Progress of Theoretical Physics, Vol. 72, No. 4, October 1984

A Coupled Base-Rotator Model for Structure and Dynamics of DNA

—— *Local Fluctuations in Helical Twist Angles and Topological Solitons* ——

Shigeo HOMMA and Shozo TAKENO*

Department of Applied Physics, Faculty of Engineering
Nagoya University, Nagoya 464
**Department of Physics, Kyoto Technical University, Kyoto 606*

(Received April 23, 1984)

Structure and dynamics of DNA are studied by using a model dynamical system in which each base in the system, coupled with its complementary base and nearest neighbours in the same strand by the hydrogen-bonding and the stacking energy, respectively, is allowed to rotate in a plane perpendicular to the helical axis. The potential energy of the base system is taken to be composed of two parts, intra-strand base-base interaction energy and inter-strand one, in which base-sequence variation of interaction constants is neglected. When the intra-strand interactions are much larger than the inter-strand ones, a continuum approximation can be used, and the model system admits various topological solitons propagating along the helical axis. By studying numerically nonlinear difference equations determined from the extrema of a potential function of the model system in its simplified yet nontrivial case, it is shown that there exist fairly large local fluctuations in helical twist angles from one base pair to the next characterized by commensurate, incommensurate and chaotic phases, in addition to the conventional, idealized B-form.

§ 1. Introduction

The conventional structure of DNA is the so-called B-form proposed more than thirty years ago by Watson and Crick.[1] Despite its widespread acceptance, however, this model was based on x-ray diffraction studies of fibers with limited resolution. Further examination of x-ray patterns from crystals, fibers and thin films of natural DNA revealed two other forms, A-form and Z-form, of the double helix, the latter being so different from the other two due to left-handed structure in its twist.[2] On the other hand, it has been known for some time that polar hydrogens of bases in DNA and synthetic duplexes exchange with solvent hydrogens under conditions in which these molecules are ordered and remote from any denaturation transitions.[3] Free bases are able to exchange their N-H with solvent much more rapidly than those coupled with their complementary bases by hydrogen bonding. This has led to the proposal that ordered helices contain small amounts of open states, in which bases are unpaired, and that these open states mediate exchange of otherwise inaccessible hydrogen-bonded protons. It is therefore likely that aside from several forms of averaged, overall structure, there exist local fluctuations in the structure of DNA. Englander et al.[4] and Yomosa[5] suggested that open segments in DNA may propagate as solitons. In arriving at such a result, these workers employed a continuum approximation to model equations describing the dynamical properties of DNA.

In a previous paper, which will hereafter be referred to as (I),[6] we made a brief report of our theory based on a dynamic plane base-rotator model. Main results obtained there are: (1) There exist in the continuum approximation topological solitons having the properties somewhat different from those obtained previously.[4),5] (2) Numerical solu-

tions to original, static, spatially discrete model field equations give fairly large local variations in helical-twist angles from one base pair to the next characterized by commensurate, incommensurate and chaotic phases. The results in (I) are based on a highly idealized model of DNA, in which non-periodicity of force field along the helical axis due to a particular base-pair sequence is neglected. The qualitative features of the second of the above-mentioned results, i.e., local fluctuations in the helical twist angles may remain unchanged or even be enhanced if we consider more realistic situations by taking sequence-dependence of the field into account and by allowing several other degrees of freedom for the dynamics of bases.

It is the purpose of this paper to make a more detailed study of the problem outlined in (I). This is done by (1) improving our previous plane base-rotator model and (2) doing more extensive numerical and analytical calculations of the structural and dynamical properties of DNA. When studying the problem as described above, our recognition is that the discreteness as well as nonlinearity appears to be the key elements to study the properties of DNA and other biologically important macromolecules from the viewpoint of mathematical physics. By reducing model field equations to a discrete version of the double sine-Gordon equation, it is shown that numerical solutions to a static form of the equations give a variety of local fluctuations in the structure of bases, in addition to the idealized B-form and that topological solitons described by the double sine-Gordon equation[7] or by its generalized version can exist, provided the stacking energy of the bases is much larger than the hydrogen bonding energy.

While this manuscript was under preparation, the present authors became aware of a recent review article by Dickerson,[8] who reported on the results of recent structural analysis of short double helical DNA molecules. It was shown that there exist various forms of local fluctuations in the helix structure of three forms of DNA, such as helical twist angle, base inclination, propeller twisting of bases, base rolling, etc. The first of these, which was shown to be largest among the types of fluctuations given above, is nothing but the principal result of this paper. Introducing sum functions and using heuristic arguments, Dickerson has suggested that local variations in helical twist angles and base-roll angles of very short DNA molecules are induced by particular base sequence. The conclusion arrived at in this paper for long DNA molecules with non-periodicity of force field along the helical axis neglected is that the nonlinearity and discreteness are another important ingredient to be taken into account for the study of local fluctuations in helical twist angles.

This paper is organized as follows. In the next section we introduce a model dynamical system to simulate a DNA macromolecule. In §3 we discuss the formal properties of the structure and dynamics of our model system by confining ourselves to its simplest case, i.e., dynamic plane base-rotator model. In §4 a continuum approximation is employed to show the existence of various types of topological solitons. In §5 we present the results of numerical calculations for the structural properties of our model system determined from original, discrete model. The last section is devoted to concluding remarks on the results presented in this paper.

§ 2. Model Hamiltonian

We consider the Watson-Crick model of DNA[1] with helical axis taken in the z-

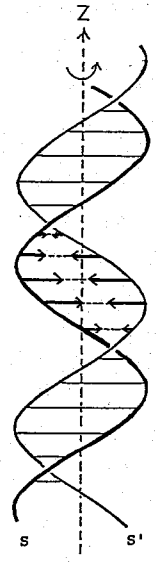


Fig. 1. Schematic feature of the B-form of DNA. The sugar-phosphate backbones of the double helix are represented as ribbons S and S' and the rung-like base pairs connecting them as arrows.

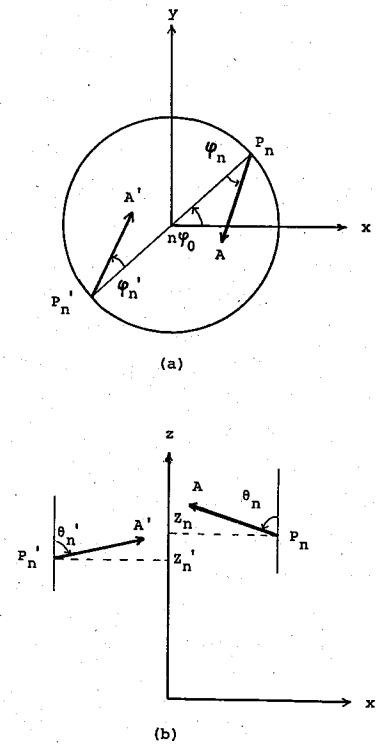


Fig. 2. Projection of the nth pair of bases (arrows) (a) in the xy-plane and (b) in the xz-plane.

