepair differ most from AT in the constant $\gamma$, where the ratio is 0.68, the ratios being 0.74 and 0.8 for the cubic and quartic terms respectively.

(c) Localisation in a linear model with impurity

Here we summarise the results of Wattis et al. (2001); in which a model of the form (1.4)–(1.5) is solved with $\eta = \nu = 0 = k$; the defect in the along-chain interactions ($k$) being considerably smaller than that in $\gamma$. Being linear and sufficiently simple, the model is fully solvable, that is, all the normal modes of the system can be found; both their frequencies and their spatial structure were determined. The majority are delocalised modes, in which displacements are spread around the lattice, but there are also a couple of modes which show increased displacements at the defect site. The calculations were carried out for a finite lattice with periodic
boundary conditions, and results illustrated for a system of 12 base pairs. Statistical information about the behaviour of the system was determined by inserting $\frac{1}{2}k_BT$ of energy into each mode and observing how the displacements of each base pair varied over time. Although a qualitative observation, it is clear that the defect site displayed greater displacements than base pairs away from the defect; the RMS deviation also shows this.

Measurements of displacement were taken at 1 picosecond intervals for 10 nanoseconds, and histograms of frequency against size of displacement were calculated. These histograms displayed a clear Gaussian shape. If a breathing event was defined as an opening of 11.5Å, it was found that the AF base pair breathed four times in the 10 ns simulation, as was observed in the simulations of Cubero et al. (1999), also from the Gaussian distribution it can be extrapolated that an AT breathing event would happen approximately four times a microsecond. A second test of the data is to calculate the free energy of opening based on the ratio of time spent open to close, this gave a value of 4.2 Kcal/mol. If a breathing event is defined as an opening distance of 4.8Å then the energy of opening is reduced to 1 Kcal/mol. Such results are surprisingly good given the approximation of complex inter-base interactions by harmonic potential energy functions, and that the results are based on the extrapolation of distributions from rare breathing events. In the current paper nonlinearity prevents performing such a detailed normal mode analysis.

(d) Localisation in a homogeneous nonlinear model

Here we summarise the method of multiple scales asymptotics as originally outlined for the current problem by Remoissenet (1986). The method of multiple scales is a method of solving problems in which the quantities of interest vary over widely differing timescales. It draws on ideas from both the WKBJ method and the boundary layer theory for constructing uniformly valid approximations. In its simplest form, it is useful for solving ordinary differential equations which are close to resonance; in such systems a quantity oscillates on one timescale, while its amplitude of oscillation slowly varies over a much longer timescale. Multiple scales
Nonlinear breathing modes due to a defect in a DNA chain

techniques can find how the amplitude changes over the slow timescale, due to the combined effects of resonance and higher-order nonlinear terms. For an introduction to these techniques, see Bender & Orzsag (1978).

In our system of equations we require three timescales, to fully resolve the kinetics of the breathing motion, and to determine the shape of the wave, we need to use multiple-scales in space as well as time. We assume the solution of (1.4) takes the form

$$y_n(t) = e^{i\omega t + i\eta n} F(x, \tau, T) + e^{i2\omega t + i\eta 2n} G(x, \tau, T) + e^{i2 H(x, \tau, T) + c.c.}$$  \hspace{0.5cm} (1.7)

where $\tau = \epsilon t$, $x = \epsilon n$, $T = \epsilon^2 t$, and c.c. stands for complex conjugate. We thus have two space scales ($n$ and $x$) and three timescales ($t$, $\tau$ and $T$). Substituting (1.7) into (1.4) and equating terms of the same order in $\epsilon$ and the same leading order frequency ($e^{i\omega t}$) leads to

$$O(\epsilon e^{i\omega t} F) : \hspace{0.5cm} \omega^2 = 2\gamma + 2k(1 - \cos p)$$

$$O(\epsilon^2 e^{i\omega t}) : \hspace{0.5cm} 0 = -2\gamma H + 2\eta|F|^2$$

$$O(\epsilon^2 e^{2i\omega t}) : \hspace{0.5cm} 2i\omega F_\tau = 2ikF_x \sin p$$

$$O(\epsilon^3 e^{3i\omega t}) : \hspace{0.5cm} -4\omega^2 G = -2\gamma G + 2\eta F^2 - 2kG(1 - \cos 2p)$$

$$O(\epsilon^3 e^{4i\omega t}) : \hspace{0.5cm} 2i\omega F_T + F_{\tau\tau} = 6\nu|F|^2 F + 4\eta(FH + FH^* + F^*G) + kF_{xx} \cos p.$$  \hspace{0.5cm} (1.8)

