Conformational solitons in stacked systems

D. Hofmann, W. Förner, and J. Ladik
Chair for Theoretical Chemistry and Laboratory of the National Foundation for Cancer Research, Friedrich-Alexander-Universität Erlangen—Nürnberg, D-8520 Erlangen, Egerlandstrasse 3, Federal Republic of Germany
(Received 18 May 1987; revised manuscript received 8 December 1987)

With the help of numerical simulations using an extended Su-Schrieffer-Heeger-type Hamiltonian, the existence of solitary waves in a polyformamide stack is shown. Each formamide unit in the stack has three geometrical degrees of freedom during the time simulation. The passing of a solitary wave through the stack and its reflection at the end of the chain is shown. In addition, the collision of two solitary waves is discussed.

INTRODUCTION

The activation of oncogenes has been discussed since their discovery around 1970. One of the unsolved problems in this context is chemical carcinogenesis, the reason for about 80% of all tumors.[1] Many theories have been proposed which are based on local effects of carcinogens bound to DNA. Simple statistical considerations, however, show that the probability of removing a histone by local effects is far too small to explain carcinogenesis through chemicals.[2] Three different possibilities for activation of oncogenes by long-range effects are given in the literature.

(1) The tertiary structure of DNA is changed by the binding of a metabolized carcinogen to DNA. This is accompanied by a corresponding change in the surrounding structure of water and the distribution of K^+ ions.[3] All of these cause a change in that DNA-protein interaction which regulates an oncogene.

(2) The DNA could become a conductor via an internal "charge transfer"[4] (CT) between the DNA and the protein. The binding of a carcinogen can cause a complete filling of the valence bands of one or both of the interacting chains. Therefore, as second-order perturbation theory shows it, the dispersion and polarization forces between the two chains strongly decrease.[5,6] At the same time, the periodicity of the backbone of the polynucleotide is broken and its band structure is destroyed.[3] Both electronic effects cause the release of the protein from DNA.

(3) Let us assume that a bulky carcinogen is bound to a nucleotide base. Certainly the structure of DNA in the neighborhood of the attached carcinogen will change. *In vivo* repair enzymes can remove the carcinogen within a few hours and start a solitary wave.[8] Due to their long lifetime, they can cover large distances in DNA (e.g., see Ref. 7). This long-range effect can influence the interaction between the DNA and the histone to a great extent.[8,9]

Solitons are localized excitations rather than wave packets of large amplitude. They propagate without change of shape or loss of energy. They are stable to small perturbations. Moreover, they have the remarkable property of surviving collisions with other solitons. Solitons arise in systems which are both nonlinear and dispersive.[10] Waves are called solitary waves, if they do not exactly fulfill the soliton conditions (for instance, they have long but not infinite lifetimes).

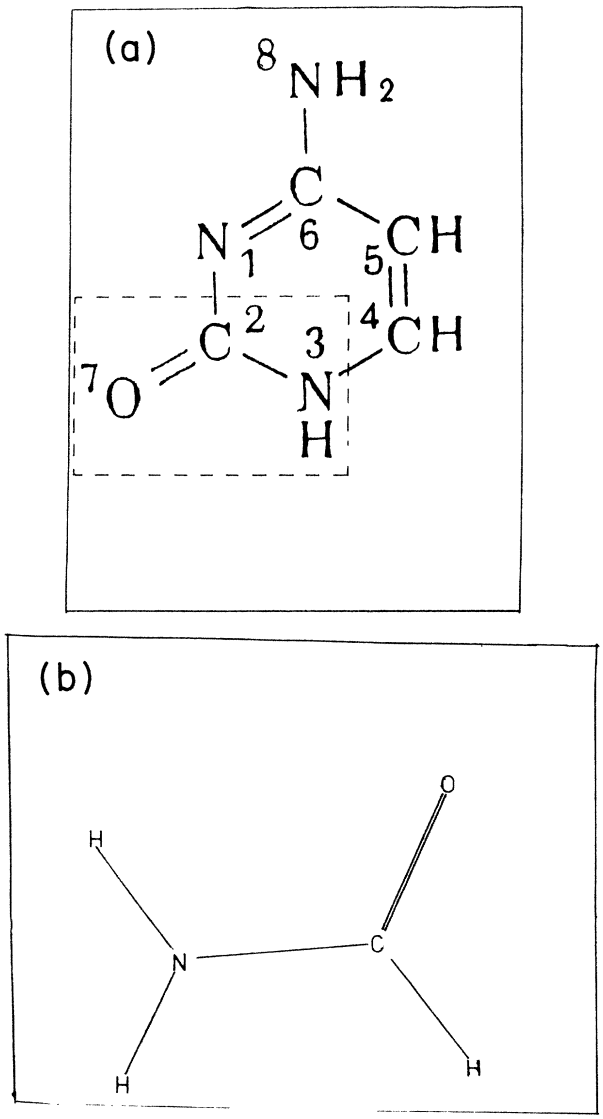

FIG. 1. (a) Chemical structure of cytosine. (b) Chemical structure of formamide.

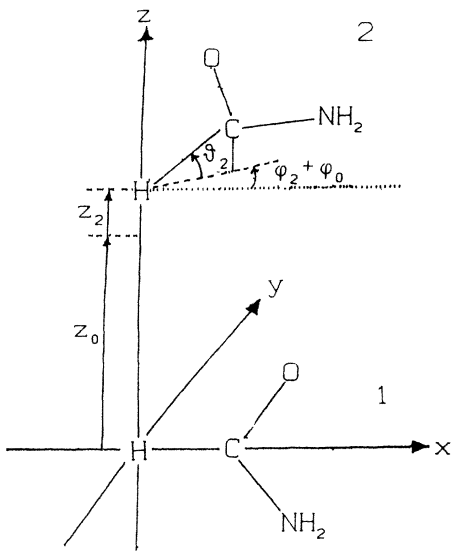

FIG. 2. Geometry of a formamide dimer used for the determination of the electron-phonon coupling constant.

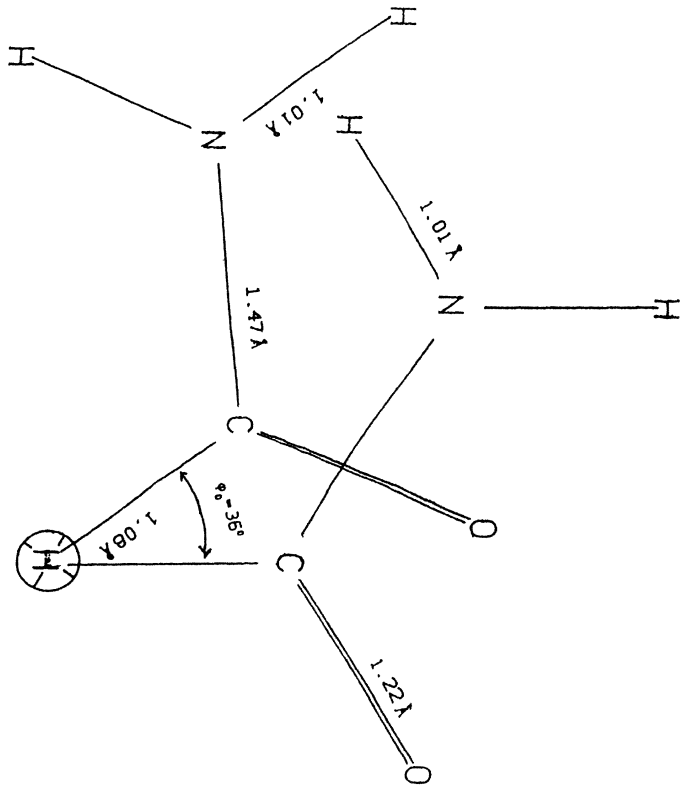

FIG. 3. Projection of the calculated formamide dimer (all angles 120°, except φ_0) along the helix axis. This axis goes through the H atoms of the C=O group of the formamide molecules.

