Equations of motion for polymer chains in a thermostat

Aurel Bulgac\textsuperscript{1} and H Eduardo Roman\textsuperscript{2}

\textsuperscript{1} Department of Physics, University of Washington, PO Box 351560, Seattle, WA 98195, USA
\textsuperscript{2} Dipartimento di Fisica, Università di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy
E-mail: bulgac@phys.washington.edu and Eduardo.Roman@mib.infn.it

Abstract. A constrained dynamics method suitable for molecular dynamics simulations is considered to study the long-time dynamics of polymer chains. The method is initially discussed on the basis of the Lagrangian and Hamiltonian formalisms for isolated polymer chains with fixed monomer–monomer links. Subsequently, the corresponding equations of motion are obtained for describing the dynamics of such polymer chains in the presence of a thermostat. The approach is applied to a few typical cases to illustrate how the formalism is implemented numerically and to elucidate its convergence properties when studying such systems in equilibrium. As an example, we consider the problem of reconstructing the backbone structure (chain of C\textsubscript{\alpha} atoms) of protein Rubredoxin from its contact matrix. It is shown that the target structure is successfully reached after a long transient regime (typically in the range from $10^6$ to $10^8$ integration steps). A particular attractive extension of the algorithm presented here is to environment-dependent couplings, which could allow the study of the long-time polymer dynamics in realistic environments. The present method is thus expected to have useful applications in the modelling of the complex dynamics of bio-polymers such as proteins, and also in the context of nanoscale polymer materials.
1. General remarks

Numerical simulations of the dynamics of polymer chains have been addressed by studying transport equations either in the form of a Fokker–Planck equation (often referred to as Smoluchowski equation also) or of Langevin type [1], by means of Monte Carlo (MC) methods [2] or molecular dynamics (MD) techniques [1]–[8]. In all of the above cases, one of the major challenges is a suitable description of the links between subsequent monomers in a polymer chain. One approach is to introduce a relatively stiff potential between monomers [1, 2]. In actual simulations, it leads to rather small integration time steps and, as a result, one is forced to spend most of the time in simulating the polymer links rather than the relatively slow polymer chain dynamics. One obvious solution is to introduce holonomic constraints [4]–[6]; another is to define generalized coordinates [1, 7, 8] and thus deal only with the real degrees of freedom.

A further requirement in the case of polymers in solutions is to have a suitable description of the solvent and its interaction with the polymer. That can be done either by an explicit treatment of the solvent molecules [2, 4] (which requires at least an order of magnitude or more molecules than the number of monomers) or by the introduction of a phenomenological friction and fluctuating forces as in the Langevin equation. The introduction of constraints and a phenomenological description of the solvent molecules allows one to concentrate one’s attention on the most interesting feature, the polymer dynamics.

In this paper, we discuss the main theoretical aspects involved in the study of polymer chains (with either fixed bond lengths, or fixed bond lengths and fixed angles between successive links) coupled to a thermostat in the framework of a MD approach, in which hydrodynamic interactions are neglected. A very similar Lagrange and Hamiltonian dynamics for systems with holonomic constraints was developed independently in [9]. The Hamiltonian formalism developed here, however, is somewhat more transparent and the manner in which the holonomic constraints...
are introduced is more direct and in clear relationship with Dirac’s formalism [10]. The main difference between our approach and that of [9], for example, is in the way we couple the polymer chain to the thermostat. The coupling to the thermostat described here is more flexible in several aspects, it allows a controlled exchange of the energy flow between the system and the thermostat both in the time and in the spatial domains. In particular, it is very easy in our approach to allow the thermostat to couple to the system only at its surface, for example, which in many instances is clearly the manner by which the energy is exchanged between the system and the thermostat.

The paper is organized as follows: in section 2 we shall briefly review the Lagrangian dynamics of a polymer chain. In section 3 we discuss the Hamiltonian dynamics of a polymer chain, using the so-called Dirac–Poisson brackets [10], a method to treat holonomic and nonholonomic constraints, developed for the needs of quantum field theory. In section 4 we show how one can couple a polymer chain to a thermostat and in this way have a phenomenological description of the solvent and of the polymer–solvent interaction, which obviates an explicit treatment of the solvent molecules with very little effort. We shall briefly exemplify our methods with numerical simulations of the polymer dynamics in section 5 and present our main conclusions in section 6. A number of rather technical aspects of our approach are relegated to appendices A–E.

2. Lagrangian formalism

Let us assume that a polymer chain with \( N + 1 \) monomers with coordinates \( r_k \), where \( k = 0, \ldots, N \), is described by the Lagrangian \( \mathcal{L} \)

\[
\mathcal{L} = \sum_{k=0}^{N} \frac{m \dot{r}_k^2}{2} - U, \tag{1}
\]

where \( U \) is the potential energy. (The formalism is identical for the case of unequal monomer masses.) If the links of the chain are rigid, one can derive the following Lagrange equations of motion for the chain

\[
m\ddot{r}_k = -\frac{\partial U}{\partial r_k} - \lambda_k (r_k - r_{k-1}) + \lambda_{k+1} (r_{k+1} - r_k) \tag{2}
\]

\[
m\ddot{r}_k = -\frac{\partial U}{\partial r_k} - \lambda_k \rho_k + \lambda_{k+1} \rho_{k+1}, \tag{3}
\]

where we have introduced the link vectors \( \rho_k \) and the constraints \( \lambda_k \) (for \( k = 1, \ldots, N \))

\[
\rho_k = r_k - r_{k-1}, \tag{4}
\]

\[
\lambda_k = \rho_k^2 - l^2 = 0. \tag{5}
\]

In the above equations, \( \lambda_k \)'s are Lagrange multipliers, corresponding to the fixed bond lengths.

By using the second time derivative of the constraints

\[
\rho_k^2 + \ddot{\rho}_k \cdot \dot{\rho}_k \equiv 0, \tag{6}
\]
after some straightforward manipulations of the equations of motion, one can derive the following set of equations for the Lagrange multipliers $\lambda_k$:

$$-\rho_{k+1} \cdot \rho_k \lambda_{k+1} - \rho_{k-1} \cdot \rho_k \lambda_{k-1} + 2\dot{p}_k^2 \lambda_k = m \ddot{r}_k - \rho_k \cdot \left( \frac{\partial U}{\partial \dot{r}_k} - \frac{\partial U}{\partial \dot{r}_{k-1}} \right).$$  

(7)

As a last remark, once a chain has been specified, the initial velocities of the monomers cannot be specified in an arbitrary way, but have to be consistent with the constraints, namely

$$\dot{r}_k \cdot \rho_k = (r_k - \dot{r}_{k-1}) \cdot (r_k - \dot{r}_{k-1}) \equiv 0,$n

for $k = 1, \ldots, N$. In this way, the length of each particular link will be conserved and also the relative velocity of two consecutive monomers will always be orthogonal to the instantaneous link vector $\rho_k = r_k - r_{k-1}$.