direction. Its exact B-form is schematically shown in Fig. 1, where each base is depicted by an arrow with unit length, and complementary base pairs are indicated by conjugated arrows directed inward. Three fundamental assumptions are employed to study structure and dynamics of DNA: (1) The essential points of the problem can be gained by solely paying attention to bases in the double strands. (2) Fluctuations of the positions of the bases take place through their rotational motion around the points where they are attached to the strands. (3) These fluctuations give rise, under certain circumstances, to breaking of hydrogen bonding of the bases, thereby inducing unwinding of double strands to which the bases are held. Let the coordinates of P_n and P_n', which are the points where the nth base pair is attached to the one of the strand (referred to as S) and to the other strand (referred to as S') be $(a\cos n\varphi_0, a\sin n\varphi_0, z_n)$ and $(a\cos(n\varphi_0+\pi), a\sin(n\varphi_0+\pi), z_n')$, respectively, (Fig. 2) with $\varphi_0=2\pi/p (p=10)$. Here a is the radius of the circle depicted in Fig. 2, and p is the number of bases per turn in S and S'. Also, let (θ_n, φ_n) and (θ_n', φ_n') be two angles of rotation of the nth base pair around the points P_n and P_n', respectively. The quantities $(\theta_n, \varphi_n, z_n)$ and $(\theta_n', \varphi_n', z_n')$ then measure the deviation of the positions of the nth pair of bases from the exact B-form. What is the type of base-base interactions in such a model DNA system? As a preliminary to later discussion, we first employ a heuristic argument to assume that inter-strand base-base interactions or the hydrogen-bonding energy of a given base pair depend on their distance. In terms of the variables given above, the square $\overline{AA'}^2 \equiv L_n^2$ of the distance between the top of the nth arrow pair is given by

$$L_n^2=2+4a^2+(z_n-z_n')^2+2[\sin\theta_n\sin\theta_n'\cos(\varphi_n-\varphi_n')-\cos\theta_n\cos\theta_n']$$

$$-4a(\sin\theta_n\cos\varphi_n+\sin\theta_n'\cos\varphi_n')+2(z_n-z_n')(\cos\theta_n-\cos\theta_n'). \qquad (2\cdot1)$$

In this paper we limit our discussion to the case $z_n=z_n'$. This amounts to neglecting longitudinal compression waves. Here the essential ingredients of the inter-strand base-base interaction energy are the factors $\sin\theta_n\sin\theta_n'\cos(\varphi_n-\varphi_n')-\cos\theta_n\cos\theta_n'$ and $-(\sin\theta_n\cos\varphi_n+\sin\theta_n'\cos\varphi_n')$. Since we have characterized the rung-like base pairs connecting them as planks as well as sugar-phosphate backbones of the double helix as ribbons, it is not altogether meaningless to assume that the rotational motion of the bases takes place predominantly in planes perpendicular to the helical axis. The two factors then reduce to $\cos(\varphi_n-\varphi_n')$ and $-(\cos\varphi_n+\cos\varphi_n')$. It is seen that the former depends only on the relative angle of the base pair, having the tendency of keeping a pair of bases antiparallel to each other. While, the latter comes from the local-field energy responsible for the exact B-form of DNA, where helical twist angle from one base pair to the next is 36 degree, corresponding to 10 base pairs per 360-degree turn.

Under the assumption that essential features of structure and dynamics of DNA can be gained by such a plane base-rotator model, we take the inter-strand base-base interaction energy $v(\varphi_n, \varphi_n')$ of the nth base pair to be of the form

$$v(\varphi_n, \varphi_n')=h_n(2-\cos\varphi_n-\cos\varphi_n')-\lambda_n[1-\cos(\varphi_n-\varphi_n')]. \qquad (2\cdot2)$$

Here the quantities h_n and λ_n are interaction constants characterizing the local-field energy and the hydrogen-bonding energy of the nth base pair, respectively. We are next concerned with the intra-strand base-base interaction or the stacking energy of the bases. Since this has the tendency of keeping neighbouring bases parallel to each other, we take the total energy $\sum_n U(\varphi_n, \varphi_n')$ of the intra-strand interaction to be of the form

$$\sum_n U(\varphi_n, \varphi_n')=\sum_n\{J_n[1-\cos(\varphi_{n+1}-\varphi_n)]+J_n'[1-\cos(\varphi_{n-1}-\varphi_n)]\}. \qquad (2\cdot3)$$

Here J_n and J_n' are interaction constants associated with the nth bases in S and S', respectively. Stationary configurations of the bases in our model DNA system are obtained by minimizing the total potential energy with respect to φ_n and φ_n'.

The dynamical properties of our model system can be studied by adding the kinetic energy of the bases to this. Thus we take the Hamiltonian in the form

$$H=\sum_n[(I_n/2)\dot{\varphi}_n^2+(I_n'/2)\dot{\varphi}_n'^2+U(\varphi_n, \varphi_n')+v(\varphi_n, \varphi_n')]. \qquad (2\cdot4)$$

Here I_n and I_n' are the moments of inertia of the nth bases in S and S'. It is seen that here a DNA molecule is simulated as longitudinally and transversally coupled plane rotators. Equation (2·4) is obviously a generalization of the Frenkel-Kontrova model,[9] which has recently been a subject of renewed interest in solid state physics.[10] It is also a slight generalization of the model Hamiltonian employed by the present authors in (I).[6] Since the model Hamiltonian (2·4) is too restrictive in studying the structural and dynamical properties of DNA, an attempt is made in the Appendix to generalize it to the case of three-dimensional rotators.

§ 3. On-site potential

Some of the qualitative properties of our model system (2·4) can be gained by solely paying attention to the on-site potential $v(\varphi_n, \varphi_n')$, the dimensionless form of which is written as

$$\bar{v}(\varphi_n, \varphi_n') \equiv v(\varphi_n, \varphi_n')/h_n$$
$$= 2-\eta_n-\cos\varphi_n-\cos\varphi_n'+\eta_n\cos(\varphi_n-\varphi_n') \qquad (3\cdot1)$$

with

$$\eta_n=\lambda_n/h_n\,. \qquad (3\cdot2)$$

Depending upon the value of η_n, it takes several forms. To see this, we first look into $\bar{v}(\varphi_n, \varphi_n')$ for $\varphi_n=\varphi_n'$ and $\varphi_n=-\varphi_n'$. In these two specific cases the on-site potential takes the form

$$\bar{v}(\varphi_n, \varphi_n')=\begin{cases} 2(1-\cos\varphi_n) & \text{for } \varphi_n=\varphi_n' \qquad (3\cdot3a) \\ 2(1-\cos\varphi_n)-\eta_n(1-\cos2\varphi_n) & \text{for } \varphi_n=-\varphi_n\,. \qquad (3\cdot3b) \end{cases}$$

It is seen that $\bar{v}(\varphi_n, \varphi_n')$ in the direction of the lines $\varphi_n=\varphi_n'$ and $\varphi_n=-\varphi_n'$ has the form of a potential function for the sine-Gordon equation and the double sine-Gordon equation, respectively. The on-site potential for the latter case for $0\leq\varphi_n\leq2\pi$ has three forms as shown in Fig. 3, where three regions (1), (2) and (3) are defined, the boundaries (1)-(2) and (2)-(3) being at $\eta_n=-1/2$ and $\eta_n=1/2$, respectively. In region (1) $(-1/2<\eta_n<1/2)$ $\bar{v}(\varphi_n, -\varphi_n)$ has absolute minimum points at $\varphi_n=0$ and 2π and an absolute maximum point at $\varphi_n=\pi$. In region (2) $(\eta_n<-1/2)$ it has absolute and relative minimum points at $\varphi_n=0$, 2π and π, respectively, and maximum points at $\varphi_n=\cos^{-1}(1/2\eta_n)$ and $2\pi-\cos^{-1}(1/2\eta_n)$. In region (3) $(\eta_n>1/2)$ $\varphi_n=0$, 2π and $\varphi_n=\pi$ are relative and absolute maximum points, respectively, while the absolute minimum points are given by $\varphi_n=\cos^{-1}(1/2\eta_n)$ and $2\pi-\cos^{-1}(1/2\eta_n)$. It is worth noticing here that in case (3) the exact B-form of DNA having the configuration $\varphi_n=0$ or 2π for all n is no longer the ground state of the system, but it is rather an energy maximum point.