Solitons were used by Davydov to explain the transport and reserve of energy of proteins. They can travel along the hydrogen bonds perpendicular to the main polypeptide chain in a helical conformation,[11]

$$\cdots \text{H—N—C=O} \cdots \text{H—N—C=O} \cdots .$$

In this case, the coupling of the high-frequency vibrations along the C=O bond with the low-frequency vibrations

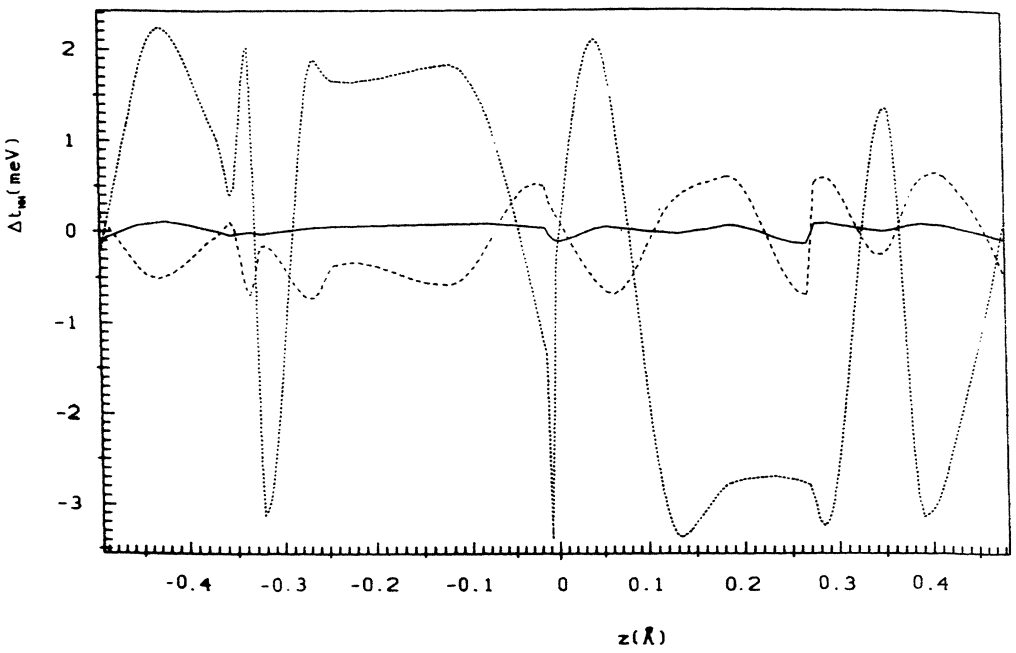

FIG. 4. Difference between the resonance integral (between the N atoms of the two formamide molecules in the dimer) calculated directly with the Hartree-Fock method and from the fitted Taylor series ($\Delta t_{N,N}$) as a function of the coordinate z ($\vartheta=\varphi=0$) for three different orders of the fit (· · · ·, fourth order; — — —, fifth order; ——, sixth order).

TABLE I. Range and largest error of the hopping integrals in 10^{-3} hartrees.

g	h	$\max(t_{g,h})$	$\min(t_{g,h})$	$\Delta\max(t_{g,h})$
N	N	35.048	0.108	0.005
N	C	36.237	1.156	0.003
N	O	25.647	−0.026	0.004
C	N	13.750	0.167	0.002
C	C	65.885	2.741	0.003
C	O	52.434	0.559	0.002
O	N	0.565	−0.092	0.010
O	C	12.338	0.235	0.002
O	O	14.876	−0.368	0.003

TABLE III. Selected force constants (10 out of 84) in 10^{-3} hartrees/bohr [i,j,k according to Eq. (8)].

i	j	k	$K_{i,j,k}$
0	0	0	−329 756.99
1	0	0	−2.03
0	1	0	−4.79
0	0	1	−2.32
2	0	0	1.96
1	1	0	9.04
1	0	1	2.32
0	2	0	11.31
0	1	1	5.97
0	0	2	1.72

along the O · · · H hydrogen bonds provides the nonlinearity which is a necessary condition for the existence of a soliton. Su, Schrieffer, and Heeger (SSH) proposed a soliton for the interpretation of the semiconductor properties of polyacetylene.[12] They define hopping integrals $t_{n,n+1}$ between two neighboring sites n and $n+1$ and postulate a linear electron-phonon coupling. The reflection[13] and the stability against small perturbations[14] have been shown with a modified SSH Hamiltonian[15] (stabilized against lattice shrinking) and in the case of soliton reflection with the help of the Pariser-Parr-Pople (PPP) Hamiltonian as well.[16]

Ladik and Čížek give the first formalism to prove the hypothesis of conformational solitary waves in DNA.[8,9] They introduce three geometrical variables for each base. Variable z_n describes the change from its equilibrium value of the stacking distance of the nth unit, variable φ_n describes the rotation of this base in the plane perpendicular to the main axis of the double helix (following Krumhansl and Alexander[17]), and the angle ϑ_n measures the tilting of the base against the plane perpendicular to the main axis (a more detailed explanation is given below). The formalism includes the spin as well as the two different strands in the DNA double helix. The electron-electron interaction is described by the PPP approximation. This Hamiltonian has been applied with small changes to a polyformamide stack as a first test for this hypothesis for chemical carcinogenesis.

In the double-helix structure of B-DNA in the B conformation, the planar nucleotide molecules are arranged in stacks. This means that the molecular planes of the bases are parallel to each other and perpendicular to the helix axis. Each base is shifted by 3.36 Å along the helix axis and rotated by 36° around it with respect to the next base. The stacked molecules are bound together by a sugar-phosphate chain, the so-called backbone.

The molecular orbitals of planar molecules such as the nucleotide bases can be divided in three classes. The inner-shell or core orbitals are localized around the nuclei and are not involved in the chemical bonding. The σ orbitals are symmetric with respect to the molecular plane. These orbitals are usually localized between two atoms. The π orbitals are antisymmetric with respect to the molecular plane and are delocalized over the whole molecule. The occupied σ and π orbitals together form the chemical bonding. Atoms whose atomic orbitals contribute to the π molecular orbitals of a molecule are called π centers. For instance, the DNA, the four different nucleotide bases C, N, and O have p_π orbitals and therefore they are "π centers." The four nucleotide bases are called guanine, cytosine, adenine, and thymine. The chemical structure of cytosine is shown in Fig. 1. To reduce the computational effort, we have used, instead of cytosine, formamide, which is a section of the cytosine molecule as shown in Fig. 1. C(2) and N(3) are saturated by two additional hydrogen atoms. The sugar-phosphate backbone of the DNA is simulated by a correction term in the potential.

METHOD

Because of the large distances between the molecules within a stacked arrangement similar to B-DNA and the

TABLE II. Electron-phonon coupling constants up to second order in hartrees/bohr.

i	j	k	$t_{i,j,k,\mathrm{N,N}}$	$t_{i,j,k,\mathrm{N,C}}$	$t_{i,j,k,\mathrm{N,O}}$	$t_{i,j,k,\mathrm{C,C}}$	$t_{i,j,k,\mathrm{O,O}}$
0	0	0	0.0035	0.0083	0.0021	0.0158	0.0001
1	0	0	−0.0048	−0.0101	−0.0035	−0.0176	−0.0008
0	1	0	−0.0167	−0.0200	−0.0121	−0.0361	−0.0037
0	0	1	−0.0096	−0.0006	0.0061	−0.0086	−0.0021
2	0	0	0.0029	0.0053	0.0026	0.0084	0.0010
1	1	0	0.0202	0.0209	0.0174	0.0351	0.0074
1	0	1	0.0137	0.0009	−0.0100	0.0101	0.0043
0	2	0	0.0341	0.0157	0.0283	0.0404	0.0153
0	1	1	0.0471	0.0024	−0.0320	0.0206	0.0150
0	0	2	0.0071	−0.0087	−0.0011	−0.0042	−0.0018