We thus have

$$\omega^2 = 2\gamma + 4k \sin^2 \left(\frac{1}{2}p\right), \hspace{0.5cm} G = \frac{-\eta F^2}{3\gamma + 8k \sin^2 \left(\frac{1}{2}p\right)}, \hspace{0.5cm} H = \frac{\eta|F|^2}{\gamma},$$  \hspace{0.5cm} (1.9)

from the first, second and fourth equations above, and two equations for $F(x, \tau, T)$. The $O(\epsilon^3 e^{3i\omega t})$ equation implies

$$F(x, \tau, T) = F(x - u\tau, T) = F(z, T), \hspace{0.5cm} \text{where} \hspace{0.5cm} u = \frac{-k \sin p}{\omega}. \hspace{0.5cm} (1.10)$$

Thus to obtain modes which are stationary and centred near $n = 0$ ($x = 0$) we require $u = 0$, which implies $z \equiv x$, and either $p = 0$ or $p = \pi$; the former we refer to as 'in-phase breathers' and the latter, 'out-of-phase breathers'.

For general $p$ the velocity $u$ is nonzero and the $O(\epsilon^3 e^{4i\omega t})$ equation implies

$$2i\omega F_T + 2\nu F_{zz} + v_{eff}(p)|F|^2 F,$$  \hspace{0.5cm} (1.11)

which has the form of a nonlinear Schrödinger equation (NLS), in which

$$v_{eff}(p) = 6\nu + \frac{8\eta^2}{\gamma} - \frac{4\eta^2}{3\gamma + 8k \sin^2 \left(\frac{1}{2}p\right)}. \hspace{0.5cm} (1.12)$$

The moving soliton solution of (1.11) is $F = Ae^{-\alpha T + iRz} \text{sech}(\beta(z - cT))$ where

$$\beta = A \sqrt{\frac{\nu_{eff}(p)}{2(k \cos p - u^2)}}, \hspace{0.5cm} R = \frac{-\omega c}{k \cos p - u^2}, \hspace{0.5cm} \Omega = \frac{A^2 \nu_{eff}(p)(k \cos p - u^2)^2 - 2\omega^2 c^2}{4\omega (k \cos p - u^2)}.$$  \hspace{0.5cm} (1.13)
In order for $\beta$ to be real, there are restrictions on $p$. In the next sections we seek the forms of stationary localised modes in which energy is localised near the defect. For static ($u = c = 0$) in-phase ($p = 0$) solutions, we find $\Omega = A^2 \nu_{\text{eff}}^0/4\omega$, $\beta = A\sqrt{\nu_{\text{eff}}^0/2\gamma}$, $u = R = 0$, $\omega = \sqrt{2\gamma}$, and $\nu_{\text{eff}}^n = 6\nu + 20\eta^2/3\gamma$. Thus

$$y_n \sim 2\varepsilon \text{sech} \left( \sqrt{2\gamma} \nu_{\text{eff}}^0 \right) \cos \left( \sqrt{2\gamma} \left( 1 - \varepsilon^2 A^2 \nu_{\text{eff}}^0 \right) t \right) + \frac{2\varepsilon^2 A^2 \nu_{\text{eff}}^0}{\gamma} \left( \sqrt{2\gamma} \nu_{\text{eff}}^0 \right) \left[ 1 - \frac{1}{3} \cos \left( \sqrt{2\gamma} \left( 1 - \frac{\varepsilon^2 A^2 \nu_{\text{eff}}^0}{8\gamma} \right) t \right) \right].$$  \hspace{1cm} (1.14)

This solution corresponds to the low frequency breather since it lies just below the dispersion relation for linear waves. Note that this solution exists only for $\nu > -10\eta^2/9\gamma$. For static ($u = c = 0$) out-of-phase ($p = \pi$) solutions, we find

$$F = Ae^{-i\xi A^2 \nu_{\text{eff}}^0 \sigma/4\omega} \text{sech} \left( Ax \sqrt{\frac{-\nu_{\text{eff}}^0}{2k}} \right),$$  \hspace{1cm} (1.15)

where $u = R = 0$, $\omega = \sqrt{2\gamma + 4k}$ and $\nu_{\text{eff}}^0 = 6\nu + 8\eta^2/\gamma - 4\eta^2/(3\gamma + 8k)$, and thus