With the above preliminary consideration for the on-site potential, we show in Fig. 4

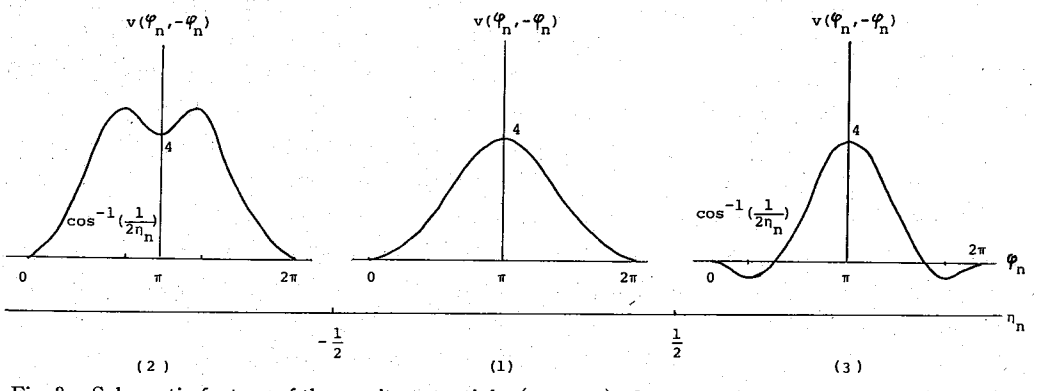


Fig. 3. Schematic feature of the on-site potential $\bar{v}(\varphi_n,-\varphi_n)=2-\eta_n-2\cos\varphi_n+\eta_n\cos2\varphi$ for (1) $-1/2<\eta_n<1/2$, (2) $\eta_n<-1/2$ and (3) $\eta_n>1/2$.

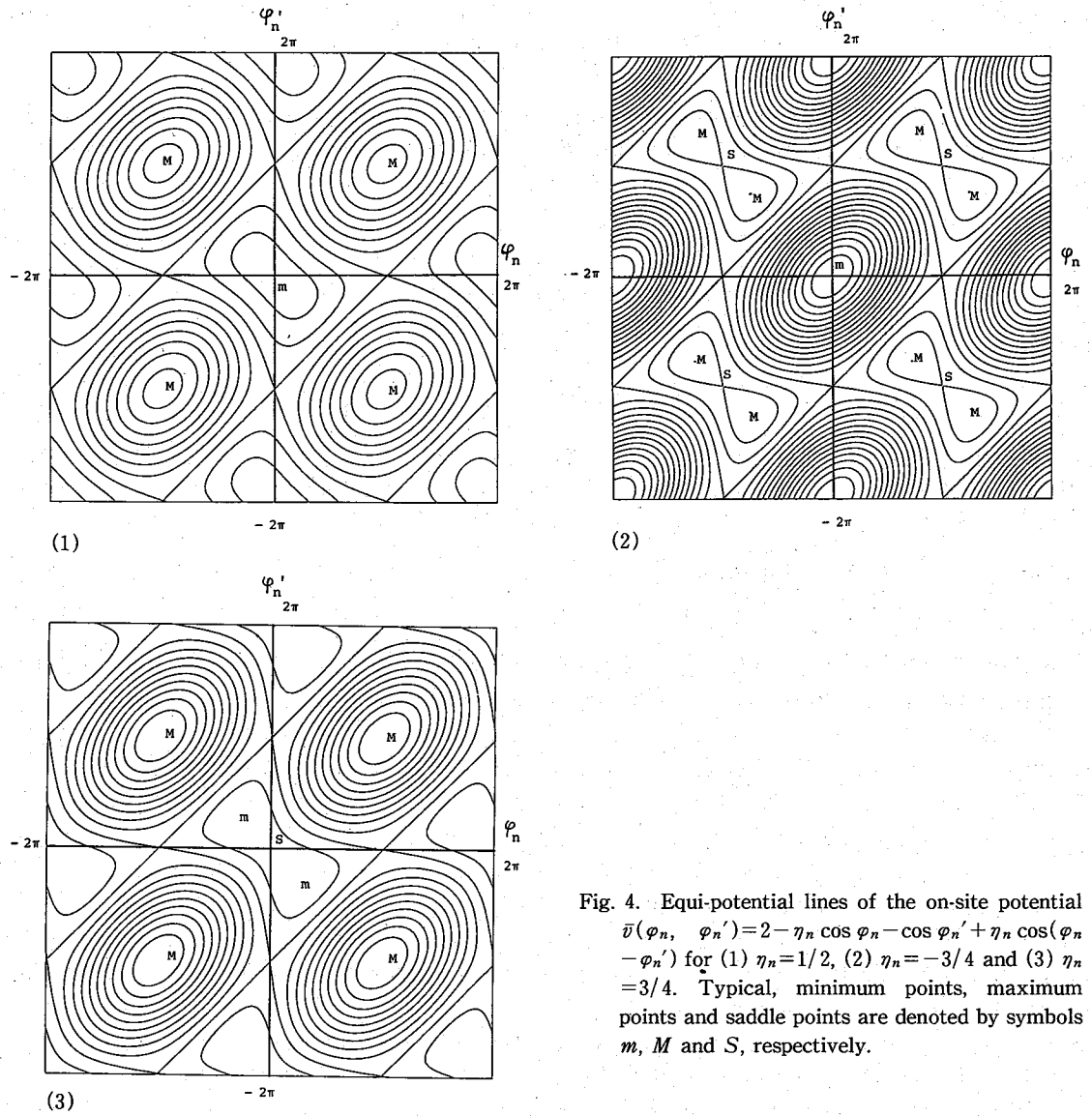


Fig. 4. Equi-potential lines of the on-site potential $\bar{v}(\varphi_n, \varphi_n') = 2 - \eta_n \cos\varphi_n - \cos\varphi_n' + \eta_n \cos(\varphi_n - \varphi_n')$ for (1) $\eta_n = 1/2$, (2) $\eta_n = -3/4$ and (3) $\eta_n = 3/4$. Typical, minimum points, maximum points and saddle points are denoted by symbols m, M and S, respectively.

equi-potential lines of $\bar{v}(\varphi_n, \varphi_n')$ for $\eta_n = -3/4$, $1/2$ and $3/4$ which correspond to regions (2), (1) and (3), respectively, for the case $\varphi_n = -\varphi_n'$. It is seen that essential features of contours of the potential $\bar{v}(\varphi_n, \varphi_n')$, i.e., minimum and maximum points, absolute or relative, are contained in the specific cases $\varphi_n = -\varphi_n'$ and $\varphi_n = \varphi_n'$. Figures 3 and 4 are useful to study qualitatively the nature of solitons in our model dynamical system when the continuum approximation is employed. This will be done in §4.