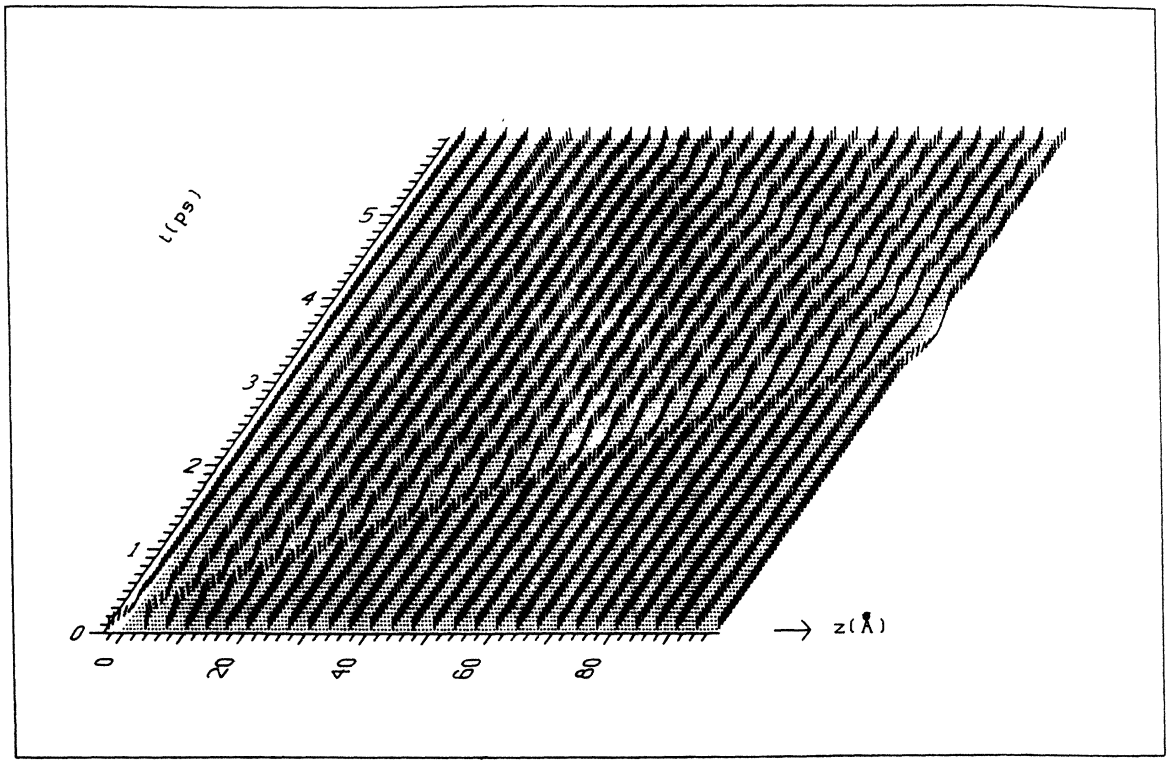

FIG. 5. Propagation of a solitary wave as a function of all geometrical variables and of time.

localized character of the core and σ electrons, one can assume[8,9] that interactions between the molecules are, first of all, due to the localized π electrons. Therefore it seems to be justified to separate the π electrons from the σ system. While the influence of the core and σ electrons can be described with a classical potential, the π electrons must be treated explicitly with a tight-binding-type model. The electron-electron interactions are included only implicitly via the parametrization.

Consequently, the Hamiltonian is divided into six different parts:

$$\hat{H}=\hat{H}_{\text{kin}}+\hat{H}_0+\hat{H}_{\text{el}}+\hat{H}_{\text{el-conf}}+\hat{H}_{\text{conf}}+\hat{H}_{\text{bb}} \ . \tag{1}$$

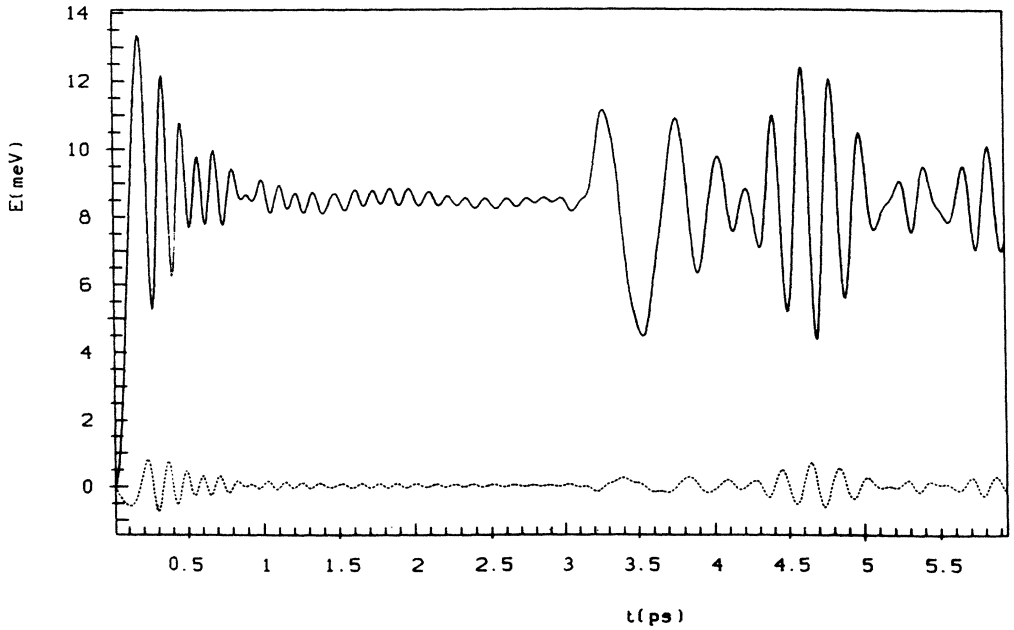

FIG. 6. Development of the kinetic (upper curve) and total (lower curve) energy in time for the propagation of a solitary wave.

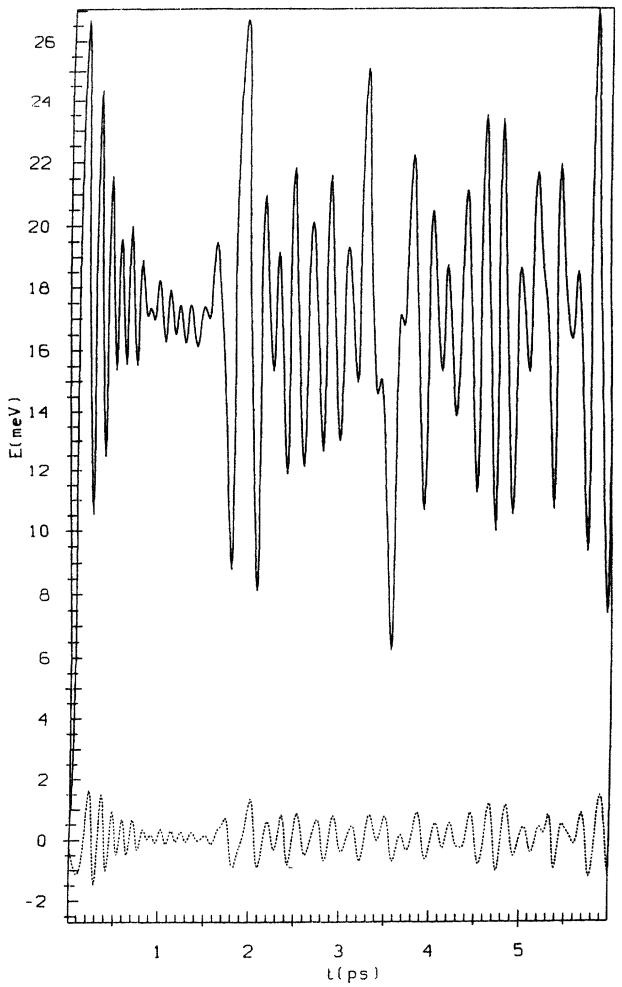

FIG. 7. Development of the kinetic (upper curve) and total (lower curve) energy in time for the collision of two solitary waves.

$\hat{H}_{\text{kin}}$ gives the kinetic energy of the system,

$$\hat{H}_{\text{kin}}=\tfrac{1}{2}\sum_{n=1}^{N}(M\dot{z}_n^2+\theta_\vartheta\dot{\vartheta}_n^2+\theta_\varphi\dot{\varphi}_n^2)\ . \tag{2}$$

E_{kin} is the kinetic energy due to $\hat{H}_{\text{kin}}$. It is given by

$$E_{\text{kin}}=\langle\psi\,|\,\hat{H}_{\text{kin}}\,|\,\psi\rangle\ . \tag{3}$$