### 3. Hamiltonian formalism

In order to derive the Hamiltonian equations of motion for a polymer chain with fixed links, we shall use Dirac’s formalism [10]. The reader who is not interested in the formalism and the derivation of the equations of motion can simply skip the whole section and refer only to the last equations in this section, equations (33) and (34).

According to Dirac’s formalism, besides the so-called primary constraints $\mathcal{X}_k = 0$, (see (6)), one has to introduce also the so-called secondary constraints $\mathcal{Y}_k = 0$ (for $k = 1, \ldots, N$)

$$\mathcal{Y}_k = \{\mathcal{X}_k, \mathcal{H}\} = \frac{2}{m} \rho_k \cdot \pi_k,$$  

(9)

$$\pi_k = p_k - p_{k-1},$$  

(10)

where $p_k$ are canonical conjugate momenta to coordinates $r_k$ with respect to the usual Poisson bracket between two arbitrary functions $A$ and $B$ on the phase space $r_k, p_k$

$$\{A, B\} = \sum_{k=0}^{N} \left( \frac{\partial A}{\partial r_k} \cdot \frac{\partial B}{\partial p_k} - \frac{\partial A}{\partial p_k} \cdot \frac{\partial B}{\partial r_k} \right).$$

(11)

Following [10], let us gather together both primary and secondary constraints and denote them as $\mathcal{Z}_\mu$, with $\mu = 1, \ldots, 2N$, where the first $N$ quantities are the primary constraints $\mathcal{X}_k$ and the following $N$ quantities are the secondary constraints $\mathcal{Y}_k$. The Hamiltonian dynamics in the presence of constraints is now completely defined by the so-called Dirac–Poisson brackets

$$\{A, B\}_D = \{A, B\} - \sum_{\mu, \nu=1}^{2N} \{A, \mathcal{Z}_\mu\} \mathcal{D}_{\mu\nu}\{\mathcal{Z}_\nu, B\},$$  

(12)

$$\mathcal{C}_{\mu\nu} = \{\mathcal{Z}_\mu, \mathcal{Z}_\nu\}, \quad \sum_{\gamma=1}^{2N} \mathcal{C}_{\mu\gamma} \mathcal{D}_{\gamma\nu} = \delta_{\mu\nu}. $$

(13)
It is convenient to introduce the following antisymmetric $N \times N$ matrices:

\[ U_{kl} = \{X_k, Y_l\}, \quad (14) \]

\[ \sum_{m=1}^{N} U_{km} V_{ml} = \delta_{kl}, \quad (15) \]

\[ T_{kl} = \{Y_k, Y_l\}. \quad (16) \]

As we will show below, see equations (24)–(27), the matrices $U$ and $T$ have a rather simple structure. Then the $2N \times 2N$ matrices $C$ and $D$ can be easily shown to have the following structure:

\[ C = \begin{pmatrix} 0 & U \\ -\tilde{U} & T \end{pmatrix}, \quad (17) \]

\[ D = \begin{pmatrix} \tilde{V} T \gamma & -\tilde{V} \\ \gamma & 0 \end{pmatrix}, \quad (18) \]

where $\tilde{U}$ stands for the transpose of $U$. The Dirac–Poisson bracket has the property that for an arbitrary Hamiltonian

\[ \{Z_\mu, \mathcal{H}\}_D \equiv 0. \quad (19) \]

Since the dynamical evolution of any quantity $A$ is now defined as [10]

\[ \frac{dA}{dt} = \frac{\partial A}{\partial t} + \{A, \mathcal{H}\}_D = \frac{\partial A}{\partial t} + \{A, \mathcal{H}\} - \sum_{\mu,\nu=1}^{2N} \{A, Z_\mu\} D_{\mu,\nu} \{Z_\nu, \mathcal{H}\} \]

\[ = \frac{\partial A}{\partial t} + \{A, \mathcal{H}\} - \sum_{\mu=1}^{2N} \{A, Z_\mu\} e_\mu, \quad (20) \]

\[ = \frac{\partial A}{\partial t} + \{A, \mathcal{H}\} - \sum_{\mu=1}^{2N} \{A, Z_\mu\} e_\mu, \quad (21) \]

where all the primary and secondary constraints are automatically satisfied along the trajectory of the system, if they were satisfied initially. In other words, both primary and secondary constraints are integrals of motion of the Dirac–Poisson dynamics. Above we have introduced the quantities $e_\mu$ as

\[ e_\mu = \sum_{\nu=1}^{2N} D_{\mu,\nu} \{Z_\nu, \mathcal{H}\}. \quad (22) \]

The set of quantities $e_\mu$ is thus entirely defined in terms of the phase space variables $r_k, p_k$ and the Hamiltonian of the system. Alternatively, the quantities $e_\mu$ can be determined as the solution of the following set of linear equations:

\[ \sum_{\mu=1}^{2N} C_{\nu\mu} e_\mu = \sum_{\mu=1}^{2N} \{Z_\nu, Z_\mu\} e_\mu = \{Z_\nu, \mathcal{H}\}, \quad (23) \]
which follows from equations (13) and (22). It is straightforward to show that only the following Poisson brackets among the constraints are nonvanishing:

\[
\{X_k, Y_k\} = \frac{8}{m} X_k + \frac{8}{m} l^2 = \frac{8}{m} \rho_k^2, \tag{24}
\]

\[
\{X_k, Y_{k+1}\} = -\frac{4}{m} \rho_k \cdot \rho_{k+1}, \tag{25}
\]

\[
\{X_k, Y_{k-1}\} = -\frac{4}{m} \rho_k \cdot \rho_{k-1}, \tag{26}
\]

\[
\{Y_k, Y_{k+1}\} = \frac{4}{m} (\pi_{k+1} \cdot \rho_k - \pi_k \cdot \rho_{k+1}), \tag{27}
\]

and as a result the previous linear system of equations for \(e_{\mu}\) has essentially a tridiagonal form. After defining \(a_k = e_k\) and \(b_k = e_{N+k}\), equations (23) can be rewritten as follows (with the convention that \(a_0 = a_{N+1} = b_0 = b_{N+1} \equiv 0\):

\[
-\rho_k \cdot \rho_{k-1} b_{k-1} + 2l^2 b_k - \rho_k \cdot \rho_{k+1} b_{k+1} = \frac{1}{2} Y_k, \tag{28}
\]

\[
-\rho_k \cdot \rho_{k-1} a_{k-1} + 2l^2 a_k - \rho_k \cdot \rho_{k+1} a_{k+1} = -\frac{m}{4} \{\{Y_k, H\} + \{b_{k-1} Y_{k-1} + b_{k+1} Y_{k+1}, Y_k\}\}, \tag{29}
\]

where

\[
\{Y_k, H\} = \frac{2}{m^2} \pi_k^2 - \frac{2}{m} \rho_k \left( \frac{\partial U}{\partial r_k} - \frac{\partial U}{\partial r_{k-1}} \right). \tag{30}
\]

In equation (29) one can introduce the quantities \(b_k\) under the Poisson bracket because \(Y_k = 0\) along the physical trajectory.