As preliminaries to the discussion given in §5, we describe here a procedure to investigate the structural properties of our model DNA system from Hamiltonian (2·4). A brief discussion on the dynamical properties using a continuum approximation will be given in §4. In doing this we limit our discussion henceforth to the case in which all the J_n's, J_n''s, I_n's, I_n''s, h_n's, λ_n's are taken to be independent of the site index n, thus omitting the subscript n attached to these quantities hereafter. We also confine ourselves to the case $J = J'$ and $I = I'$. This is a highly idealized model of DNA, in which the difference

of inter-strand and intra-strand interaction constants for different base sequence is neglected.

Stationary spatial configurations of the bases are determined by the extrema of the potential function $V(\varphi_n, \varphi_n')\equiv U(\varphi_n, \varphi_n')+v(\varphi_n, \varphi_n')$ in Eq. (2·4) with respect to φ_n and φ_n'. These are determined by the equations

$$\sin(\varphi_{n+1}-\varphi_n)-\sin(\varphi_n-\varphi_{n-1})=g\{\sin\varphi_n-\sin(\varphi_n-\varphi_n')\}, \tag{3·4}$$

$$\sin(\varphi_{n+1}'-\varphi_n')-\sin(\varphi_n'-\varphi_{n-1}')=g\{\sin\varphi_n'-\sin(\varphi_n'-\varphi_n)\}, \tag{3·5}$$

where

$$g=h/J\,. \tag{3·6}$$

Equations (3·4) and (3·5) constitute simultaneous nonlinear difference equations, which, in addition to trivial solutions $\varphi_n=\varphi_n'=0=2\pi$ corresponding to the perfect B-form of DNA, yield a number of solutions. The solutions are generally classified into two types, the ones yielding metastable configurations of the bases and the others corresponding to unstable ones. This implies that there exists a variety of local fluctuations of helical twist angles from one base pair to the next. In the specific cases $\varphi_n=\varphi_n'$ and $\varphi_n=-\varphi_n'$ Eqs. (3·4) and (3·5) are decoupled to give a single equation

$$\sin(\varphi_{n+1}-\varphi_n)-\sin(\varphi_n-\varphi_{n-1})=g\sin\varphi_n \qquad \text{for } \varphi_n=\varphi_n', \tag{3·7}$$

$$\sin(\varphi_{n+1}-\varphi_n)-\sin(\varphi_n-\varphi_{n-1})=g(\sin\varphi_n-\eta\sin2\varphi_n) \qquad \text{for } \varphi_n=-\varphi_n'\,. \tag{3·8}$$

Equation (3·7) is similar to those treated by Aubry[10] and others and by Greene,[11] Chirikov[12] and others to study the Frenkel-Kontrova model in solid state physics and the motion of particles in plasmas, respectively. One important difference here is that the difference factor on the left-hand side is of the form $\sin(\varphi_{n+1}-\varphi_n)-\sin(\varphi_n-\varphi_{n-1})$, in contrast to the conventional one $\varphi_{n+1}+\varphi_{n-1}-2\varphi_n$. Equations (3·7) and (3·8), which are specific forms of Eqs. (3·4) and (3·5), are not generally analytically tractable. Numerical solutions to the former two equations will be studied in §5.

§ 4. Continuum approximation and topological solitons

Equations of motion of the bases are readily obtained from Eq. (2·4) as follows:

$$I\ddot{\varphi}_n=J[\sin(\varphi_{n+1}-\varphi_n)-\sin(\varphi_n-\varphi_{n-1})]+\lambda\sin(\varphi_n-\varphi_n')-h\sin\varphi_n\,, \tag{4·1}$$

$$I\ddot{\varphi}_n'=J[\sin(\varphi_{n+1}'-\varphi_n')-\sin(\varphi_n'-\varphi_{n-1}')]+\lambda\sin(\varphi_n'-\varphi_n)-h\sin\varphi_n'\,. \tag{4·1'}$$

In studying the properties of our model dynamical system, we limit our discussion to the specific case in which the intra-strand interaction energy J is much larger than the local field energy h and the hydrogen-bonding energy λ. We can then employ a continuum approxition to reduce Eqs. (4·1) and (4·1') to

$$\frac{\partial^2\varphi}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2}=\frac{1}{l_0^2}[\sin\varphi-\eta\sin(\varphi-\varphi')], \tag{4·2}$$

$$\frac{\partial^2\varphi'}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2\varphi'}{\partial t^2}=\frac{1}{l_0^2}[\sin\varphi'-\eta\sin(\varphi'-\varphi)], \tag{4·2'}$$

where

$$c^2=Ja^2/I\,,\qquad l_0{}^2=Ja^2/h\,, \tag{4·3}$$

in which a is defined by $z_{n+1}-z_n=a$. We assume that φ and φ' depend on z and t only through the variable $z-vt$, where v is a constant. Equation (4·2) and (4·2′) can then be integrated once to give

$$\frac{l_0{}^2}{2}\left[\left(\frac{d\varphi}{d\bar{z}}\right)^2+\left(\frac{d\varphi'}{d\bar{z}}\right)^2\right]+[-\bar{v}(\varphi,\ \varphi')]=0\,, \tag{4·4}$$

where

$$\bar{z}=[1-(v^2/c^2)]^{-1/2}(z-vt) \tag{4·5}$$

is the rest-frame coordinate for the system which is Lorentz-invariant. Here we have chosen the integral constant suitably to get $-\bar{v}(\varphi,\ \varphi')$ defined by Eq. (3·1). We observe that Eq. (4·4) is analogous to the classical equation of motion for a Newtonian particle with mass $l_0{}^2$ under the influence of a potential field $-\bar{v}(\varphi,\ \varphi')$ in a two-dimensional space $(\varphi,\ \varphi')$ with $\bar{z}$ playing the role of time. It is seen from Eq. (4·4) and Fig. 4 that here topological solitons or kinks which correspond to excitations connecting degenerate minima of $\bar{v}(\varphi,\ \varphi')$ exist. In the region $-1/2<\eta<1/2$ the solitons are due to the existence of degenerate, absolute minima of $\bar{v}(\varphi,\ \varphi')$. In the regions $\eta<-1/2$ and $\eta>1/2$ situations are somewhat more involved, since there exist absolute and relative minima and maxima for the former and the latter, respectively, for $\varphi=-\varphi'$.