The same is valid for E_{conf}, and so on. The index for the site is n. N gives the number of molecules in a stack. The geometry variables z, φ, and ϑ are defined in the same way as by Ladik and Čížek[8,9] (see Fig. 2). Figure 2 shows a stacked formamide dimer. The variables z_n, φ_n, and ϑ_n are defined as the deviations from their equilibrium values $(n-1)z_0$, $(n-1)\varphi_0$, and ϑ_0. These are given as in B-DNA by $z_0=3.36$ Å, $\varphi_0=36°$, and $\vartheta_0=0$. The helix axis is denoted by z. In Fig. 2, molecule 1 is in its equilibrium position, while for molecule 2, a deviation of z_2, φ_2, and ϑ_2 is sketched. Figure 3 shows a projection of a stacked dimer in its equilibrium geometry onto the plane perpendicular to the z axis. The bond lengths are given in angstroms, and all bond angles are assumed to be 120°. M gives the total mass of a molecule. θ_φ and θ_ϑ, respectively, are the inertial momenta for the two different rotational axes (for φ, the rotational axis is the main axis z of the formamide helix, and for ϑ, it is the y axis; see Fig. 2). The matrix elements formed with the help of atomic orbitals of a basis set, $\langle\chi_r\,|\,\hat{H}_0\,|\,\chi_s\rangle$, create a block-diagonal matrix. Each block describes a single molecule (only π electrons are considered explicity). For N single molecules, $\hat{H}_0$ is defined as

$$\hat{H}_0=\sum_{g=1}^{G}\sum_{n=1}^{N}\alpha_g\hat{C}^{\dagger}_{g,n}\hat{C}_{g,n}+\sum_{g,h=1}^{G}\beta_{gh}\sum_{n=1}^{N}(\hat{C}^{\dagger}_{g,n}\hat{C}_{h,n}+\hat{C}^{\dagger}_{h,n}\hat{C}_{g,n})\ . \tag{4}$$

$\hat{C}^{\dagger}_{g,n}$ and $\hat{C}_{g,n}$ are creation and annihilation operators for π electrons at atom g in molecule n. α_g stands for the Coulomb integral of atom g. The β's are the different resonance integrals between atoms g and h in one molecule. The values of α_g and β_{gh} are chosen to be equal to the corresponding Fock matrix elements of an *ab initio* calculation of a formamide molecule using a minimal STO-3G (Slater-type orbitals contracted from three Gaussian functions) basis set. All Hartree-Fock calculations have been performed with the program package GAUSSIAN-74.[18] The exponents and contraction coefficients of the applied atomic basis functions have been taken from the basis-set library of this program package. The indices g and h number the π centers in the molecule. G is the total number of π centers in it. The electron-electron interaction is not explicitly considered for this first calculation. The field-operator formulation of quantum mechanics was used by Su, Schrieffer, and Heeger[12,13] in their work on topological solitons in polyacetylene, in which they had to introduce only one geometrical variable. In order to facilitate comparison between these two different kinds of solitons, the field-operator formulation was also adapted by Ladik and Čížek[8,9] in their original theory of conformational solitons in DNA (in which they had to use, instead of one, three geometrical variables). However, one should keep in mind that the wave-function formulation provides a completely equivalent description of the problem.

$\hat{H}_{\text{el}}$ describes the interaction of the π electrons along the stack including only first neighbors,

$$\hat{H}_{\text{el}}=\sum_{g,h=1}^{G}t_{0,0,0,g,h}\sum_{n=1}^{N-1}\left[\hat{C}^{\dagger}_{n+1,g}\hat{C}_{n,h}+\hat{C}^{\dagger}_{n,h}\hat{C}_{n+1,g}\right]\ . \tag{5}$$

$t_{0,0,0,g,h}$ is the resonance integral of the undistorted stack between atoms g and h in neighboring molecules. They have been determined together with the electron-phonon coupling constants.

The term $\hat{H}_{\text{el-conf}}$ represents the electron-phonon coupling. The coupling constants are expanded into a Taylor series. Δz_n is defined as

$$\Delta z_n=z_{n+1}-z_n\ , \tag{6}$$

and analogously for $\Delta\vartheta_n$ and $\Delta\varphi_n$. Then, one can write $\hat{H}_{\text{el-conf}}$ as follows:

$$\hat{H}_{\text{el-conf}}=\sum_{i+j+k>0}^{i+j+k\le\rho}\sum_{g,h=1}^{G} t_{i,j,k,g,h}\sum_{n=1}^{N-1}\Delta z_n^i\Delta\vartheta_n^j\Delta\varphi_n^k(\hat{C}^{\dagger}_{n+1,g}\hat{C}_{n,h}+\hat{C}^{\dagger}_{n,h}\hat{C}_{n+1,g}) . \tag{7}$$

We have performed calculations on formamide dimers (see Fig. 2) treating the two molecules together as a supermolecule using the above-mentioned STO-3G minimal basis set for different sets of coordinates z, ϑ, and φ. These calculations give the matrix elements of $\hat{H}_{\text{el-conf}}$ which are equal to the corresponding Fock matrix elements between the $2p_z\pi$ atomic basis functions. To obtain these matrix elements as an analytical function of the geometrical variables, they have been expanded into a Taylor series. The coefficients of this series, which are the electron-phonon coupling constants $t_{i,j,k,g,h}$, have been determined by a least-squares fit. In Fig. 4, the deviation of this Taylor series from the exact value is displayed as a function of z for the resonance integral between the two nitrogen atoms of the dimer for different orders of the Taylor series ($\rho=4,5,6$). Obviously, the deviations in the sixth-order curve are negligible. The points in which Hartree-Fock calculations have been performed are chosen randomly. Comparative calculations on point sets of different sizes (200, 400, 600, and 800 points) have shown that a 200-point set is not dense enough to allow a reliable fit, while 400 points are enough for this purpose.[19] For more details see below.

The operator $\hat{H}_{\text{conf}}$ contains effects of the σ-electron system and the nuclear repulsion:

$$\hat{H}_{\text{conf}}=\sum_{n=1}^{N-1}\sum_{i,j+k\ge 0}^{i+j+k\le\rho} K_{i,j,k}\Delta z_n^i\Delta\vartheta_n^j\Delta\varphi_n^k . \tag{8}$$

One obtains the coefficients $K_{i,j,k}$ in the following way: At first, a potential surface scan (STO-3G basis) is performed for the dimer. Then, from the hypersurface, the π-electron energy ($E_0+E_{\text{el}}+E_{\text{el-conf}}$) is subtracted. Last, the coefficients are fitted to this modified hypersurface with the least-squares method, as described above for the coupling constants.

The numerical results show that formamide stacks are not stable in the chosen conformation on the *ab initio* level. In DNA, the stacked nucleotide bases are bound together covalently by sugar-phosphate groups called the backbone (BB). Therefore it seems to be reasonable to add a term to the potential which simulates such a backbone. In the equilibrium geometry, the gradients of the total potential energy of a formamide stack turned out to vanish, with the exception of the ones at the two terminal molecules. These two sets of nonvanishing gradients are equal in absolute value but of opposite sign. Therefore an additional term $\hat{H}_{\text{BB}}$, leading to an energy contribution E_{BB}, can be designed, which shifts the two nonzero sets of gradients in the equilibrium geometry ($z_n=\vartheta_n=\varphi_n=0$, symbolized by eq) to zero. For this purpose, a linear ansatz for E_{BB} can be used:

$$E_{\text{BB}}=A_z\sum_{n=1}^{N-1}\Delta z_n+A_\vartheta\sum_{n=1}^{N-1}\Delta\vartheta_n+A_\varphi\sum_{n=1}^{N-1}\Delta\varphi_n$$

$$=A_z(z_N-z_1)+A_\vartheta(\vartheta_N-\vartheta_1)+A_\varphi(\varphi_N-\varphi_1) . \tag{9}$$