The equations of motion for the phase space variables \(r_k, p_k\) can be written as follows:

\[
\dot{r}_k = \frac{p_k}{m} - \sum_{\mu=1}^{2N} \{r_k, Z_{\mu}\} e_{\mu} = \frac{p_k}{m} - \frac{2}{m} (b_k \rho_k - b_{k+1} \rho_{k+1}), \tag{31}
\]

\[
\dot{p}_k = -\frac{\partial U}{\partial r_k} - \sum_{\mu=1}^{2N} \{p_k, Z_{\mu}\} e_{\mu} = -\frac{\partial U}{\partial r_k} + 2(a_k \rho_k - a_{k+1} \rho_{k+1}) + \frac{2}{m} (b_k \pi_k - b_{k+1} \pi_{k+1}). \tag{32}
\]

A further simplification of the equations of motion occurs if one remembers that along the physical trajectory \(Y_k = 0\). In such a case, \(b_k = 0\), see equation (33), and the Hamiltonian equations for the phase space variables \(r_k, p_k\) become

\[
\dot{r}_k = \frac{p_k}{m}, \tag{33}
\]

\[
\dot{p}_k = -\frac{\partial U}{\partial r_k} + 2(a_k \rho_k - a_{k+1} \rho_{k+1}) = -\frac{\partial U}{\partial r_k} + 2 \sum_{j=1}^{N} (\rho_k Y_{k,j} - \rho_{k+1} Y_{k+1,j}) \{Y_j, H\}. \tag{34}
\]
These are the Hamiltonian equations of motion we shall be using later. The quantities $\rho_k$ have been introduced in equation (4), $Y_l$ in equation (9) and $V_l$ in equations (14) and (15), respectively. It is not obvious that the system of linear equations (28) is always degenerate and the only solution is $b_k = 0$. In particular, for a perfectly straight polymer chain this system of linear equations has a nontrivial solution, even though the right-hand side is vanishing identically. We shall neglect here such refinements of the theory, which can however be incorporated into the formalism [10]. It is somewhat surprising to find out that along the physical trajectory one has the usual relationship between velocities and momenta, see equations (33), as in the absence of constraints. For example, it is easy to convince oneself that $r_k$ is not canonically conjugate to $p_k$ with respect to the Dirac–Poisson bracket (12), see appendix A. Because of this somewhat unexpected coincidence, the right hand side of equation (34), has exactly the same form as in the Lagrangian formalism.

4. Coupling the polymer chain to a thermostat

Once the Hamiltonian dynamics in the presence of constraints is constructed, the coupling of the polymer dynamics to a thermostat can be performed as described in [11], which represents a generalization of the improved Nosé–Hoover dynamics [12] developed in [13]. In [11] two different types of couplings were introduced, so-called projected and twisted. We shall use here the twisted coupling of the polymer to the thermostat, since it is much easier to implement and it is also extremely efficient.

One way of introducing the coupling to a thermostat can be described by adding to the Hamiltonian equations of motion couplings through the so-called pseudo-friction coefficients according to the following rule:

$$
- \frac{\partial U}{\partial r_k} \implies - \frac{\partial U}{\partial r_k} - \frac{\tau}{t_0 \sqrt{2N}} \zeta p_k, \quad (35)
$$

where $\zeta$ is a pseudo-friction coefficient, $\tau$ is a dimensionless constant of order one, characterizing the strength of the coupling to the thermostat, and $t_0$ is a time constant of the order of the shortest period of the polymer chain. The reason for having the factor $\sqrt{2N}$ in the above relations will be explained below. The form of this coupling can vary, for example, one can consider a different pseudo-friction coefficient for each cartesian direction and/or consider different other functions of $p_k$ and $\zeta$, e.g., $\tau \zeta^3 p_k^3$. Even though we shall use more complicated couplings in the simulations, see appendix C, in order not to clutter the paper with complicated relations, that can be derived in a straightforward manner and have been discussed previously [11, 13, 14], we shall exemplify the method in this section only with the above coupling. Rather often ergodicity is not achieved if only one pseudo-friction coefficient is introduced and for that reason a coupling with several pseudo-friction coefficients is preferable [11, 13, 14] and such a scheme will be used in actual calculations.

In the presence of a thermostat the equations of motion become

$$
\dot{r}_k = \frac{p_k}{m}, \quad (36)
$$
\[ \dot{p}_k = -\frac{\partial U}{\partial r_k} + 2 \sum_{l=1}^{N} (\rho_k V_{k,l} - \rho_{k+1} V_{k+1,l}) \{ \mathcal{V}_l, \mathcal{H} \} - \frac{\tau}{t_0 \sqrt{2N}} \zeta p_k \]

\[ = -\frac{4\tau \zeta}{mt_0 \sqrt{2N}} \sum_{l=1}^{N} (\rho_k V_{k,l} - \rho_{k+1} V_{k+1,l}) \rho_l \cdot \pi_l \]

\[ = -\frac{\partial U}{\partial r_k} + 2 (a_k \rho_k - a_{k+1} \rho_{k+1}) - \frac{\tau}{t_0 \sqrt{2N}} \zeta p_k, \quad (37) \]

\[ \dot{\zeta} = \frac{\tau}{t_0 \sqrt{2N}} \left( \sum_{k=0}^{N} \frac{p_k^2}{mT} - 2N \right). \quad (38) \]

Note that the pseudo-friction coefficient \( \zeta \) becomes a dynamical variable in this approach. \( T \) is the temperature of the thermostat (here in energy units, Boltzmann constant \( k_B = 1 \)). In the equation for \( p_k \), we have taken into account that \( \rho_l \cdot \pi_l = 0 \) for all \( l = 1, \ldots, N \). \( (2N + 3) \) represents the number of degrees of freedom of a polymer chain with \( N + 1 \) monomers and \( N \) fixed bond lengths and thus \( NT \) is the canonical average of the total internal kinetic energy. In this way the average of the term \( \sum_{k=0}^{N} p_k^2 / mT \) is equal to \( 2N \) if the internal motion is ergodic. If the dynamics is ergodic, the quantity in parentheses in equation (38) is a sum of independent random variables, with zero mean and unit variance. Thus the mean amplitude of this quantity is \( \sqrt{2N} \) and in this way the mean value of \( \dot{\zeta} \) is \( \approx \tau / t_0 \). Note that this type of coupling to the thermostat conserves the total linear momentum and also the direction of the total angular momentum of the chain (see appendix C).