Since Eqs. (4·2) and (4·2′) are not generally analytically tractable, we content ourselves here with considering the case $\varphi=-\varphi'$ to obtain explicit soliton solutions in our model dynamical system. Equations (4·2) and (4·2′) reduce to the double sine-Gordon equation

$$\frac{\partial^2\varphi}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2}=\frac{1}{l_0{}^2}(\sin\varphi-\eta\sin 2\varphi)\,. \tag{4·6a}$$

A corresponding equation for the case $\varphi=\varphi'$ is given by

$$\frac{\partial^2\varphi}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2}=\frac{1}{l_0{}^2}\sin\varphi\,, \tag{4·6b}$$

which is nothing but the sine-Gordon equation. Since a detailed discussion on the properties of soliton solutions to the double sine-Gordon equation has been given by Condat, Guyer and Miller,[13] we present here only the result of calculations, treating the cases $\eta<-1/2$, $-1/2<\eta<1/2$ and $\eta>1/2$ separately.

(1) region(1): $-1/2<\eta<1/2$, *2π-kinks*

This is the simplest case of all these three cases, corresponding to region (1) in Fig. 3. The potential function $\bar{v}(\varphi,\ -\varphi)$ is a simple sinusoidal function, and 2π-kink soliton solutions connecting the absolute mimima $\varphi=0$ and 2π is given by

$$\varphi\equiv\varphi_1{}^{(1)}=2\tan^{-1}[\pm(1-2\eta)^{1/2}\mathrm{cosech}\{(1-2\eta)^{1/2}\bar{z}\}]. \tag{4·7}$$

Here the plus and minus signs denote kink and anti-kink solutions, respectively. For $\eta=0$ the solutions reduce to the sine-Gordon kink solutions.

(2) region (2): $\eta<-1/2$, *2π-kinks and bubbles*

Here we have two types of solutions

$$\varphi\equiv\varphi_1^{(2)}=2\tan^{-1}[\pm(1-2\eta)^{1/2}\mathrm{cosech}\{(1-2\eta)^{1/2}\bar{z}\}], \tag{4·8}$$

$$\varphi\equiv\varphi_2^{(2)}=2\tan^{-1}[\pm(-1-2\eta)^{1/2}\cosh\{(1-2\eta)^{1/2}\bar{z}\}]. \tag{4·9}$$

Solution (4·9), which continuously goes over to Eq. (4·7) is, as before, represents 2π-kink solutions. Equation (4·9) represents finite fluctuations of φ from the local minimum point $\varphi=\pi$ of $\bar{v}(\varphi, -\varphi)$. Excitations corresponding to the latter, called critical bubbles are, however, unstable.

(3) region (3): $\eta>1/2$, *large kinks and small kinks*

As mentioned before, a characteristic feature is that the conventional perfect B-form is not the ground state, but the points $\varphi=0$ and 2π are rather energy local maximum points. Here two types of kink solutions exist:

$$\varphi\equiv\varphi_1^{(3)}=2\tan^{-1}\left[\pm\left(\frac{2\eta-1}{2\eta+1}\right)^{1/2}\coth\left\{\left(\frac{4\eta^2-1}{8\eta}\right)^{1/2}\bar{z}\right\}\right], \quad \text{(large kinks)} \tag{4·10}$$

$$\varphi\equiv\varphi_2^{(3)}=2\tan^{-1}\left[\pm\left(\frac{2\eta-1}{2\eta+1}\right)^{1/2}\tanh\left\{\left(\frac{4\eta^2-1}{8\eta}\right)\bar{z}\right\}\right]. \quad \text{(small kinks)} \tag{4·11}$$

The former and the latter correspond to kink excitations which connect two degenerate relative minimum points $\varphi=\cos^{-1}(1/2\eta)$ and $2\pi-\cos^{-1}(1/2\eta)$ across the absolute maximum point $\varphi=\pi$ and the relative maximum point $\varphi=0$, respectively, of $\bar{v}(\varphi, -\varphi)$.

§ 5. Local fluctuations in helical twist angles

In the preceding section we studied the dynamical properties of our model DNA system by using the continuum approximation. By such an approximation procedure, however, several of the essential features of the problem are lost. For example, Eqs. (4·6a) and (4·6b) are always integrable, exhibiting propagating topological soliton solutions, whereas their original, discrete form, Eqs. (4·1) and (4·1′) for $\varphi_n=-\varphi_n'$ and $\varphi_n=\varphi_n'$, is not generally integrable. In view of this non-integrability of the equations, we study in this section numerical solutions to Eqs. (3·7) and (3·8) and their implications to the structural properties of DNA. We follow here the conventional procedure to recast Eqs. (3·7) and (3·8) in the form of mapping

$$\sin\omega_{n+1}=\sin\omega_n+g\begin{cases}\sin\varphi_n & \text{for Eq. (3·7)} \quad (5\cdot1\cdot1)\\ \sin\varphi_n-\eta\sin2\varphi_n & \text{for Eq. (3·8),} \quad (5\cdot1\cdot2)\end{cases}$$

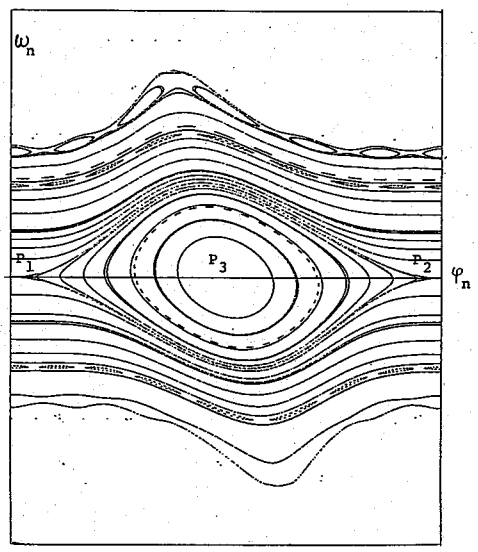
$$\varphi_{n+1}=\varphi_n+\omega_{n+1}\,. \tag{5·1′}$$

Numerical solutions to Eq. (5·1) can then be generated by starting at one point in (φ, ω) space and iterating the equations. When the spatial variation of the φ_n's is fairly smooth, which is realized for small g, the quantity $\sin\omega_n$ in Eqs. (5·1·1) and (5·1·2) can be replaced by ω_n itself. Equations (5·1·1) and (5·1′) then reduce to equations in the Frenkel-Kontrova model[9),10)] or standard map equations studied extensively as a model for certain dynamical systems[11),12)] (for example, the motion of ions in a plasma). Three

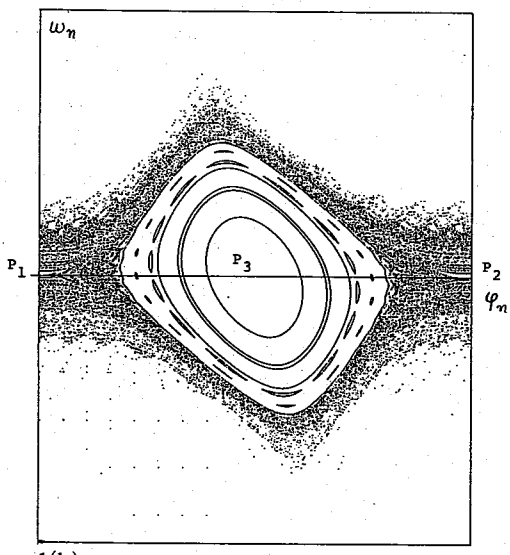
features exist for solutions to Eq. (5·1) depicted in (φ, ω)-space: (i) fixed points, hyperbolic or elliptic, giving commensurate phases, (ii) invariant trajectories yielding incommensurate or commensurate phases and (iii) chaotic or disordered phases corresponding to clouds of points that are not analytical curves. The qualitative properties of these features are then deduced from the theory developed by Kolmogorov, Arnold and Moser.[14] Namely, for small g all trajectories which exist in integral systems in the continuum limit persist, albeit deformed. As g increases, fewer and fewer of the KAM trajectories remain, and their disappearance is connected with the appearance of chaotic states.