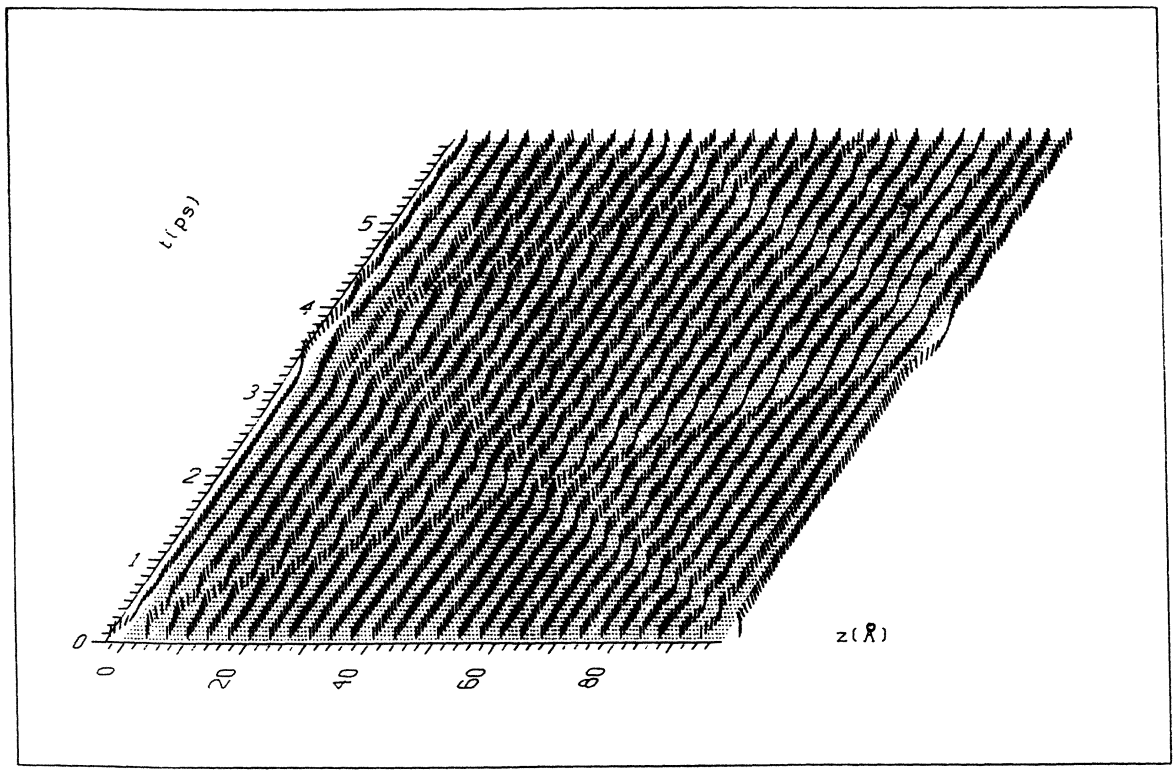

FIG. 8. Passing of two solitary waves as a function of all geometrical variables and of time.

The requirement that the equilibrium geometry represent the minimum of the total potential energy E_p leads to

$$\left.\frac{\partial E_p}{\partial z_1}\right|_{\text{eq}} = \left.\frac{\partial}{\partial z_1}(E_{\text{BB}}+E_{\text{conf}}+E_{\text{el-conf}})\right|_{\text{eq}} = 0 \ , \tag{10a}$$

and finally, together with (9), to

$$A_z = \left.\frac{\partial}{\partial z_1}(E_{\text{conf}}+E_{\text{el-conf}})\right|_{\text{eq}} \ . \tag{10b}$$

The potential-energy terms not appearing in (10a) are geometry independent. Since the derivative of $E_{\text{conf}}+E_{\text{el-conf}}$ with respect to z_n has the same value as that with respect to z_1, but with opposite sign, (10b) determines A_z uniquely. For A_ϑ and A_φ, the same considerations apply. As shown below, (10b) is equivalent to

$$A_z = -K_{1,0,0} - 2\sum_{g,h=1}^{G} t_{1,0,0,g,h}P_{2,g,1,h} \ . \tag{11}$$

(For the definition of $P_{2,g,1,h}$, see below.)

To solve the dynamical problem, one needs the gradients for a geometry $\{z\}_i$, $\{\vartheta\}_i$, and $\{\varphi\}_i$ at a time z_i to calculate the new configuration $\{z\}_{i+1}$, $\{\vartheta\}_{i+1}$, and $\{\varphi\}_{i+1}$. These are computed analytically. Because the formulas for all three variables are analogous, only the derivative of the energy with respect to $\{z\}$ is given.

The backbone term is only for the molecules at the chain ends of interest [see Eq. (9)],

$$\frac{\partial}{\partial z_N}E_{\text{BB}} = -\frac{\partial}{\partial z_1}E_{\text{BB}} = A_z \ . \tag{12}$$

From the classical term $\hat{H}_{\text{conf}}$, one gets

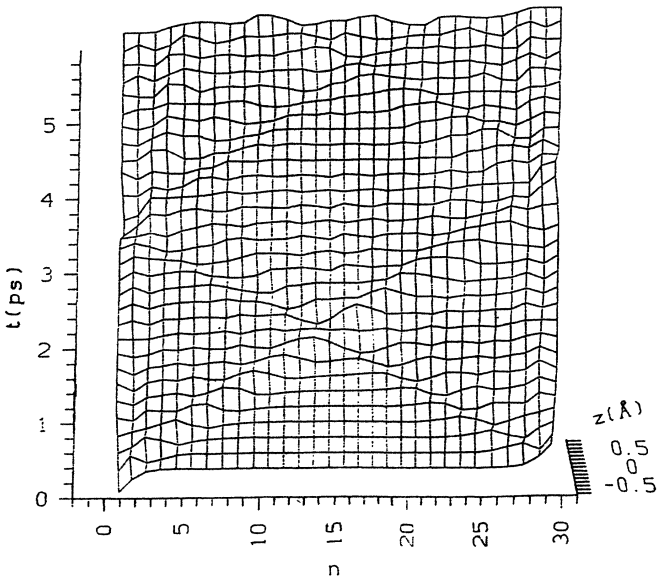

FIG. 9. Passing of two solitary waves as a function of z and time.

$$\frac{\partial}{\partial z_n}E_{\text{conf}} = \sum_{\substack{j,k\geq 0\\ i>0}}^{i+j+k\leq\rho} iK_{i,j,k}(\Delta z_{n-1}^{i-1}\Delta\vartheta_{n-1}^{j}\Delta\varphi_{n-1}^{k} - \Delta z_n^{i-1}\Delta\vartheta_n^{j}\Delta\varphi_n^{k}) \ . \tag{13}$$

The electronic part simplifies to

$$\frac{\partial}{\partial z_n}E_{\text{el-conf}} = 2\sum_{\substack{j,k\geq 0\\ i>0}}^{i+j+k\leq\rho}\sum_{g,h=1}^{G} it_{i,j,k,g,h}(P_{n,g,n-1,h}\Delta z_{n-1}^{i-1}\Delta\vartheta_{n-1}^{j}\Delta\varphi_{n-1}^{k} - P_{n+1,g,n,h}\Delta z_n^{i-1}\Delta\vartheta_n^{j}\Delta\varphi_n^{k}) \ . \tag{14}$$

Here the elements of the density matrix $\underline{P}$ are defined by

$$P_{r,s} = 2\sum_{i=1}^{n^*} c_{i,r}c_{i,s} \tag{15}$$

(n^* denotes the number of occupied orbitals), where $c_{i,r}$ is the rth coefficient of the ith eigenvector. This can be verified by using the fact that

$$\frac{\partial}{\partial z_n}E_0 = \frac{\partial}{\partial z_n}E_{\text{el}} = 0 \tag{16}$$

and (see, e.g., Ref. 19) that

$$\frac{\partial}{\partial z_n}\langle\psi|\hat{H}_0+\hat{H}_{\text{el}}+\hat{H}_{\text{el-conf}}|\psi\rangle = \frac{\partial}{\partial z_n}\sum_{r,s}(P_{r,s}H_{r,s}) = \sum_{r,s}P_{r,s}\frac{\partial}{\partial z_n}H_{r,s} \tag{17}$$

is valid (ψ is the wave function).

Assuming the determined gradients over a timestep τ_0 to be constant, one can integrate the classical equations of motion for each molecule at a time τ_i,

$$F_n(i) = M\ddot{z}_n(i) \ , \tag{18}$$

$$\dot{z}_n(i+1) = \dot{z}_n(i) + \tau_0 F_n(i)/M \ . \tag{19}$$

For the determination of the new coordinates, the formula of Su and Schrieffer was used,[13]

$$z_n(i+1) = z_n(i) + \tau_0\dot{z}_n(i+1) \ . \tag{20}$$

As already said, a formamide stack was chosen as a model for a nucleotide base stack to avoid its complexity. Formamide can be considered as a section of a cytosine molecule [N(3)-C(2)-O(7)]. The structure of the stack should imitate the structure of B-DNA.[20] Therefore the equilibrium stacking distance is 3.36 Å, $\varphi_0 = 36°$, and the plane of the molecule is perpendicular to the main axis. The sugar-phosphate chain is simulated by $\hat{H}_{\text{BB}}$. Formamide has three π centers ($G=3$). The mass of formamide is $M = 45.0$ g/mol. $\theta_\vartheta = 115.6$ g Å^2 and