5. Numerical simulations of polymer-chain dynamics

5.1. The dynamics of a monomeric chain

We illustrate the method in the case of a chain of \( N = 50 \) identical monomers of unit mass \( M = 1 \), connected by rigid bonds of unit length. The interaction potential between monomers is taken as

\[ V(r) = R_0^4 - \varepsilon R_0^2 / r^2, \quad (39) \]

where \( R_0 = 1 \) and \( \varepsilon = 0.3 \) yields the strength of the attractive interaction. The minimum of \( V \) occurs at \( r_{\text{min}} = \sqrt{2/\varepsilon} \approx 2.582 \), yielding \( V_{\text{min}} = -\varepsilon^2 / 4 = -0.0225 \). The temperature is taken as \( T = 0.025 \) and the friction \( \zeta \) obeys the equation of motion (see appendix C): 

\[ \dot{\zeta}_x = \frac{\tau}{t_0} \sqrt{\frac{1}{2N} \left( \sum_{k=0}^{N} \frac{p_k^2}{mT} - \frac{2N}{3} \right)}, \quad (40) \]

with similar expressions for the \( y \) and \( z \) components. Here we use \( t_0 = 1 \) and \( \tau = 2.42 \).

Typical chain configurations attained in the early stages of the simulation are shown in figure 1. After about \( 10^7 \) integration steps, corresponding to a time \( t = 150 \), the chain attains a
Figure 1. The configurational dynamics of the 50 monomer chain. The starting configuration is shown on top. Each plot has been obtained after every $2 \times 10^6$ iterations, using the Adams–Bashforth–Moulton scheme [15] with an integration step $\Delta \tau = 1.5 \times 10^{-5}$, corresponding to time intervals $\Delta t = 30$. The different chain configurations have been shifted from each other for clarity.

rather compact structure. The resulting radius of gyration at equilibrium is $R_G \approx 2.65$ in units of bond length. Applications of this method to study protein structure will be considered in the next section.

The temporal behaviour of friction variables $\zeta_x$, $\zeta_y$, and $\zeta_z$ are shown in figure 2 for the simulation trajectory considered here. The probability distribution functions (PDFs) for friction and momenta have been calculated by recording all the corresponding values for times $t > 150$. While the standard deviation for each friction component is unity (cf equation (38) and text), the standard deviation of a monomer momentum component is given by,

$$\sigma = \sqrt{\frac{2}{3} \frac{N}{N+1} MT}, \quad (41)$$

a value which is used in figure 3 in order to display the numerically obtained PDF. As one can see from figure 3, the PDF of the momentum components have reached their asymptotic form, while the PDF for the frictions still display departures from Gaussian behaviour for times $t = 510$.

5.2. Reconstruction of the backbone structure of a protein

In the following, we apply the present method to the reconstruction of the backbone structure of a protein, represented by the linear chain connecting its $C_{\alpha}$ atoms. The structure of the backbone uniquely determines the contact matrix of the protein, from which the positions of all atoms in

New Journal of Physics 7 (2005) 2 (http://www.njp.org/)
Figure 2. Values of the frictions components (from top to bottom) $\zeta_x$, $\zeta_y$ and $\zeta_z$ versus time. The data correspond to a simulation of $34 \times 10^6$ integration steps, i.e., $t = 510$. The data sets have been shifted along the $y$-axis for clarity. The horizontal lines represent the values $\zeta = 0$. The chain configurations shown in figure 1 correspond to times $t \leq 150$.

Figure 3. Probability distribution functions $P(x)$ versus $x$ for each component of (normalized) momentum and friction: (a) $p = p_x/\sigma$ ($\square$), $p_y/\sigma$ ($\bigcirc$), and $p_z/\sigma$ ($\triangle$), where $\sigma \approx 0.128$ (cf equation (41)). (b) $\zeta_x$ ($\square$), $\zeta_y$ ($\bigcirc$), and $\zeta_z$ ($\triangle$). The normal distribution is displayed by continuous solid line.

the native structure can be obtained (see e.g., [16, 17]). We consider as target structure a (wild-type) Rubredoxin protein, choosing in particular the 1e8j structure [18]. From the all-atoms’ coordinates [19], we extract the positions of the $N = 51$ $C_\alpha$ atoms, which define our target backbone chain (see figure 4). For convenience we take $R_u$ as the unit length, the $C_\alpha-C_\alpha$ mean distance in 1e8j which turns out to be $R_u = 3.8$ Å.

We define two $C_\alpha$ atoms, $i$ and $j$, to be in contact when their distance in space is smaller than the cutoff $R_c = 1.9$ (in units of $R_u$), and their positions along the linear chain obey $|i - j| \geq 3$. 

New Journal of Physics 7 (2005) 2 (http://www.njp.org/)
Figure 4. The native backbone structure of Cα atoms in Rubredoxin 1e8j. In the second panel, we display a typical reconstructed structure obtained using the present algorithm.

In this way, the target structure is characterized by having 93 contacts. For each pair \( i, j \) that is in contact with the target structure, we assume that the following interaction potential \( V(r) \) is acting between them:

\[
V(r) = \frac{R_a^4}{r^4} - \varepsilon \frac{R_r^2}{r^2},
\]

where \( R_a = 0.8, R_r = 0.5 \) and \( \varepsilon = 0.7 \), independently of the type of amino acid, and \( r \) is the actual distance between monomers \( i, j \) at any instant of time. For two amino acids that are not in contact with the target structure, the interaction potential is given by equation (42) with \( \varepsilon = 0 \), i.e., we assume their interaction to be purely repulsive. With this simple scheme, we aim to reconstruct the target structure, using the thermostat rules described in the previous section, by starting from an elongated chain configuration. Here we use a mass \( M = 1 \) for all monomers, the temperature \( T = 0.005 \), the thermostat parameters \( \tau = 26 \) and \( t_0 = 1 \), and the integration step \( \Delta \tau = 10^{-6} \).