$$y_n \sim 2\varepsilon A \left( -1 \right)^n \text{sech} \left( \sqrt{2\gamma + 4k} \nu_{\text{eff}}^0 \right) \cos \left( \sqrt{2\gamma + 4k} \left( 1 - \varepsilon^2 A^2 \nu_{\text{eff}}^0 \right) t \right) + \frac{2\varepsilon^2 A^2 \nu_{\text{eff}}^0}{\gamma} \left( \sqrt{2\gamma + 4k} \nu_{\text{eff}}^0 \right) \left[ 1 - \frac{1}{3} \cos \left( \sqrt{2\gamma + 4k} \left( 1 - \frac{\varepsilon^2 A^2 \nu_{\text{eff}}^0}{8\gamma + 16k} \right) t \right) \right],$$  \hspace{1cm} (1.16)

for the discrete Klein-Gordon breather mode. This is a high-frequency breather since it lies just above the dispersion relation for small amplitude linear waves; it exists only for $\nu > -2\eta^2(5\gamma + 16k)/(3\gamma + 8k)$.

2. A point defect in the inter-strand interaction

We now construct the solution for a breather in a nonlinear system in which there is also a defect. We shall assume initially that there is a point defect in $\gamma$ at $n = 0$, and that all other parameters are uniform across the whole lattice. We define a solution $y_n^+ (t)$ which is defined for all $n$ and will be the solution to the equation of motion in $n \geq 1$; similarly $y_n^- (t)$ is defined for all $n$ and is the solution in $n < 0$. Each satisfies the homogeneous equation (1.4) but not the equation at $n = 0$. The solutions $y_n^\pm (t)$ thus have the form (1.14) or (1.16) with amplitudes $A_\pm$ and displacements $n_0 = n_{\pm}$.

In order for the frequencies of $y_n^\pm$ to be the same we require $A_\pm^2 = A_\mp^2$. Applying the equation (1.4) at $n = 1$ means that $y_1 = y_0^+$ and applying (1.4) at $n = -1$ implies $y_{-1} = y_0^-$; thus we have $y_0^+ = y_0^-$ and so $A_-$ and $A_+$ must have the same sign and so are equal. We define $A = A_+ = A_-$, and the remainder of this section is involved with determining $n_{\pm}$.

We know that $y_n^\pm$ satisfies

$$\ddot{y}_n^\pm = k(y_{n+1}^\pm - 2y_n^\pm + y_{n-1}^\pm) - 2\gamma y_n^\pm + 2\eta(y_n^\pm)^2 + 2\nu(y_n^\pm)^2,$$  \hspace{1cm} (2.1)
for all \( n \), and we wish to find \( n_\pm \) such that at \( n = 0 \) we have
\[
y_0^\pm = k(y_1^+ - 2y_0^+ + y_-^+) - 2\gamma y_0^+ + 2\eta(y_0^+)^2 + 2\nu(y_0^+)^3 - 2\gamma y_0^+ + 2\eta(y_0^+)^2 + 2\nu(y_0^+)^3 , \tag{2.2}
\]

taking the differences of these two equations yields the condition
\[
0 = ky_{-1} - ky_1 - 2\gamma y_0^+ + 2\eta(y_0^+)^2 + 2\nu(y_0^+)^3 . \tag{2.3}
\]

(a) In-phase breathers at a defect in \( \gamma \)

For in-phase breathers the solution can be written as
\[
y_n^\pm (t) = 2\varepsilon A \pm \text{sech}(\theta A \pm (n-n_\pm)) \cos(W \pm t) +
\frac{2\varepsilon^2 A^2}{\gamma} \text{sech}^2(\theta A \pm (n-n_\pm))(1 - \frac{1}{4} \cos(2W \pm t)) + \ldots \tag{2.4}
\]
in the regions \( n > 0 \) and \( n < 0 \), where
\[
W = \sqrt{2\gamma} \left(1 - \frac{\varepsilon^2 A^2 \nu_{\text{eff}}^2}{8 \gamma}\right) , \quad \theta = \varepsilon \sqrt{\frac{\nu_{\text{eff}}^2}{2k}} . \tag{2.5}
\]

By considering the equation of motion at \( n = \pm 1 \) we obtain the condition \( y_0^+ = y_0 = y_0^- \) and hence \( \text{sech}(\pm \theta n_\pm) = \text{sech}(\pm \theta n_-) \); there are two cases to consider: \( n_+ = n_- \) and \( n_+ = -n_- \). The former yields \( y_0^+ = y_0^- \) and so there is no modification to the wave form by the presence of the defect, this leads to the contradiction \( \gamma = 0 \). Thus we have \( n_+ = n_0, n_- = -n_0 \), and using (2.3) we find \( \gamma = k\theta A \tanh(\theta A n_0) \) which simplifies to
\[
n_0 = \frac{1}{\varepsilon A} \sqrt{\frac{2k}{\nu_{\text{eff}}^2}} \tanh^{-1} \left( \frac{\gamma}{\varepsilon A} \sqrt{\frac{2}{kv_{\text{eff}}^2}} \right) . \tag{2.6}
\]