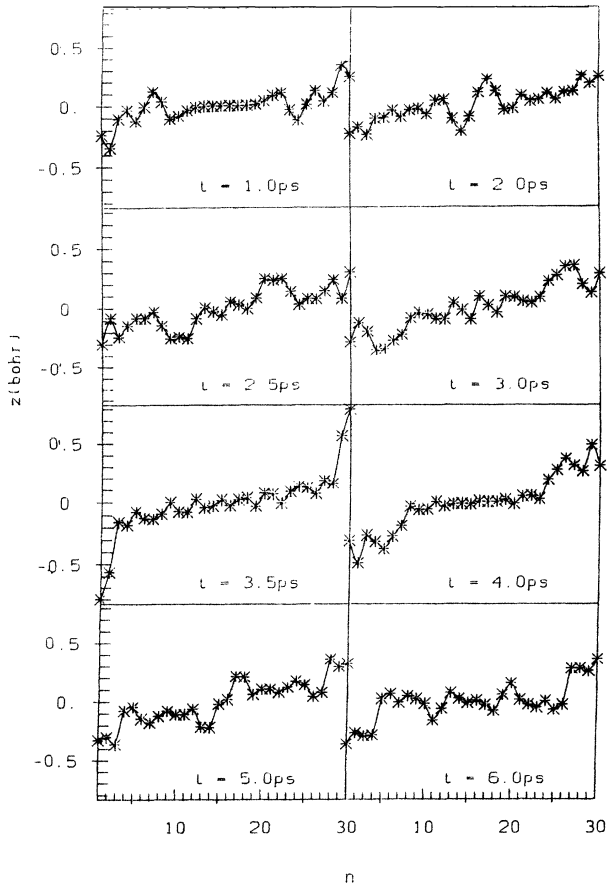

FIG. 10. Passing of two solitary waves as a function of z (τ gives the time in ps).

$\theta_\varphi = 162.4$ g Å^2 results for the inertial momenta by the given choice of the axis (see Figs. 1 and 2). Thirty molecules have been chosen for the simulation ($N = 30$). The values of the Coulomb and resonance integrals are $\alpha_C = -0.139$ hartrees, $\alpha_N = -0.340$ hartrees, $\alpha_O = -0.184$ hartrees, $\beta_{CN} = -0.195$ hartrees, and $\beta_{CO} = -0.362$ hartrees.

Four hundred points of the potential hypersurface of the dimer were computed with the help of GAUSSIAN-74 to get the hopping integrals and electron-phonon coupling constants. There were 340 points created randomly between the limits $\Delta z = \pm 0.5$ Å, $\Delta\vartheta = \pm 15°$, and $\Delta\varphi = \pm 15°$. The other 60 points were necessary to control the fit (see, e.g., Fig. 3). As one can see in Fig. 3, there is agreement between the exact curve and the sixth-order fit. Therefore $\rho = 6$ was used.

Table I shows, for all hopping integrals, the range and the greatest error. One can see that the error compared to the range is very small in the sixth order. The range is rather large due to the large changes in the geometry within the set of 400 points. Only the first ten coupling constants for the five most important interactions are given as examples because of the large number of them (756) in Table II.

Knowing these constants, we can determine the classical constants K. At first, the energy of interaction of the σ-electron system was calculated (E_{HF} represents the Hartree-Fock energy):

$$E_\sigma = E_{HF} - E_\pi = E_{HF} - E_0 - E_{el} - E_{el\text{-}conf} \,. \tag{21}$$

The σ energy was expanded in a Taylor series up to sixth order. The range for the σ energy was [−329.736,

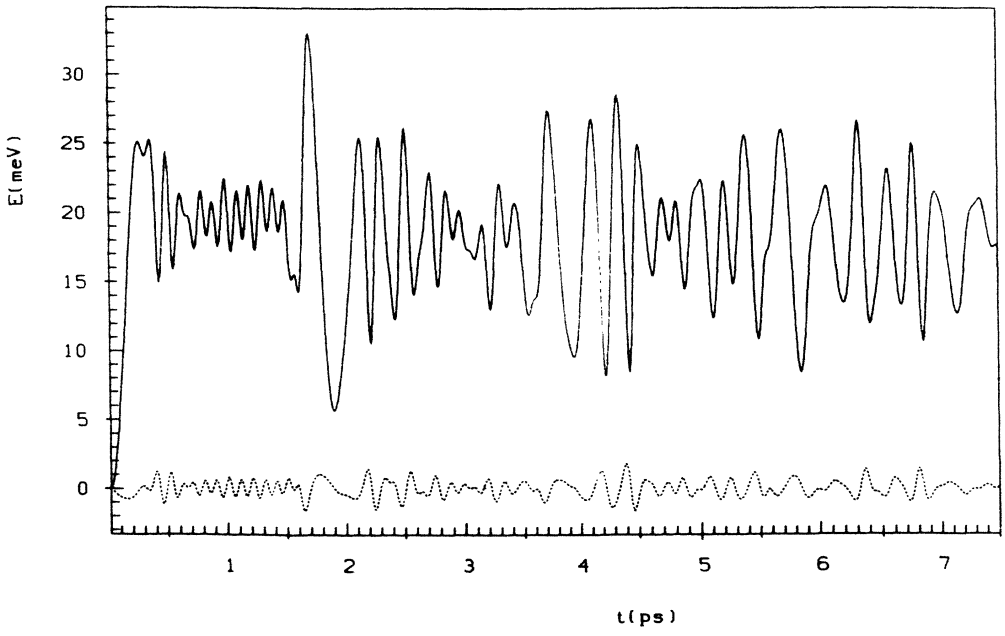

FIG. 11. Development of the kinetic (upper curve) and total (lower curve) energy in time for simultaneous changes of the starting values of all three geometrical variables.

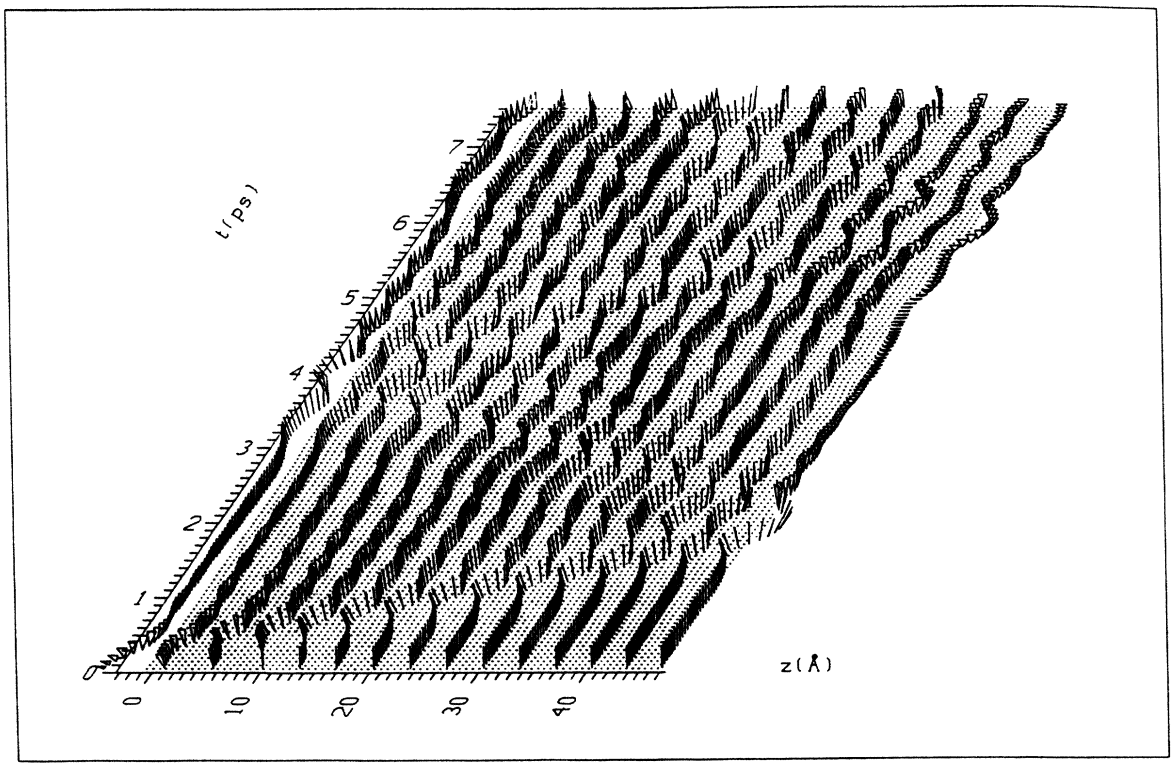

FIG. 12. Reflection of a solitary wave as a function of all geometrical variables and of time.