We show in figure 5 a typical result of the reconstruction simulation by starting from an elongated chain configuration. As indicated in the figure, both the number of native and wrong contacts grow initially as a function of time. The latter remains bounded below about 30, while the true contacts are continuously formed up to a time of about 200, when a plateau is reached. This plateau is found in all simulations performed and seems to be typical of the dynamical process. After the search period is over, at \( t \approx 350 \), the number of true contacts increases rapidly again and enters the (small) corridor delimited by the dashed lines in figure 5. These lines represent characteristic bounds for the fluctuations observed in separate simulations in which the initial configuration coincides with the target structure itself. This is done to control the stability of the relaxed structure and permits to determine the model parameters appropriately. One observes that the fluctuations in both the number of true and wrong contacts are consistent with the ones found...
Figure 5. The number of contacts between monomers are shown as a function of time, for a simulation dynamics of the linear polymer by starting from an elongated chain configuration. The wrong contacts between two monomers, i.e., those which are present during the dynamical evolution of the chain but are not present in the target structure, are indicated in the lower part of the picture. The native contacts between two monomers, which are formed during the reconstruction dynamics are indicated by the growing curves. The latter tends to reach the maximum value of 93 contacts for times $t > 350$. The broken line denotes the typical bounds observed during a relaxation dynamics of the chain structure by starting from the target configuration.

for a relaxed structure. A typical reconstructed structure obtained using the present algorithm is displayed in figure 4.

The actual contacts taking place in a relaxation and a reconstruction simulation are shown in figure 6, where one can distinguish more clearly the different local environments displayed by the different structures. The examples shown in the figure report two cases having the same number of true and wrong contacts.

6. Final remarks

We have discussed a new efficient algorithm aimed at studying the long-time dynamics of polymer chains in the presence of a thermostat, by keeping the monomer–monomer link distances fixed over the whole dynamical trajectory of the system. A particular attractive extension of the algorithm presented here is to environmental-dependent couplings, which could allow the study of the long-time polymer dynamics in realistic environments. The well-known Nosé–Hoover thermostat couples all the particles of a system at all times with the same strength to a fictitious thermostat. In this manner all the parts of a system exchange energy with the thermostat at all times and at the same average rate. A polymer however, depending on its instantaneous structure, can
Figure 6. Contact matrix for a typical relaxed target structure and for a reconstructed structure. The native contacts are indicated by open squares, which are filled by full circles when they are present in the structure. The wrong contacts are indicated by ♦. In the two examples shown here, there occurs a single missing contact, (32) and (36) for the relaxed structure and the (15) and (27) for the reconstructed one, and the same number, i.e., 20, of wrong contacts.

have either its whole length or only a small fraction of its monomers exposed to the surrounding medium. As a result, the polymer dynamics is expected to depend dramatically on what fraction of its monomers can exchange energy with the surrounding medium in a given time interval. This allows for the investigation of the long timescale dynamics of polymer chains, which can not be otherwise accessible with traditional methods. The present results should also be useful in the context of protein structure predictions and in the modelling of nanoscale polymer materials.

Acknowledgments

We thank U Bastolla, W Dietrich, D Kusnezov, F Marini, M Porto, B Spivak and M Vendruscolo for discussions. We would also like to thank F Marini for help.

Appendix A. A few Poisson and Dirac–Poisson brackets

In the following formulae, the subscripts for $r_k$ and $p_k$ take values in the interval $k = 0, \ldots, N$, while for the quantities $X_k, Y_k$ only in the interval $k = 1, \ldots, N$ and for $D_{k,l}$ the indices take values in the interval $1, \ldots, 2N$. If a subscript is outside these intervals the corresponding quantity is identically vanishing. Only the following Poisson brackets between $r_k$, $p_k$ and the constraints $X_k, Y_k$ are nonvanishing:

$$\{ r_k, Y_k \} = \frac{2}{m} \rho_k, \quad (A.1)$$

$$\{ r_k, Y_{k+1} \} = -\frac{2}{m} \rho_{k+1}, \quad (A.2)$$
The explicit form of the Dirac–Poisson brackets demonstrates that the phase space variables \( r_k \) and \( p_k \) are not canonically conjugate with respect to the Dirac symplectic structure:

\[
\{ r_k, r_l \}_D = 0,
\]

\[
\{ r_k, p_l \}_D = I \delta_{kl} - \frac{4}{m} D_{k+N,l} \rho_k \otimes \rho_l + \frac{4}{m} D_{k+1+N,l} \rho_{k+1} \otimes \rho_l + \frac{4}{m} D_{k+N,l+1} \rho_k \otimes \rho_{l+1}
- \frac{4}{m} D_{k+1+N,l+1} \rho_{k+1} \otimes \rho_{l+1},
\]

\[
\{ p_k, p_l \}_D = 4 D_{k,l} \rho_k \otimes \rho_l - 4 D_{k+1,l} \rho_{k+1} \otimes \rho_l + 4 D_{k+1,l+1} \rho_{k+1} \otimes \rho_{l+1} - \frac{4}{m} D_{k+N,l+1} \pi_k \otimes \rho_l + \frac{4}{m} D_{k+1+N,l+1} \pi_{k+1} \otimes \rho_l
- \frac{4}{m} D_{k+N,l} \pi_k \otimes \rho_{l+1} + \frac{4}{m} D_{k+1+N,l+1} \pi_{k+1} \otimes \rho_{l+1}
+ \frac{4}{m} D_{k,l+1+N} \rho_k \otimes \pi_l - \frac{4}{m} D_{k+1,l+1+N} \rho_{k+1} \otimes \pi_l
- \frac{4}{m} D_{k,l+1+N} \rho_k \otimes \pi_{l+1} + \frac{4}{m} D_{k+1,l+1+N} \rho_{k+1} \otimes \pi_{l+1}.
\]

Here \( I \) is the \( 3 \times 3 \) identity matrix and \( \otimes \) stands for the direct product of two \( 3 \)-vectors, i.e. a symmetric second-rank tensor.

**Appendix B. An alternative derivation of the Lagrangian and Hamiltonian dynamics**

Let us define the following \( 3(N + 1) \)-dimensional vectors:

\[
\mathcal{R}_0 = (r_0, r_1, \ldots, r_N),
\]

\[
\mathcal{R}_1 = (-\rho_1, \rho_1, 0, \ldots, 0),
\]

\[\text{New Journal of Physics 7 (2005) 2 (http://www.njp.org/)}\]
\[ R_2 = (0, -\rho_2, \rho_2, 0, \ldots, 0), \quad (B.3) \]
\[ \vdots \]
\[ R_N = (0, \ldots, -\rho_N, \rho_N), \quad (B.4) \]
\[ P_0 = (p_0, p_1, \ldots, p_N), \quad (B.5) \]
\[ P_1 = (-\pi_1, \pi_1, 0, \ldots, 0), \quad (B.6) \]
\[ P_2 = (0, -\pi_2, \pi_2, 0, \ldots, 0), \quad (B.7) \]
\[ \vdots \]
\[ P_N = (0, \ldots, -\pi_N, \pi_N). \quad (B.8) \]