With this definition equation (2.4) can be rewritten, the leading order term becomes
\[
y_n = 2\varepsilon A \text{sech}(\theta A(n_0 + |n|)) \cos(W t) + \mathcal{O}(\varepsilon^2) .
\]

For small amplitude defects, where \( \gamma \ll \varepsilon \), the quantity \( n_0 \) is linearly related to the amplitude of the defect and \( \varepsilon n_{0} \sim \gamma/\varepsilon \ll 1 \). However, since the oscillation has a width of \( \mathcal{O}(\varepsilon^{-1}) \), only when the width is altered by an amount of this size that the effect of the defect is noticed at leading order, that is \( n_0 = \mathcal{O}(\varepsilon^{-1}) \) is the interesting size of defect. This corresponds to a defect of strength \( \gamma = \mathcal{O}(\varepsilon) \). For breather-defect combinations satisfying \( \gamma/\varepsilon \ll 1 \) the defect will be almost invisible to the breather, whereas for \( \gamma/\varepsilon \gg 1 \) the defect will radically alter the shape of the breather, and there is a smooth transition between these two limits which occurs when \( \gamma/\varepsilon = \mathcal{O}(1) \).

There is an abrupt change in \( y_{n+1} - y_n \) at the defect, observable as a ‘sharp corner’ in the profile of the wave as illustrated in Figure 3(a); parameter values are given in Table 1. Negative values of \( \gamma \) lead to single-humped modes with a sharp corner at \( n = 0 \) (corresponding to \( n_0 < 0 \)), whereas \( \gamma > 0 \) gives \( n_0 > 0 \) yielding double-humped solutions with maxima both sides of \( n = 0 \), namely at \( n = \pm n_0 \). As \( \gamma \rightarrow -\varepsilon A \sqrt{kv_{\text{eff}}^2/2} \), we have \( n_0 \rightarrow -\infty \) and the shape of the solution approaches the \( e^{-\lambda |n|} \) observed in the linear chain with a defect (Wattis et al. 2001).
As $\tilde{\gamma} \to \epsilon A \sqrt{(k_B T / \theta)}$, we have $n_0 \to \infty$, and for larger values of $\tilde{\gamma}$ such solutions cease to exist.

The energy in these solutions can be calculated and the energy contained in such a mode observed would typically be $k_B T / 2$ where $k_B$ is Boltzmann’s constant ($1.38 \times 10^7$ J K$^{-1}$) and $T$ is the Temperature, which we shall take as 293 K so that $k_B T = 4.1 \times 10^{-21}$ J. The energy of the solution can be calculated by substituting the solution (2.4) into (1.1). To leading order, we obtain

$$ H = 4\gamma \epsilon^2 A \theta \left( 1 + \frac{\tilde{\gamma}}{k_B T} \right), \quad (2.7) $$

and by putting $H = \frac{1}{2}k_B T$ we gain the amplitude

$$ \epsilon A = \frac{k_B T \theta}{8\gamma} - \frac{\epsilon \tilde{\gamma}}{k_B T}. \quad (2.8) $$

The solution so determined is illustrated in Figure 3(b), although it should be noted that in this case the calculated amplitude is large enough that the condition $\epsilon A \ll 1$ is of dubious validity. A comparison with the solution of the linear problem (given by equations (3.28) and following from Wattis et al. (2001)) are shown and denoted by $y_l$, from which it is seen that the inclusion of nonlinear terms produce a dramatic reduction in the width of the breathing mode, but little difference is made to its amplitude.

(b) Out-of-phase breathers at a defect in $\gamma$

Here we define the out-of-phase breather solutions $y^+_n$ and $y^-_n$ to be valid in $n \geq 1$ and $n \leq -1$ respectively, by

$$ y^+_n(t) = 2(-1)^n \epsilon A \text{sech}(\theta A (n-n_\pm)) \cos(W \pm t) + $$

$$ + \frac{2\epsilon^2 A^2 \eta \gamma}{\gamma} \left( 1 - \frac{1}{3\gamma + 8k} \cos(2W \pm t) \right) \text{sech}^2(\theta A (n-n_\pm)) + \ldots \quad (2.9) $$

---

*(Figure 3. In-phase breather modes.