We confine ourselves to Eq. (5·1·2) (and Eq. (5·1′)) in doing numerical calculations, since the case of Eq. (5·1·1) is qualitatively similar to the former in region (1) $(-1/2 < \eta < 1/2)$. Three cases $g=0.1$, 0.2 and 0.3 are considered as examples for each of $\eta = -1.0$, $1/2$ and 2.0, which correspond to regions (2), (1) and (3), respectively, of the potential function $\bar{v}(\varphi, -\varphi)$(see Fig. 3). The results of numerical calculations are shown in Figs. 5.1, 5.2 and 5.3. When discussing the obtained results, we first consider the case $\eta=1/2$ which is nearest to the original perfect B-form. Here $P_1(0, 0)$ and $P_2(2\pi, 0)$ are hyperbolic (fixed) points corresponding to the exact B-form, while $P_3(\pi, 0)$ is an elliptic point giving configurations in which all the base pairs are directed in outward direction. It is seen from Fig. 5.1(a) that for $g=0.1$ corresponding to strong stacking energy or weak hydrogen-bonding energy there appear smooth invariant curves as expected and several islands outside the separatrix which is a trajectory connecting the points P_1 and P_2. Some of the curves here appear to be dashed lines. This is because the iterations were stopped before the curves were traced out. Of various types of the trajectories the separatrix corresponds to solitons which exist in the continuum, integrable limit. The trajectories outside the separatrix correspond to multi-periodic phases, while those encirculating the point P_3 are unstable states corresponding to local fluctuations of bases about the energy maximum point, having energy higher than those outside the separatrix. As g increases (Fig. 5.1(b)) the trajectories around the point P_3 persist, and there appear islands corresponding to new multi-periodic phases. In contrast to this, only traces of the separatrix are seen near P_1 and P_2, and the trajectories and the islands near and outside the separatrix become diffuse. It is worth mentioning here that when the initial configuration is chosen sufficiently close to P_1 or P_2, the trajectories change in a dramatic way, and the points do not form invariant curves but tend to fill out a finite area in (φ, ω) space. Such area represents chaotic states which correspond to the randomness of the helical twist angles from one base pair to the next. It is seen that invariant trajectories having lower energies are strongly influenced by the increase of g, while those with higher energy around the point P_3 are little affected by the discreteness effect. As g increases further (Fig. 5.1(c)) the numerical result becomes rather simple. Namely, clouds of points in the vicinity of P_1 and P_2 and of the separatrix which exist in the case $g=0.1$ decrease so much as they are nearly countable, including multi-periodic phases, though the area encircled by such trajectories decreases slightly. These results imply that higher excited states corresponding to unstable states are less influenced by the discreteness effect, while those metastable states in the vicinity of P_1 and P_2 and of the separatrix are strongly perturbed.

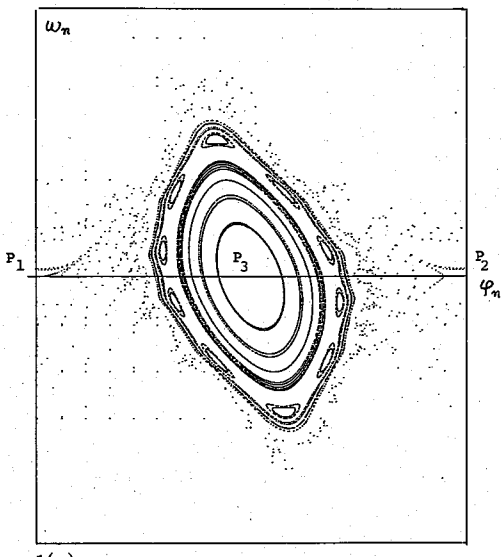
The above-mentioned features of the numerical solutions remain unchanged for the region $\eta < -1/2$ (Fig. 5.2(a)~(c)). Three points are worth noticing here: (1) A separatrix connecting P_1 and P_2 corresponds to 2π-kinks which exist in the continuum



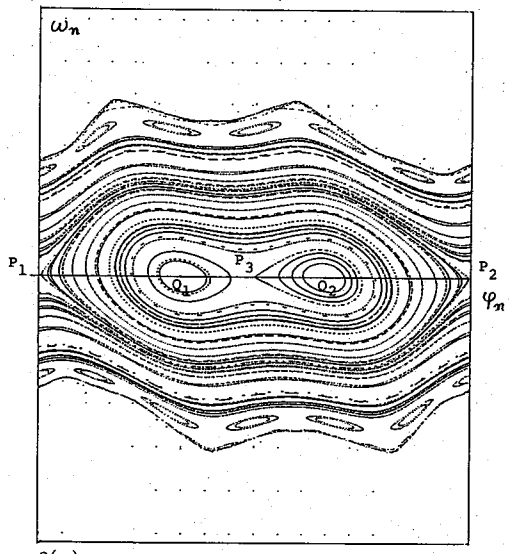
1(a) $\eta = 0.5,\ g = 0.1$



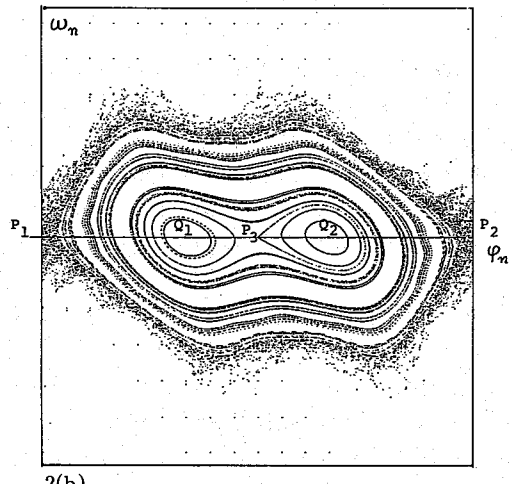
1(b) $\eta = 0.5,\ g = 0.2$



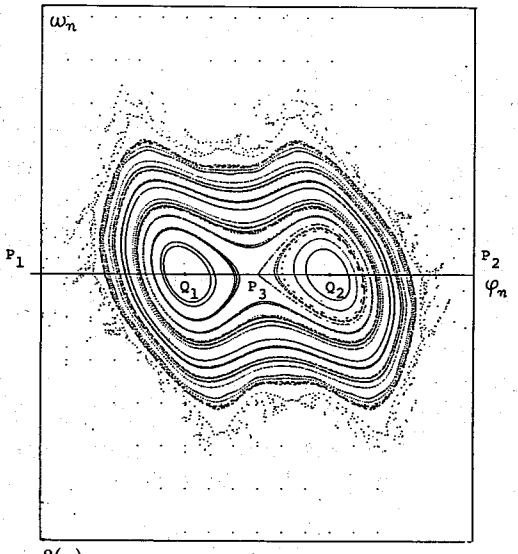
1(c) $\eta = 0.5,\ g = 0.3$



2(a) $\eta = -1.0,\ g = 0.1$



2(b) $\eta = -1.0,\ g = 0.2$



2(c) $\eta = -1.0,\ g = 0.3$

Fig. 5. (continued on the next page)