The primary and secondary constraints can then be written as follows:
\[ \chi_k = \langle R_k | R_0 \rangle - l^2 = \frac{1}{2} \langle R_k | R_k \rangle - l^2 = 0, \quad (B.9) \]
\[ \gamma_k = \frac{2}{m} \langle R_k | P_0 \rangle = \frac{2}{m} \langle P_k | R_0 \rangle = 0, \quad (B.10) \]

and the trivial scalar product in this 3\((N + 1)\) space is defined as expected
\[ \langle R_0 | P_0 \rangle = \sum_{k=0}^{N} r_k \cdot p_k. \quad (B.11) \]

A new Lagrangian, in which the constraints are taken into account automatically can be defined as follows:
\[ \mathcal{L}_0 = \sum_{k=1}^{N} \mathcal{I}_{kl} \langle \dot{R}_k | \dot{R}_l \rangle, \quad (B.12) \]
\[ L_0 = \frac{m}{2} \left[ \langle \dot{R}_0 | \dot{R}_0 \rangle - \sum_{k,l=1}^{N} \langle \dot{R}_0 | R_k \rangle J_{kl} \langle R_l | \dot{R}_0 \rangle \right] - U = \frac{m}{2} \langle \dot{R}_0 | \dot{R}_0 \rangle - U = \mathcal{L}, \quad (B.13) \]

where
\[ \mathcal{I}_{kl} = \langle R_k | R_l \rangle, \quad \sum_{n=1}^{N} \mathcal{I}_{kn} \mathcal{J}_{nl} = \delta_{kl}. \quad (B.14) \]

\( \mathcal{L}_0 = \mathcal{L} \) in equation (B.13) only when the constraints are fulfilled and as a result \( \langle \dot{R}_0 | R_k \rangle = 0 \) for all \( k \). The matrix \( \mathcal{I} = m/4U \) is identical with the matrix appearing in equations (8) for the Lagrange multipliers \( \lambda_k \) or the matrix appearing in equations (28) and (29) for the quantities \( a_k \) and \( b_k \). Thus any components of the velocity vector \( \dot{R}_0 \) along any vector \( R_k \) do not give any contribution to the kinetic energy. In this way, we have simply projected out of the kinetic energy the contributions corresponding to link changes and they do not affect the dynamics. Thus the primary constraints (more exactly, its time derivatives) have been absorbed into the definition of

New Journal of Physics 7 (2005) 2 (http://www.njp.org/)
this new Lagrangian. One can now derive the Lagrange equations of motion corresponding to
this new Lagrangian. When considering the variation of the action, we have to make sure that
no link variations are allowed. This can be achieved simply by replacing the variations of the
coordinates according to the simple prescription
\[ \delta R_0 \rightarrow \delta R_0 - \sum_{k,l=1}^{N} R_k J_{kl} (R_l | \delta R_0), \]  
which is equivalent to considering only those coordinate variations that leave the links fixed.
It is a relatively straightforward exercise to show that in this way one recovers the Lagrange
equations of motion derived in the main text. The momenta are defined as
\[ P_0 = \frac{\partial L}{\partial \dot{R}_0} = m \left( \dot{R}_0 - \sum_{k,l=1}^{N} R_k J_{kl} (R_l | \dot{R}_0) \right) = m \dot{R}_0. \]
The last equality holds because only such motions where \( \langle \delta R_0 | R_k \rangle = 0 \) are considered. It is
obvious also from the above relations that
\[ \langle P_0 | R_k \rangle = \langle P_k | R_0 \rangle = \frac{m}{2} \dot{\gamma}_k = 0, \]  
and consequently momenta are not at all independent and they have to satisfy the secondary
constraints. The set of secondary constraints has thus been obtained in a somewhat unusual
fashion, by not allowing explicitly any dynamics in the Lagrangian \( L_0 \) in those directions, which
 correspond to the primary constraints. The new Hamiltonian of the system now reads as
\[ H_0 = \frac{1}{2m} \left( \langle P_0 | P_0 \rangle - \sum_{k,l=1}^{N} \langle P_0 | R_k \rangle J_{kl} (R_l | P_0) \right) + U = \frac{1}{2m} \langle P_0 | P_0 \rangle + U = H. \]
The last two equalities in equation (B.18) hold only when the secondary constraints are fulfilled,
i.e., \( \langle P_0 | R_k \rangle = 0 \) for all \( k \).

Appendix C. Coupling of a polymer chain to a thermostat

If one desires to couple the polymer with a thermostat and study its dynamics, the knowledge
of the Hamiltonian dynamics becomes imperative. A way to avoid the study of the Hamiltonian
dynamics in the presence of constraints is to use generalized coordinates [1, 7, 8]. One aspect
that is often overlooked in this case is the fact that the introduction of generalized or internal
coordinates makes the description very complicated. The main reason is the appearance of
nonabelian gauge fields, a fact which has been brought to light only relatively recently in
molecular physics [20]. The presence of nonabelian gauge fields means that in the equations
of motion one has strong singularities (poles), which requires quite a sophisticated treatment
(use of several different coordinate atlases), not always simple to implement and to deal with.
In this respect the situation is similar to the treatment of rigid body rotations, a much simpler
problem, which requires the introduction of quaternions [21] instead of the well-known Euler

New Journal of Physics 7 (2005) 2 (http://www.njp.org/)
angles, i.e., again another form of constraints that can be dealt with in a similar way which we have chosen in the present work [22]. The treatment of holonomic constraints in a Langevin approach is a challenging problem, and accurate numerical solutions of such stochastic equations are not a trivial problem. On one hand, the presence of white noise requires in principle an infinitesimally small integration time step. On the other hand, very high frequencies should not be relevant for the physics on the present computer.

In the following we discuss the treatment of a thermostat within the Hamiltonian dynamics by making use of the Dirac–Poisson brackets formalism. Since some of the ensuing equations are rather long if one uses standard notations, we shall introduce here a somewhat more compact one.

Let us denote the phase space variables \( r_k, p_k \) with \( k = 0, \ldots, N \) by \( z_\alpha \), where \( \alpha = 1, \ldots, 6(N + 1) \). The Dirac–Poisson bracket between two arbitrary functions, see equation (12), can then be written as

\[
\{ A, B \}_D = \sum_{\alpha, \beta = 1}^{6(N+1)} G_{\alpha\beta}(z) \frac{\partial A}{\partial z_\alpha} \frac{\partial B}{\partial z_\beta},
\]

where \( G_{\alpha\beta}(z) \equiv \{ z_\alpha, z_\beta \}_D = -G_{\beta\alpha}(z) \) (C.2)
and the dependence on \( z \) stands for dependence on all phase space variables \( z_\alpha \). The Hamiltonian equation of motion reads as

\[
\dot{z}_\alpha = \{ z_\alpha, \mathcal{H} \}_D = \sum_{\beta = 1}^{6(N+1)} G_{\alpha\beta}(z) \frac{\partial \mathcal{H}}{\partial z_\beta},
\]

(C.3)

When coupled to a thermostat through one pseudo-friction coefficient only the equations of motion become

\[
\dot{z}_\alpha = \sum_{\beta = 1}^{6(N+1)} G_{\alpha\beta}(z) \frac{\partial \mathcal{H}}{\partial z_\beta} - \kappa g(\zeta) \sum_{\beta = 1}^{6(N+1)} G_{\alpha\beta}(z) A_\beta(z),
\]

where \( g(\zeta) \) is a function of the pseudo-friction coefficient \( \zeta \), \( A_\beta(z) \) is a vector function of the phase space variables \( z \) and \( \kappa \) is a constant, all of these are to be determined later. This type of coupling to a thermostat was called twisted in [11].