((a) Illustrations for $\tilde{\gamma} = -0.01$ J m$^{-2}$ (single-humped) and $\tilde{\gamma} = 0.04$ J m$^{-2}$ (double-humped), with parameter values $\epsilon A = 0.4$ Å. (b) Illustration of the breather mode for $\tilde{\gamma} = -0.016$ J m$^{-2}$, with amplitude determined by the energy of the mode being $\frac{1}{2}k_B T$, compared with the solution of the linear system (3.28) in Wattis et al. (2001).))

Article submitted to Royal Society)
Figure 4. Out-of-phase breather modes.
(Parameters: \( \gamma = -0.02 \text{ J m}^{-2}, A = 0.3 \text{ Å}, \) single-humped and denoted by \( y_1 \); and \( \gamma = 0.02 \text{ J m}^{-2}, A = 0.4 \text{ Å}, \) double-humped and denoted by \( y_2 \).)

where

\[
W_\pm = \sqrt{2\gamma + 4k + \left(1 - \frac{\varepsilon^2 A_2^2 \nu_{\text{out}}^2}{8\gamma + 16k}\right)} , \quad \theta = \varepsilon \sqrt{\frac{-\nu_{\text{out}}^2}{2k}}.
\]  \hspace{1cm} (2.10)

The argument proceeds as for in-phase breathers. The condition \( W_+ = W_- \) implies \( A_+^2 = A_-^2 \) as before; and by considering the equation (1.4) at \( n = \pm 1 \) we obtain the condition \( y_0 = y_0^0 = y_0 \). Hence \( A_+ \) and \( A_- \) must have the same sign and so \( A = A_+ = A_- \); it then only remains to determine \( n_\pm \). The condition \( y_0^1 = y_0^2 \) implies \( n_+ = n_- \), which ultimately leads to a contradiction, or \( n_+ = -n_- \). We define \( n_0 = n_+ = -n_- \) and determine \( n_0 \) from the condition (2.3) which implies

\[
n_0 = -\frac{1}{\theta A} \tanh^{-1} \left( \frac{\gamma}{\theta A k} \right).
\]  \hspace{1cm} (2.11)

Again the crucial size of defect is when \( \gamma \) becomes \( O(\varepsilon) \), the size of displacements \( \pm n_0 \sim O(1/\varepsilon) \) are the same order of magnitude as the width of the wave, so are noticeable at leading order. For smaller defects the displacements \( n_\pm = \pm n_0 \) scale linearly with the defect strength according to \( n_0 \sim 2\gamma/\varepsilon^2 A_2^2 \nu_{\text{out}}^2 \) but are negligible at leading order. Breathers of the form (2.9) with (2.11) are illustrated in Figure 4. The lines corresponding to \( y_1 \) are the leading order approximations (correct to \( O(\varepsilon) \)) which are symmetric about the zero displacement line; \( y_2 \) indicates the approximation including the first correction terms (correct to \( O(\varepsilon^2) \)). The trailing \( \pm \) indicates the sign of \( \gamma \). Parameter values are given in Table 1. Figure 5 illustrates the spatiotemporal form of both the in-phase and out-of-phase single-humped breather solutions.

3. A point defect in the along-chain interactions

In this case we consider the case where there is one base-pair interaction where the spring constant is given by \( k + k \) instead of \( k \). We assume that the inter-chain interactions \( (\gamma, \eta, \nu) \) are uniform along the chain. Equation (1.4) holds at all sites except \( n = 0 \) and \( n = 1 \). As in the previous section we define two solutions denoted by \( y_{1\pm}^k \), each defined for all \( n \) but only relevant in one half of the domain. We have

\[ Article submitted to Royal Society\]
$y_n = y_n^+$ at all sites $n \geq 1$ and $y_n = y_n^-$ when $n \leq 0$. If we are considering in-phase modes, $y_n^+$ are given by (1.14) with $A = A_\pm$ and $n_0 = n_\pm$; for out-of-phase modes, then $y_n^-$ are given by (1.16). At sites $n = 0, 1$ we have to satisfy

$$y_0 = k(y_1 - 2y_0 + y_{-1}) + \hat{k}(y_1 - y_0) - 2\gamma y_0 + 2\eta y_0^2 + 2\nu y_0^3,$$

(3.1)

$$y_1 = k(y_2 - 2y_1 + y_0) - \hat{k}(y_1 - y_0) - 2\gamma y_1 + 2\eta y_1^2 + 2\nu y_1^3.$$  

(3.2)

These lead to conditions on $A_\pm, n_\pm$. By subtracting (1.4) at $n = 0, 1$ from each of the above equations we obtain the conditions

$$y_1^+ - y_1^- = -\frac{\hat{k}}{k}(y_1^+ - y_0^-) = y_0^+ - y_0^-.$$  

(3.3)

Requiring the temporal frequency $W$ to be the same in $n_1$ and $n_0$ implies $A_+^2 = A_-^2$ for both in- and out-of-phase breathers.