$-329.758]$ hartrees. The greatest error was 0.0095×10^{-3} hartrees. Table III gives the first ten classical constants. It is now easy to show that $A_z=1.31\times10^{-3}$ hartrees/bohr, $A_\vartheta=3.45\times10^{-3}$ hartrees, and $A_\varphi=2.00\times10^{-3}$ hartrees.

RESULTS AND DISCUSSION

Three computations were carried out. The first shows the propagation of a solitary wave, the second the crossing of two different waves, and the last one the reflection of a solitary wave from the chain end. For the first calculation in the starting configuration, all variables were set to zero except $z_1=-0.5$ Å.

In Fig. 5, all three geometrical variables $\{\Delta z\}$, $\{\Delta\vartheta\}$, and $\{\Delta\varphi\}$ at the time are shown. The ordinate gives the time and the abscissa $\{z\}$ and $\{\Delta z\}$. Each molecule is symbolized by a triangle. The backbone corresponds to the dashed line. The rotation axes are displayed in a way that the ϑ axis would be perpendicular to the paper plane and the φ axis parallel to the z axis (not true for the real stack). To allow us to distinguish between positive and negative φ, the triangle is rotated in the latter case about 180°. In the first case, the triangle stands, consequently, on the bottom; in the latter, on the top. The variables were multiplied by 6 for a better visualization.

The solitary wave passes, in approximately 3.6 ps, the whole stack (98 Å). This corresponds to a velocity of 2.7 km/s, where only small dispersion occurs. The mean value of the kinetic energy is 8.4 meV. Together with the velocity, this leads to an upper limit of the kinetic mass of roughly 400 electron masses (m_e).

The conservation laws are examined to exclude errors in the numerical integration. The momentum and angular momentum is, over the whole time, zero in the accuracy of the computer. The fluctuations in the total energy compared to the total kinetic energy (see Fig. 6) appear to be small.

In the second simulation, a collision between two soli-

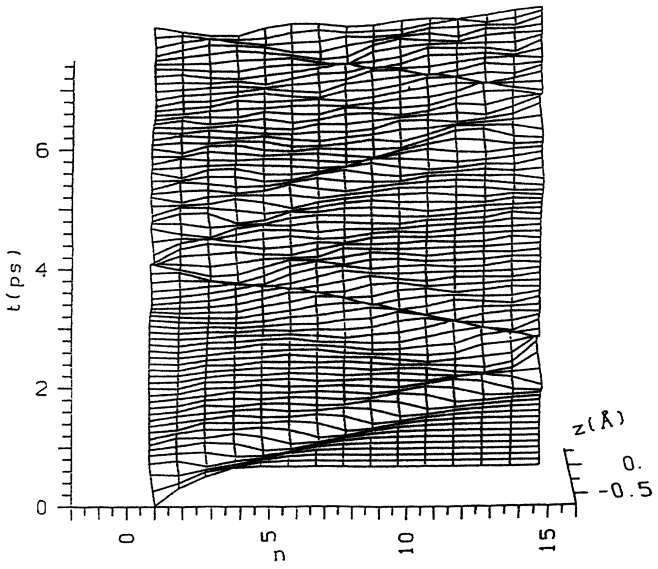

FIG. 13. Reflection of a solitary wave as a function of z and time.

tary waves starting from different chain ends is studied. This is obtained with an initial excitation of

$$z_1 = -0.5 \text{ Å} , \quad z_{30} = 0.5 \text{ Å} . \tag{22}$$

Figure 7 shows the time development of the total and kinetic energy after this excitation. As in the former case, the total energy fluctuations remain an order of magnitude smaller than the kinetic energy. The mean value of the kinetic energy, averaged over 6 ps, is 16.8 meV. The velocities of the two solitons are equal, and they pass each other without any perturbation.

By looking at Fig. 8, we see clearly that z contains the most information (ϑ about the same, φ not recognizable). This justifies Fig. 9, where z is drawn from up to down (in B), the index n of the molecule from left to right, and the time (in ps) from front to back. The passing of the solitary waves can be observed easily. This figure suggests that some other vibrations had been excited besides the two solitary waves from the starting configuration, especially after reflections at the chain ends. Figure 10 shows this also. In this drawing, z is given as a function of n. The time t is given in ps. The width of the solitary waves can be estimated to six sites.

In the last calculation, the solitary wave emitted by a more complex excitation has been studied. To save computation time, the length of the stack had been limited to 15 units. The first three molecules were gradually excited in all three variables ($z = -0.8$, -0.4, and -0.2 Å, and $\vartheta, \varphi = -12°$, $-6°$, and $-3°$). For the time step, 0.01 ps has been chosen. The constancy of the energy is not worse than before, but the kinetic energy is larger than that seen previously (Fig. 11).

The high kinetic energy can also be seen in Fig. 12. Many vibrations are visible, but the three-dimensional projection (Fig. 13) makes the reflection very clearly recognizable.

CONCLUSION

It has been shown numerically that solitary waves exist in polyformamide with the described formalism. This has been proven, because only small dispersion occurs during the propagation of the wave. Furthermore, these waves are reflected and can pass through each other. The geometry imitates the DNA B. Therefore similar long-range effects are to be expected for DNA.

Investigations of the effective mass of the solitary wave and its stability against impurities, as well as the introduction of electron-electron interactions in the PPP approximation, are in progress in our laboratory.

[1] E. Boyland, in Proceedings of the Israel Academy of Sciences, Symposium of Carcinogenesis, Jerusalem, 1968 (unpublished).

[2] J. Ladik, Int. J. Quant. Chem. Quant. Biol. Symp. **13**, 307 (1986).

[3] J. Ladik, S. Suhai, and M. Seel, Int. J. Quant. Chem. Quant. Biol. Symp. **5**, 35 (1978).

[4] A. K. Bakhshi, J. Ladik, M. Seel, and P. Otto, Chem. Phys. **108**, 233 (1986).

[5] K. Laki and J. Ladik, Int. J. Quant. Chem. **20**, 683 (1976).

[6] F. Beleznay, S. Suhai, and J. Ladik, Int. J. Quant. Chem. **20**, 683 (1981).

[7] A. S. Davydov, Phys. Scr. **20**, 387 (1979).

[8] J. Ladik and J. Čížek, Int. J. Quant. Chem. **26**, 955 (1984).

[9] J. Ladik, *Molecular Basis of Cancer, Part A: Macromolecular Structure, Carcinogens and Oncogenes* (Publisher, City, 1985), p. 343ff.

[10] M. Collins, Adv. in Chem. Phys. **53**, 225 (1983).

[11] A. S. Davydov, Usp. Fiz. Nauk **138**, 603 (1982) [Sov. Phys.—Usp. **25**, 898 (1982)].

[12] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Phys. Rev. B **22**, 2099 (1980).

[13] W. P. Su and J. R. Schrieffer, Proc. Natl. Acad. Sci. U.S.A. **77**, 5626 (1980).

[14] S. Kivelson and D. E. Heim, Phys. Rev. B **26**, 4278 (1982).

[15] W. Förner, M. Seel, and J. Ladik, J. Chem. Phys. **84**, 5910 (1986).

[16] W. Förner, C. L. Wang, F. Martino, and J. Ladik, Phys. Rev. B (to be published).

[17] J. A. Krumhansl and D. M. Alexander, in *Structure and Dynamics: Nucleic Acids and Proteins*, edited by E. Clementi and R. H. Sarma (Adenine, New York, 1983), p. 61.

[18] J. A. Pople and W. Hehre, Quantum Chemistry Program Exchange, Program No. 236, 1974 (unpublished).

[19] W. Förner (unpublished).

[20] S. Arnott, S. D. Dover, and A. J. Wonacott, Acta Crystallogr. Sec. B **25**, 2192 (1969).