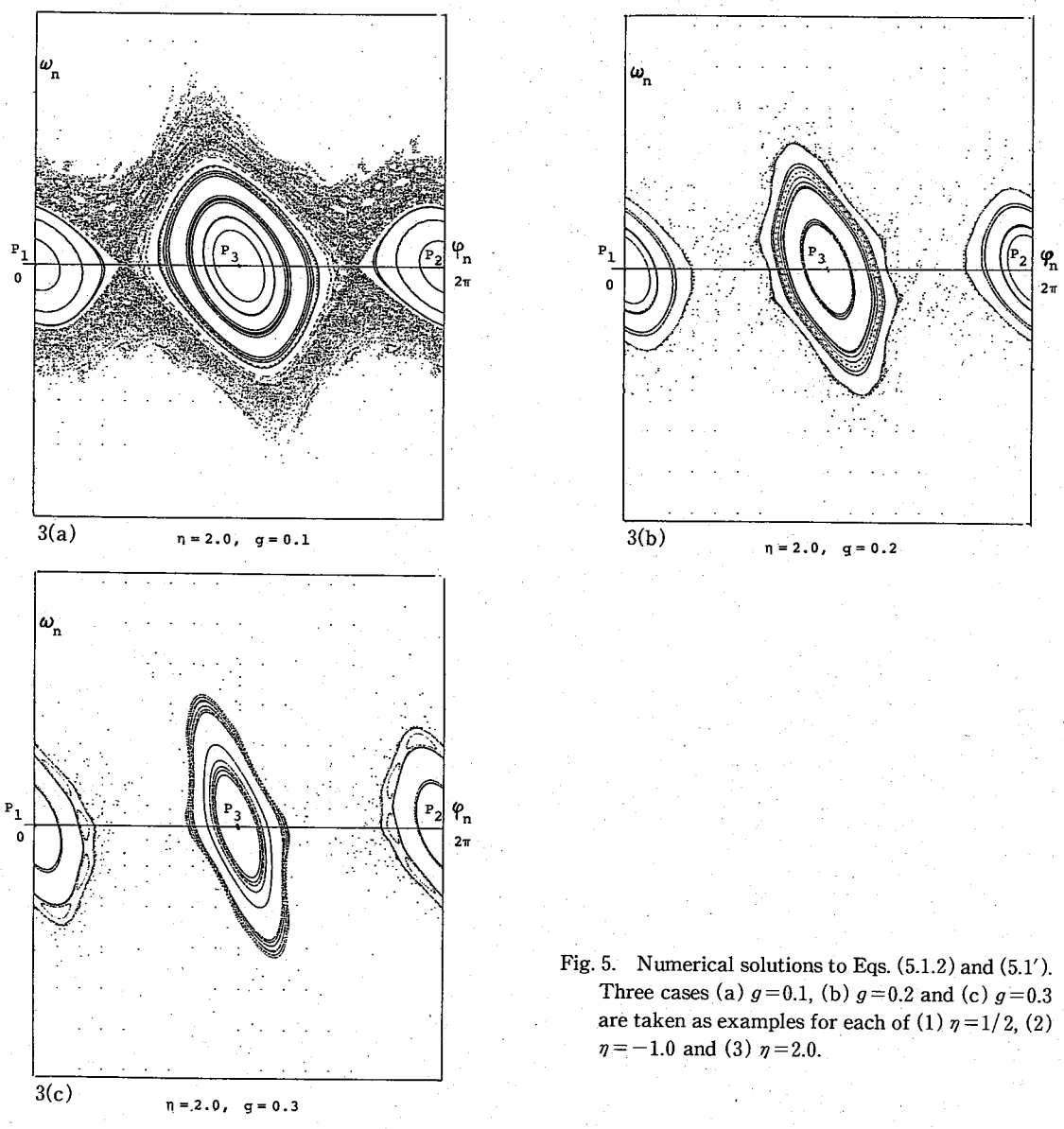


Fig. 5. Numerical solutions to Eqs. (5.1.2) and (5.1′). Three cases (a) $g=0.1$, (b) $g=0.2$ and (c) $g=0.3$ are taken as examples for each of (1) $\eta=1/2$, (2) $\eta=-1.0$ and (3) $\eta=2.0$.

limit, while those encirculating the energy-maximum point $Q_1(\cos^{-1}(1/2\eta),\ 0)$ and $Q_2(2\pi-\cos^{-1}(1/2\eta),\ 0)$ are bubble excitations. (2) It is seen from Fig. 5.2(a) that the invariant trajectories appear to be a little more influenced than those in the case $-1/2<\eta<1/2$ by the discreteness effect. (3) Since there exist two energy maximum points Q_1 and Q_2, however, the area enclosed by the trajectories encirculating these points is larger than that in region (1). This implies that as g increases the area of chaotic region gets smaller than that in case (1).

In region (3) $\eta>1/2$ P_1 and P_2 are no longer energy minimum points, but they are rather local energy maximum points. As mentioned before, here two kinds of kinks, large ones and small ones, exist in the continuum limit. As seen from Fig. 5.3(a), however, two kinds of separatrices corresponding to these kink solutions and invariant trajectories in their vicinity almost disappear even for the case $g=0.1$, leaving only

trajectories encirculating the energy absolute maximum point P_3 and energy local maximum points P_1 or P_2. It is seen that in this region of η the effect of the discreteness is largest of all these three cases. As g increases from 0.1 to 0.2 and 0.3 (Figs. 5.3(b) and (c)), situations become rather similar to the previous two cases. Namely, clouds of points which exist in the case $g=0.1$ considerably decrease as g increases, and the area occupied by invariant trajectories encirculating the energy maximum points P_3 and P_1 or P_2 decreases.

The results obtained in this section imply that in addition to the orderd B-form of DNA corresponding to the solution $\varphi_n=\varphi_n'=0 \pmod{2\pi}$, there exists a variety of local fluctuations in helical twist angles from one base pair to the next. Here configurations of physical importance are metastable states in which φ_n and φ_n' are either commensurate, incommensurate or chaotic with respect to the 2π-periodicity of the inter-strand potential $\bar{v}(\varphi_n, \varphi_n')$. For the case $\varphi_n=-\varphi_n'$ such metastable configurations correspond to points in (φ, ω)-space which lie in close vicinity of energy minimum points and of separatrices.