(a) In-phase breathers at a defect in $k$

Applying the condition $y_1^+ - y_1^- = y_0^+ - y_0^-$ from (3.3) we obtain

$$\text{sech}(\theta An_+ \tan(\theta An_-)) = \text{sech}(\theta An_- \tan(\theta An_-)),$$

(3.4)

from which we deduce that $n_+$ and $n_-$ have the same sign. By analysing (3.4) we find three possible solutions

$$e^{\theta An_+} = e^{\theta An_-} + 1 \quad \text{if} \quad n_- > 0; \quad e^{\theta An_+} = \frac{1 - e^{\theta An_-}}{1 + e^{\theta An_-}} \quad \text{if} \quad n_- < 0,$$

(3.5)

or $n_+ = n_-; \quad$ however, this last option implies $y_n^+ = y_n^-$, that is the solutions in $n \geq 1$ and $n \leq 0$ are identical and leads to the contradiction $\hat{k} = 0$, thus we ignore this possibility from hereon. A second condition, $ky_0^+ + \hat{k}y_1^+ = (k + \hat{k})y_0^-$, derived from (3.3) determines $n_-$ and hence $n_+; \quad$ for $n_+ < 0$ we have the solution

$$n_- = \frac{1}{\theta A} \log(\sqrt{2} - 1) - \frac{\hat{k}}{2(k + \hat{k})}, \quad n_+ = \frac{1}{\theta A} \log(\sqrt{2} - 1) + \frac{\hat{k}}{2(k + \hat{k})}.$$  

(3.6)
Nonlinear breathing modes due to a defect in a DNA chain

Figure 6. In-phase breather modes for a defect in the along-chain interactions.

and for $n_{\pm} > 0$ we have

$$n_+ = \frac{1}{\theta A} \log(\sqrt{2} + 1) + \frac{\hat{k}}{2(k + \hat{k})}, \quad n_- = \frac{1}{\theta A} \log(\sqrt{2} + 1) - \frac{\hat{k}}{2(k + \hat{k})}. \quad (3.7)$$

We see that to leading order, the offset of the waves is the same in both $n \geq 1$ and $n \leq 0$, that is $n_{\pm} = O(1/\varepsilon)$, specifically $\pm \log(1 + \sqrt{2})/\theta A$ so that the wave crosses the inhomogeneity at its point of inflection. There are $O(\hat{k})$ corrections to the shift of the wave, and these are different in the regions $n \geq 1$ and $n \leq 0$.

Figure 6 shows the form of the breathers in the case $\hat{k} = 0.008$ (on the left) and $\hat{k} = -0.008$ (on the right). Note that in the former case, where the interaction is stronger (larger spring constant) at the defect, the wave has a slightly increased width, whereas when the interaction is weaker ($\hat{k} < 0$) the wave is narrowed. To make this effect observable, the value of $\hat{k}$ used in Figure 6 is twice that for an AF bond; the other parameters being as given in Table 1 with $\varepsilon A = 0.5$ A. lines denoted $y_1$ denote the leading order term approximation to $y_{n_0}$, and those denoted $y_2$ include the $O(\varepsilon^2)$ correction terms also. The trailing + indicates $n_+ > 0$ and − indicates the case $n_- < 0$.

(b) Out-of-phase breathers at a defect in $k$

This case is similar to the $\gamma$-defect modes rather than the in-phase $k$-defect mode. If we seek translations of the wave by amounts $n_{\pm} = n_0 + \tilde{n}_{\pm}$, where $\tilde{n}_{\pm}$ are small corrections to $n_0$ we find the two conditions $y_1^+ - y_0^+ = y_1^0 - y_0^0$ and $ky_0^0 + \hat{k}y_1^0 = (k + \hat{k})y_0^0$ lead to

$$\text{sech}(\theta A n_+)[2 + \theta A \tanh(\theta A n_+)] = \text{sech}(\theta A n_-)[2 + \theta A \tanh(\theta A n_-)]$$

$$4\hat{k} + (k - \hat{k})\theta A \tanh(\theta A n_-) = (k + 3\hat{k})\theta A \tanh(\theta A n_+). \quad (3.8)$$

The former is automatically satisfied at leading order (and yields $\tilde{n}_- = \tilde{n}_+ - 1$ at the first correction term), and the latter yields $n_0 = (-1/\theta A)\tanh^{-1}(2\hat{k}/k\theta A)$. Thus when $\hat{k} = O(\varepsilon)$ we find solutions in which $n_{\pm} = O(1/\varepsilon)$ as illustrated in Figure 7, for the parameter values in Table 1 (with $\nu = -0.1$); the left-hand graph is for $\hat{k} > 0$ and the right-hand for $\hat{k} < 0$. 