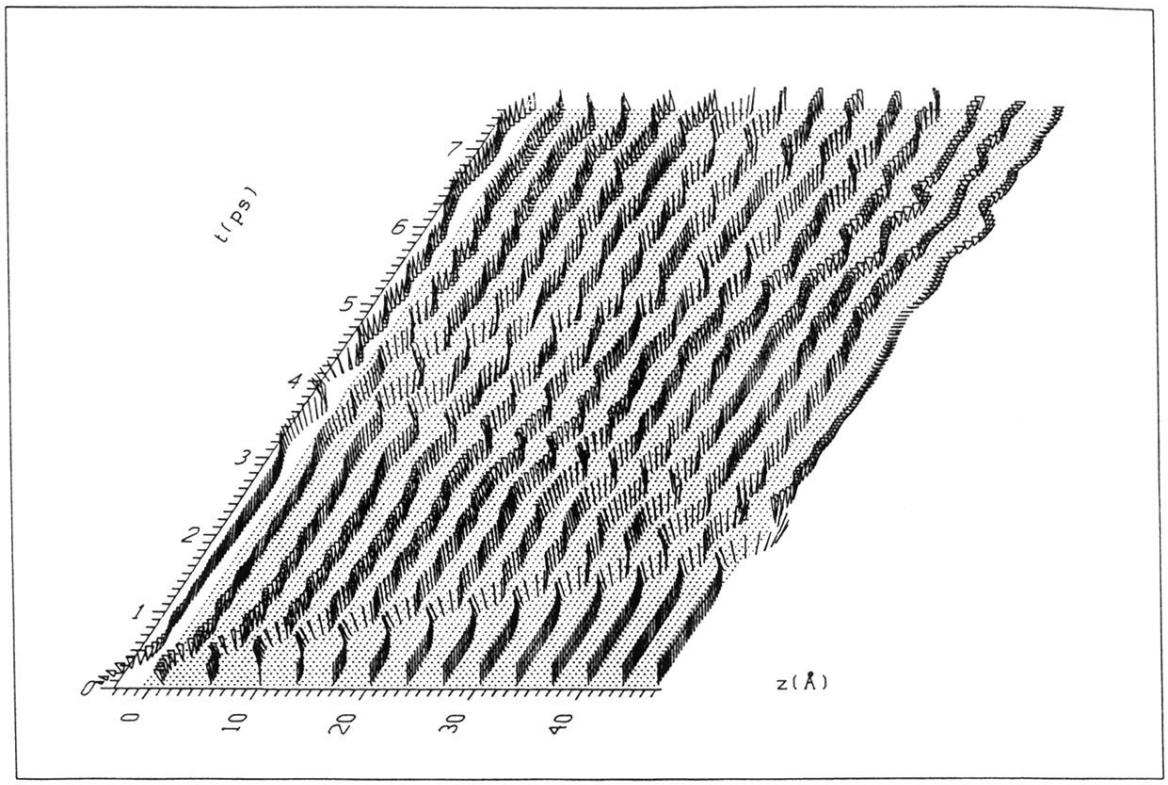

FIG. 12. Reflection of a solitary wave as a function of all geometrical variables and of time.

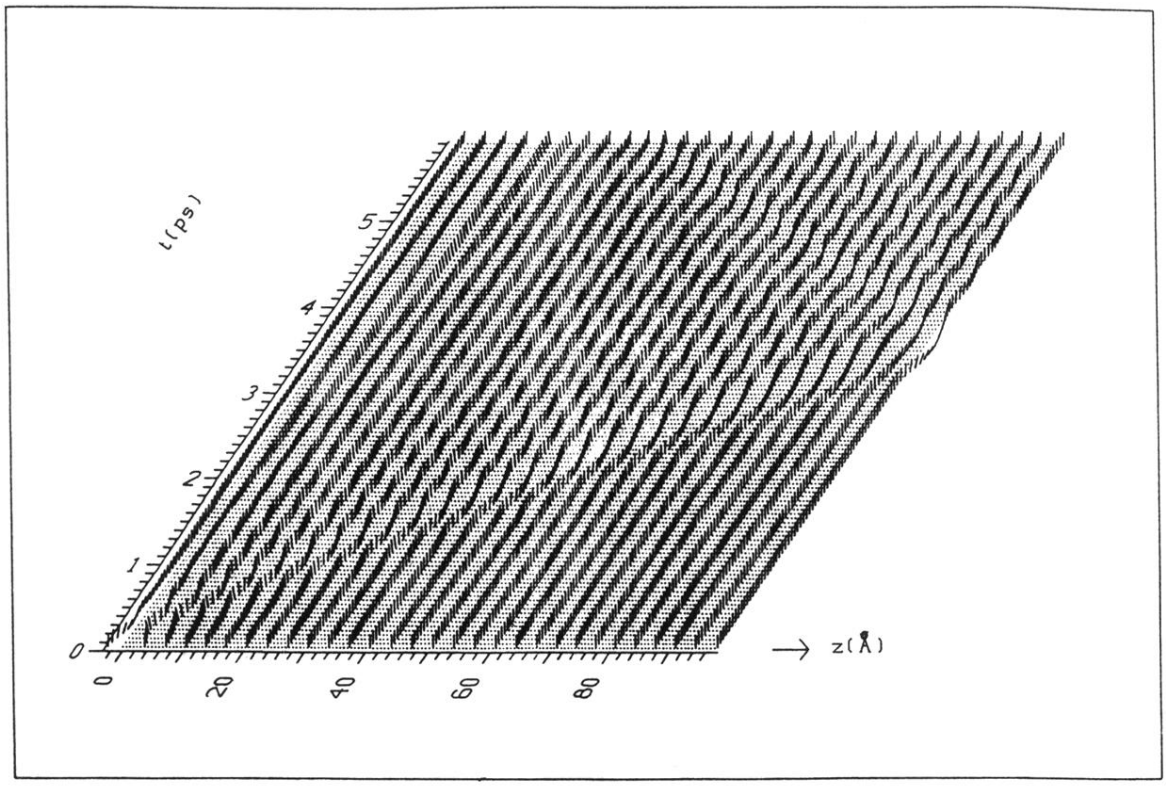

FIG. 5. Propagation of a solitary wave as a function of all geometrical variables and of time.

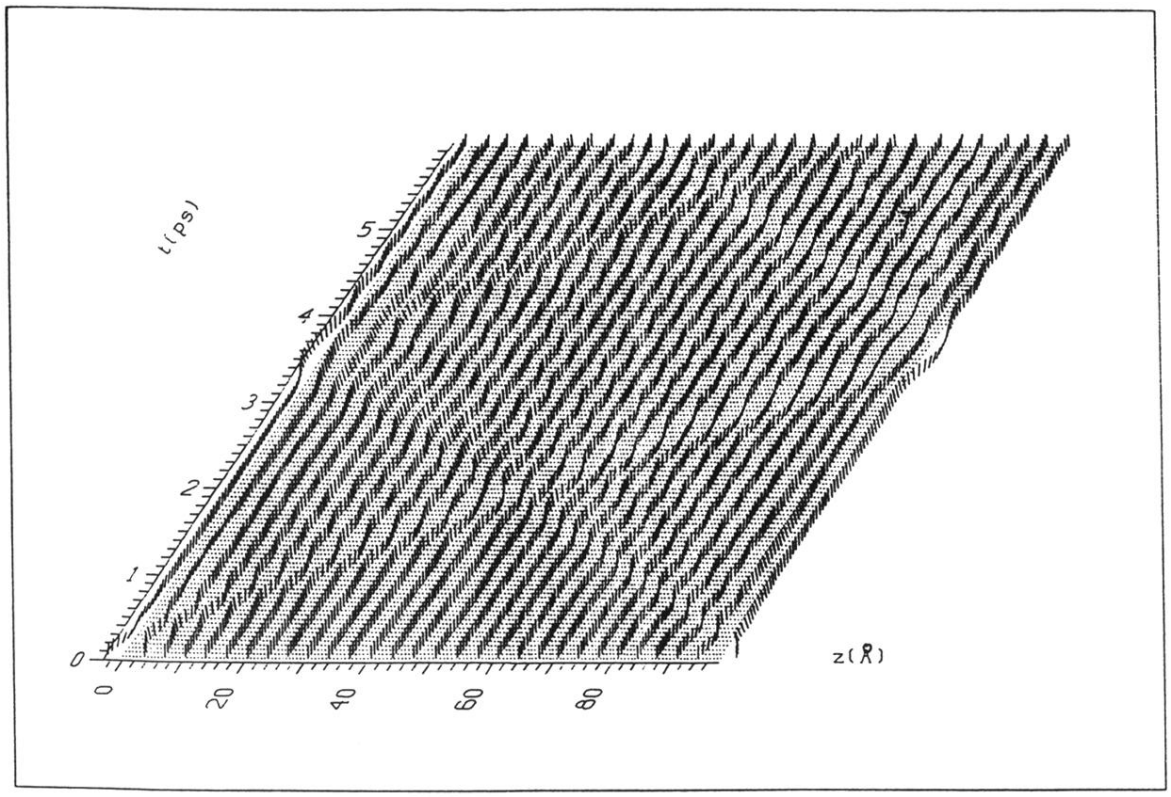

FIG. 8. Passing of two solitary waves as a function of all geometrical variables and of time.