We are interested in generating a phase space density distribution \( f(z, \zeta, t) \) satisfying the generalized Liouville equation [11]

\[
\frac{\partial f}{\partial t} + \sum_{\alpha = 1}^{6(N+1)} \frac{D(f \dot{z}_\alpha)}{Dz_\alpha} + \frac{\partial (f \dot{\zeta})}{\partial \zeta} = 0,
\]

(C.5)

where the symbol \( D/Dz_\alpha \) is a generalized partial derivative defined as

\[
\frac{Df}{Dz_\alpha} = \frac{\partial f}{\partial z_\alpha} - \sum_{\mu = 1}^{2N} \frac{\partial Z_\mu}{\partial z_\alpha} \mathbf{O}_{\mu\nu} \frac{\partial Z_\nu}{\partial \zeta},
\]

(C.6)
The angle brackets stand for the scalar product of the corresponding vectors (here, the gradients of the corresponding functions) in the \(6(N + 1)\) phase space. By definition, this generalized partial derivative satisfies the following relations:

\[
\langle \frac{D f}{D z} \mid \frac{\partial Z_{\mu}}{\partial z} \rangle = 0,
\]

(C.8)

\[
\sum_{\beta = 1}^{6(N+1)} G_{\alpha \beta}(z) \frac{\partial H}{\partial \zeta_{\beta}} = \sum_{\beta = 1}^{6(N+1)} G_{\alpha \beta}(z) \frac{\partial \mathcal{H}}{\partial \zeta_{\beta}}.
\]

(C.9)

By taking the equilibrium density distribution

\[
f(z, \zeta) \propto \exp \left[ -\frac{\mathcal{H}}{T} - G(\zeta) \right],
\]

(C.10)

to be a solution of the generalized Liouville equation, we shall derive an equation of motion for the pseudo-friction coefficient \(\zeta\). We shall also require that \([11]\)

\[
g(\zeta) = \sigma \frac{dG(\zeta)}{d\zeta},
\]

(C.11)

\[
\frac{\partial \dot{\zeta}}{\partial \zeta} = 0,
\]

(C.12)

where \(\sigma\) is a constant, the choice of which we shall discuss later. After inserting the equilibrium density distribution into the generalized Liouville equation and using the above relations one readily obtains

\[
\kappa g(\zeta) \frac{1}{T} \sum_{\alpha, \beta = 1}^{6(N+1)} G_{\alpha \beta}(z) \frac{\partial A_{\beta}(z)}{\partial \zeta_{\alpha}} + \sum_{\alpha, \beta = 1}^{6(N+1)} \frac{D}{D\zeta_{\alpha}} \left[ G_{\alpha \beta}(z) \frac{\partial \mathcal{H}}{\partial \zeta_{\beta}} \right] - \kappa g(\zeta) \sum_{\alpha, \beta = 1}^{6(N+1)} A_{\beta}(z) \frac{D G_{\alpha \beta}(z)}{D\zeta_{\alpha}}
\]

\[
- \kappa g(\zeta) \sum_{\alpha, \beta = 1}^{6(N+1)} G_{\alpha \beta}(z) \frac{\partial A_{\beta}(z)}{\partial \zeta_{\alpha}} - \sigma g(\zeta) \dot{\zeta} = 0.
\]

(C.13)

The second and the third term vanish identically, which can be checked by considering the generalized Liouville equation for an isolated system (in the absence of a thermostat) with an arbitrary Hamiltonian (in particular all Hamiltonians of the form \(\mathcal{H} = z_{a}\)). Thus one obtains

\[
\dot{\zeta} = \sigma \kappa \left[ \frac{1}{T} \sum_{\alpha, \beta = 1}^{6(N+1)} G_{\alpha \beta}(z) \frac{\partial \mathcal{H}}{\partial \zeta_{\alpha}} A_{\beta}(z) - \sum_{\alpha = 1}^{6(N+1)} G_{\alpha \beta}(z) \frac{\partial A_{\beta}(z)}{\partial \zeta_{\alpha}} \right].
\]

(C.14)

As one can see, in the final analysis for the twisted type of coupling to the thermostat all generalized partial derivatives are equivalent to ordinary partial derivatives.

New Journal of Physics 7 (2005) 2 (http://www.njp.org/)
We shall present now an explicit type of coupling to a thermostat, which, in spite of its simplicity, in most cases seems to lead to a very efficient thermalization of such systems. Even though we have derived the above equations by assuming one pseudo-friction coefficient only, in practice we use several pseudo-friction coefficients in order to ensure ergodic properties of this system of equations of motion and a more efficient coverage of the phase space. Quite often the simpler Nosé–Hoover coupling does not display an ergodic behaviour and the introduction of several pseudo-friction coefficients is the only cure of this deficiency (see [11]–[14]). In particular, the coupling we have described in the main text conserves the total linear momentum of the system and also the direction of the total angular momentum. Since we are interested in the internal dynamics of a polymer chain, the conservation of the total linear momentum is irrelevant, however the fact that at least the direction of the total angular momentum is conserved, might lead to nonergodic behaviour in the internal degrees of freedom, since overall rotational motion does not decouple exactly from the internal motion. For that reason, the coupling we shall consider amounts to modification of the usual Hamiltonian equations of motion given by (C.3) and (C.4) in the following form:

\[ -\frac{\partial U}{\partial r_k} \implies \frac{\partial U}{\partial r_k} - \kappa(\zeta_x p_{xk}, \zeta_y p_{yk}, \zeta_z p_{zk}), \]  

(C.15)

where the last term is a 3-vector. Thus we have introduced a specific vector function \( A_\beta(z) \). In particular, one has to use this prescription when one determines the constants \( a_k \) appearing in equations (29), (32) and (34), so as to have the constraints satisfied in the presence of the thermostat as well. Here \( p_{xk} \) stands for the \( x \)-component of the linear momentum \( p_k \) of the \( k \)th monomer. One can easily check that the total linear momentum of the chain is an integral of motion even in the presence of the thermostat, however the direction of the total angular momentum is not conserved anymore for this type of coupling, as opposed to the coupling described in the main text. We have assumed here that \( G(\zeta_x) = \zeta_x^2/2 \), etc. For the three pseudo-friction coefficients \( \zeta_x, \zeta_y, \zeta_z \), one thus derives the following equations of motion

\[ \dot{\zeta}_x = \sigma \kappa \left[ \sum_{k=0}^{N} \frac{p_{xk}^2}{mT} - \sum_{k=0}^{N} \{x_k, p_{xk}\}_D \right] \]  