Article submitted to Royal Society
4. Defect in both inter-chain and along chain parameters

In the above analyses we have seen two types of defect mode: in most cases we observe a symmetric mode with a corner at the defect; and in the case of a defect in the along-chain potential \( (k) \), a pair of mirror image shifted modes were obtained.

We now address the question of which type of mode occurs in a system with defects—as occurs in a DNA strand with a rogue AF-base pair (see parameters in Table 1). We aim to find the defect modes associated with defects of the form

\[
\begin{align*}
    y_1 &= k(y_2 - y_1) - (k + \tilde{k})(y_1 - y_0) - 2\gamma y_1 + 2\eta y_1^2 + 2\nu y_1^3, \\
    \tilde{y}_0 &= (k + \tilde{k})(y_2 - 2y_1 + y_0) - 2(\gamma + \tilde{\gamma})y_0 + 2\eta y_0^2 + 2\nu y_0^3, \\
    \tilde{y}_{-1} &= (k + \tilde{k})(y_0 - y_{-1}) - k(y_{-1} - y_{-2}) - 2\gamma y_{-1} + 2\eta y_{-1}^2 + 2\nu y_{-1}^3,
\end{align*}
\]

(4.1)

that is with \( k_{\pm} = k + \tilde{k} \) and \( \gamma_0 = \gamma + \tilde{\gamma} \) in (1.2). By comparing each of the above equations with (2.1) we obtain the conditions which hold at the defect, namely

\[
(k + \tilde{k})y_0 = ky_0^+ + \tilde{ky}_0^+ = ky_0^- + \tilde{ky}_0^-.
\]

(4.2)

(a) In-phase breather at a defect in \( k \) and \( \gamma \)

For an in-phase solution we define our solution by \( y_n = y_n^+ \) in \( n > 0 \), \( y_n = y_n^- \) in \( n < 0 \) with \( y_n^+ \) as in (2.4) and \( y_0 \) to be determined by (4.2); not that \( y_0 \) is not either of \( y_0^+ \). Seeking a shifted solution, that is where \( n_+ = n_0 + \tilde{n}_+ \) and \( n_- = n_0 + \tilde{n}_- \) with \( \tilde{n}_+ \ll n_0 \) yields a contradiction (since the correction terms \( \tilde{n}_\pm \) also satisfy \( \tilde{n}_+ = \tilde{n}_- \), and so the there is no modification to the wave); whereas seeking a corner solution, by substituting \( n_+ = n_0 + \tilde{n}_+ \) and \( n_- = -n_0 + \tilde{n}_- \) into the equation for \( y_0 \) ultimately leads to (2.6) for the leading order term, \( n_0 \). Thus the modes in this case will have the same form as those displayed in Figure 3.

(b) Out-of-phase breather at a defect in \( k \) and \( \gamma \)

For an out-of-phase solution we define our solution by \( y_n = y_n^+ \) in \( n > 0 \), \( y_n = y_n^- \) in \( n < 0 \) with \( y_n^+ \) as in (2.9) and \( y_0 \), as above, determined by (4.2). Again we find...
This generalisation of (2.11) yields a mode similar to those in Figures 4 and 7.

5. A step change in the interaction parameters

There is one other type of parameter changes we now consider: that is where the parameters $k, \gamma, \eta, \nu$ take one set of values in $n < 0$ and another in $n \geq 0$, we refer to this scenario as a step change in the parameters. In most cases considering the dispersion relation in $n \geq 0$ gives two frequencies for which there exist stationary localised modes, the in-phase breather solution and the out-of-phase breather solution. When considering the region $n < 0$ these two types of breather then correspond to different frequencies, and so it is impossible to piece together a single mode across the step-change. An alternative way of thinking about this scenario is to consider a frequency for which there is a localised mode in $n < 0$, then for the same frequency, in $n < 0$ this would correspond either to a mode in the interior of the dispersion relation (i.e. $0 < p < \pi$ which would give rise to a moving breather mode), or be outside the dispersion altogether.