§ 6. Concluding remarks

The inherent structural and dynamical flexibility of double-stranded DNA is important because of its potential role in the mechanisms of, for example, DNA replication, transcription, drug binding. In spite of an enormous number of degrees of freedom contained in a DNA macromolecule, a simple view of the flexibility is provided by classifying its motion principally into two types, one is associated with hydrogen-bonding and stacking and the other with sugar-phosphate backbones. In this paper we have focussed our attention to the former which arises from the motion of bases, simulating DNA as a system of transversally and longitudinally coupled base-rotators. From the viewpoint of mathematical physics, this is a generalized version of the Frenkel-Kontrova model, where a model field equation constitutes a coupled pair of nonlinear differential difference equations. In a specific yet physically interesting case, these equations are decoupled to give a discrete-version of the double sine-Gordon equation. The principal result obtained here for such a simplified model system is the existence of various types of topological solitons in the continuum approximation for the dynamical properties and the appearance of fairly large local variation in helical twist angles from one base pair to the next for the structural properties. Among various structural fluctuations those corresponding to energy-local-minimum states are, of course, most physically interesting.

Although we feel the model employed here is non-trivial for the study of the problem of DNA flexibility, it still suffers from a number of simplifications. Non-periodicity of force field along the helical axis due to a particular base sequence, treating bases as three-dimensional rotators, flexibility of sugar-phosphate backbones are all likely to improve and refine the picture presented here. In particular, the first of these may further enhance structural fluctuations, leading to sequence-induced local fluctuations. These points will be examined elsewhere. What are implications of these effects? One thing which easily comes to mind is that large local structural fluctuations or topological solitons, if any, may induce local unwinding of double strands. These and related effects on DNA problems in molecular biology are well worth studying.

Acknowledgements

One of the present authors (S. H.) would like to express his sincere thanks to Dr. T. Murai for his advice in numerical work. The numerical calculations were performed at Nagoya University, Computation Center.

Appendix

—— Classical-Spin Model of Bases in DNA ——

Here an attempt is made to generalize Eq. (2·4) to the case of three-dimensional rotators by using analogy with spin model Hamiltonian in magnetism. We observe that in terms of a quasi-spin operator $\boldsymbol{S}_n=(S_n^x, S_n^y, S_n^z)$ with

$$S_n^x=\sin\theta_n\cos\varphi_n, \qquad S_n^y=\sin\theta_n\sin\varphi_n, \qquad S_n^z=\cos\theta_n \tag{A·1}$$

for the nth base in S and a corresponding one in S', Eq. (2·1) with $z_n=z_n'$ is rewritten as

$$L_n^2=2(S_n^xS_n'^x+S_n^yS_n'^y-S_n^zS_n'^z)-4a(S_n^x+S_n'^x)+\text{const}. \tag{A·2}$$

It is seen that L_n^2 is written in the form of a generalized Heisenberg model. Then, Eqs. (2·2) and $U(\varphi_n, \varphi_n')$ in Eq. (2·3) are a specific form of $\lambda_n(S_n^xS_n'^x+S_n^yS_n'^y)-\mu_nS_n^zS_n'^z-h_n(S_n^x+S_n'^x)$ and $-J_n(S_{n+1}^xS_n^x+S_{n+1}^yS_n^y)-K_nS_{n+1}^zS_n^z-J_n'(S_{n+1}'^xS_n'^x+S_{n+1}'^yS_n'^y)-K_n'S_{n+1}'^zS_n'^z$, respectively, for $\theta_n=\theta_n'=\pi/2$. Here μ_n, K_n and K_n' are constants. To these terms we may add the anisotropy energy $A_n(S_n^z)^2+A_n'(S_n'^z)^2$ due to our characterization of rung-like base pairs as planks, where A_n and A_n' are constants. Collecting all of these terms together, we take our model base-rotator Hamiltonian to be of the form

$$\begin{aligned}H=\sum_n[&-J_n(S_{n+1}^xS_n^x+S_{n+1}^yS_n^y)-K_nS_{n+1}^zS_n^z\\&-J_n'(S_{n+1}'^xS_n'^x+S_{n+1}'^yS_n'^y)-K_n'S_{n+1}'^zS_n'^z\\&+\lambda_n(S_n^xS_n'^x+S_n^yS_n'^y)-\mu_nS_n^zS_n'^z-h_n(S_n^x+S_n'^x)\\&+A_n(S_n^z)^2+A_n'(S_n'^z)^2].\end{aligned} \tag{A·3}$$

It is understood that all the constants J_n, J_n', K_n, K_n', λ_n, μ_n, h_n, A_n, A_n' are positive. Stationary configurations of the bases in such a model DNA system are obtained by minimizing Eq. (A·3) with respect to θ_n, θ_n', φ_n, φ_n'.

Less known is the dynamical nature of such a model DNA system simulated as longitudinally and transversally coupled three-dimensional rotators. Here we merely give the remark that Eq. (2·4) itself can be derived from Eq. (A·3) for the case when the anisotropy energies A_n and A_n' are much larger than the remaining interaction constants. Namely, under the assumption that the classical equation of motion for spins:

$$\dot\theta_n=(1/\sin\theta_n)(\partial H/\partial\varphi_n), \qquad \dot\varphi_n=-(1/\sin\theta_n)(\partial H/\partial\theta_n) \tag{A·4}$$

and similar ones for $\theta_n'$and φ_n' can be used here, we get $\dot\varphi_n=2A_n\cos\theta_n$ and $\dot\varphi_n'=2A_n'\cos\theta_n'$ from the second of Eqs. (A·4) for $A_n, A_n'\gg J_n, J_n', K_n, K_n', \lambda_n, \mu_n, h_n$. In this limiting case Eq. (A·3) reduces to Eq. (2·4), where the moment of inertia $I_n(I_n')$ is

related to $A_n(A_n')$ by the relation $I_n=1/2A_n$.[15]

References

1) F. H. Crick and J. D. Watson, Proc. Roy. Soc. **A223** (1954), 80.
2) R. E. Dickerson, H. R. Drew, B. N. Conner, R. M. Wing, A. V. Fratini and M. L. Kopka, Science **216** (1982), 475 and references cited therein.
3) S. W. Englander and J. J. Englander, Proc. Natl. Acad. Sci. **53** (1965), 370.
4) S. W. Englander, N. R. Kallenbach, A. J. Heeger, J. A. Krumhansl and S. Litwin, Proc. Natl. Acad. Sci. **77** (1980), 7222.
5) S. Yomosa, Phys. Rev. **A27** (1983), 2120; **A30** (1984), 474.
6) S. Takeno and S. Homma, Prog. Theor. Phys. **70** (1983), 308.
7) See for example, R. K. Dodd, J. C. Eilbeck, J. D. Gibbon and H. C. Morris, *Solitons and Nonlinear Wave Equations* (Academic Press, 1982), chap. 10.
8) R. E. Dickerson, Scientific American, December (1983), 87.
9) J. Frenkel and T. Kontrova, J. of Phys. (USSR) **1** (1939), 137.
10) M. Peyrard and S. Aubry, J. of Phys. **C16** (1983), 1593 and also references cited therein.
11) J. M. Greene, J. Math. Phys. **9** (1968), 760; **20** (1979), 1183.
12) B. V. Chirikov, Phys. Rep. **52** (1979), 263.
13) C. A. Condat, R. A. Guyer and M. D. Miller, Phys. Rev. **B27** (1983), 474.
14) See for example, V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer-Verlag, New York-Heiderberg-Berlin, 1980), Appendix 8.
15) See also, M. Takahashi, J. Phys. Soc. Jpn **48** (1980), 746.