There is one exception to this general rule, that is when only the along chain interactions $(\kappa)$ take different values in $n < 0$ and $n \geq 0$; here the frequency of the in-phase breather does not depend on $k$ so will be the same in both regions. Defining the solution in $n \geq 0$ by $y_{n+}^+$ with $\theta = \theta_+ = \varepsilon \sqrt{(\nu_{\text{eff}}^n/2(k+\kappa))}$, and that in $n \leq 0$ by $y_{n-}^+$ with $\theta = \theta_- = \varepsilon \sqrt{(\nu_{\text{eff}}^n/2k)}$ in $n < 0$, we have

$$\ddot{y}_n^+ = (k + \kappa)(y_{n+1}^+ - 2y_n^+ - y_{n-1}^+) - 2\gamma y_n^+ + 2\eta (y_n^+)^2 + 2\nu (y_n^+)^3; \quad (5.1)$$

for all $n$, including $n = 0$ and at $n = 0$ we also have

$$\ddot{y}_0^+ = (k + \kappa)(y_1^+ - y_0^+) - k(y_0^+ - y_0^-) - 2\gamma y_0^+ + 2\eta (y_0^+)^2 + 2\nu (y_0^+)^3; \quad (5.2)$$

taking the difference of the latter and the former at $n = 0$ gives

$$k(y_1^+ - y_{-1}^-) = \kappa(y_0^+ - y_0^-). \quad (5.3)$$

This, and $y_0^+ = y_0^-$ determine the two unknowns: $n_{\pm}$. As with earlier cases, matching the frequencies in the two regions leads to $A_+^2 = A_-^2$ and the condition $y_0^+ = y_0^-$ implies that the signs of $A_+$ and $A_-$ must be the same, and so $A_+ = A_- = A$. Applying the condition $y_0^+ = y_0^-$ using (2.4) leads to $n_+ = \pm n_- \sqrt{(1 + \kappa/k)}$. Applying the second condition (5.3) leads to the contradictory conclusions of $\kappa = 0$ or $\kappa = -k$, i.e. no change in parameter values across $n = 0$, or there being no along-chain interactions at all on one side of $n = 0$. Thus there are no localised breather modes at a site where the interaction parameters $(\gamma, k, \eta, \nu)$ have a step change from one set of values in $n > 0$ to different values in $n \geq 0$. 

*Article submitted to Royal Society*
6. Conclusions

The main aim of this paper has been to derive the form of breather solutions in simple models of DNA which contain both nonlinear interactions and inhomogeneities. Each of these on its own can cause localisation of energy, and we have shown how the two effects interact. The inhomogeneities considered have been of the form of point defects affecting only one base pair interaction, or one along-chain bond. In these systems we derived expressions for the form of stationary breathers at or near such sites. There are typically two modes which satisfy the criteria of being stationary at the defect that is a low frequency 'in-phase' mode and a higher frequency 'out-of-phase' mode, corresponding to the top and bottom edges of the dispersion relation. These modes exist for point defects in the inter-chain interactions and for the along-chain interactions.

In the penultimate section we have also considered inhomogeneities in which one half of the lattice is governed by one set of parameter values and the other half by a different set of values. In such systems it is, in general, not possible to construct stationary breather modes.

We have found shifted modes for $k$-defects (that is, in the along chain interactions) with an in-phase mode; and corner modes for $\gamma$-defects (that is, in the inter-chain interactions) and out of phase $k$-defects both these types of mode were observed by Sukhorukov et al. (2001) in an NLS system. We expect these results to be indicative of the behaviour of systems in which the inhomogeneities in lattice parameters are ‘smeared out’ over several lattice sites, rather than the abrupt changes we have considered here. With realistic DNA parameters and $\frac{1}{2}k_BT$ energy in such a mode, we are at the limit of applicability of the small amplitude asymptotic expansion. Nevertheless we see that the combined effect of nonlinearity and defect causes a narrower oscillation than was seen in the pure linear system analysed previously (Wattis et al. 2001). While the out-of-phase breather mode gives another mechanism for high-frequency breathing in Klein-Gordon systems, the parameters required to realise it do not include those found in DNA.

Future work should be aimed at determining the stability of these solutions, we expect the double-humped solutions to be unstable and the single-humped to be stable. The form of localised modes for longer stretches of alternative parameters would also be interesting, for example as occurs in a sequence of TATA base pairs in a DNA strand which is predominantly composed of C’s and G’s.

I am grateful to Sarah Harris and Charlie Laughton for useful conversations. I also thank the referees for comments on the manuscript and pointing out additional references.

References


Nonlinear breathing modes due to a defect in a DNA chain


Article submitted to Royal Society