(C.16)

and similar equations for \( \zeta_y \) and \( \zeta_z \). An explicit expression for the Dirac–Poisson bracket \( \{x_k, p_{xk}\}_D \) was given in appendix A, see equation (A.8). An appropriate value for the coupling constant \( \kappa \) is

\[ \kappa = \frac{\tau}{\hbar_0} \sqrt{\frac{1}{2N}}, \]  

(C.17)

where \( \tau \) is a constant (generally of order \( O(1) \)). If chosen in this way, the phase space variables \( r_k, p_k \) and the pseudo-friction coefficients \( \zeta_x, \zeta_y, \zeta_z \) will have comparable time derivatives and in this way the energy exchange between the thermostat and the polymer chain is the most effective. One can of course consider other situations as well and it certainly makes sense to do so, however, the simulations will have to be much longer in such a case. These different types of coupling, which differ by the rate at which the energy is exchanged with the thermostat can be simulated by an appropriate choice of the coupling constants \( \sigma \) and \( \kappa \). For the time being we shall use \( \sigma = 1 \) and relate further refinements to the future. In order to speed up the simulations and without a significant loss of accuracy, one can replace the last term \( \sum_{k=0}^{N} \{x_k, p_{xk}\}_D \) by its average value \( 2N/3 \) (when the contribution of the overall translation of the chain is excluded).
Appendix D. Environment-dependent coupling of a polymer chain to a thermostat

There is a certain freedom in the coupling of a system to a thermostat, which has not been explored until now. In particular, the role of a thermostat for a polymer chain is played by solvent molecules. One might reasonably assume that the coupling of a given monomer to the thermostat should depend on how many other monomers of the chain are in the immediate vicinity of the given monomer. If the polymer chain is unfolded, all parts of the chain will be exposed to the solvent. However, after the chain has folded, either partially or fully, the monomers inside interact only very weakly with the solvent, their thermalization being thus mediated through the neighbouring monomers. One can simulate such a situation by changing the form of coupling we have described in appendix C as follows:

\[-\frac{\partial U}{\partial r_k} \Rightarrow -\frac{\partial U}{\partial r_k} - \kappa g(r_k)(\zeta_x p_{xk}, \zeta_y p_{yk}, \zeta_z p_{zk})\],

where \(0 < g(r_k) \leq 1\) is a function that depends on the environment, representing the degree of exposure of monomer \(k\) (with coordinates \(r_k\)) to the thermostat. (Implicitly \(g(r_k)\) depends on all monomer coordinates.)

Indeed, we can imagine two opposite situations. On one hand, if the monomer \(k\) has only one (if it is at either end of the polymer chain) or two (otherwise) close monomer neighbours then \(g(r_k) \approx 1\). On the other hand, if the given monomer is surrounded by many other monomers of the chain then \(g(r_k) \approx 0\). Except for the appearance of the function \(g(r_k)\), the equations of motion for the phase space variable remain the same as those in the main text or those in appendix C. The equation for the pseudo-friction coefficient \(\zeta_x\) for example becomes now

\[\dot{\zeta}_x = \sigma \kappa \left[ \sum_{k=0}^{N} g(r_k) \frac{P_{xk}^2}{mT} - \sum_{k=0}^{N} g(r_k) \{x_k, p_{xk}\}_D \right].\]  

(D.2)

Some care should be taken now in choosing appropriate values for the coupling constants \(\kappa\) and \(\sigma\), since the average magnitude of the quantity in square brackets in the above equation for \(\zeta_x\) is now environment dependent. The time dependence of this quantity is controlled now by two factors: (i) thermal fluctuations, which are relatively fast changes, and (ii) environmental changes, which are relatively slow. Previously, in appendix C, we have chosen the value for the coupling constant \(\kappa\) such that the characteristic frequency for the pseudo-friction coefficient \(\zeta_x\) was comparable with the fastest mode of the polymer chain. In this case the coupling to the thermostat is most effective. It is not difficult to realize that the formalism we have described, does not actually require the coupling constants \(\kappa\) and \(\sigma\) to be constants, and can be chosen in particular as time dependent. This is what we propose to do here, namely to choose

\[\kappa = \frac{\tau}{\hbar \sqrt{2 g_{eff}(t)}}\],

(D.3)

where the effective (time dependent) number of \(x\)-degrees of freedom coupled to the thermostat is

\[g_{eff}(t) = \sum_{k=0}^{N} g(r_k)\]  

(D.4)
and the overline stands for some type of time average over the chain configuration, consequently
over the slowest mode of the polymer chain. There is no reason to compute this average exactly,
only to have a fair guess of its value. In general, the described formalism should work even if
the coupling constants are arbitrary (as long as the dynamics is ergodic). Only the efficiency of
the simulation is affected by the actual values of these couplings, which can be varied freely,
within an order of magnitude or so from their ‘optimal’ values, without affecting significantly
the quality of a given simulation. Since we have at our disposal two ‘coupling constants’ $\kappa$ and $\sigma$,
by varying $\kappa$ or $\tau$ for example, while maintaining through equation (D.3), one can still control the
strength of the coupling of the polymer chain to the thermostat, i.e., the rate at which energy is
exchanged between the polymer chain and the solvent. Note that the results quoted in appendix
C are recovered when $g(r_k) = 1$, and equation (D.4) yields $g_{\text{eff}}(t) = N$.

The reader should have realized by now that the class of couplings to a thermostat, we
have described, allow for significant freedom in modelling various situations in a very effective
manner. Moreover, one can easily consider other schemes as well, e.g., coupling only part(s) of
the chain to a thermostat, or even coupling each monomer to its ‘own’ thermostat, or increasing
the number of pseudo-friction coefficients and considering other types of coupling as well, see
e.g., [11, 13, 14], and so on.

**Appendix E. Polymer chains with fixed bond lengths and fixed angles between bonds**

Another interesting class of polymer chains, which can be effectively treated using the present
formalism represents polymer chains with fixed bond lengths and fixed angles between successive
links. Let us define the following $2N - 1$ ($\alpha = 1, \ldots, 2N - 1$) primary constraints $X_\alpha$ as

\[ X_{2k} = \rho_k^2 - l^2 = 0 \quad \text{for } k = 1, \ldots, N \quad (E.1) \]

\[ X_{2k-1} = (\rho_{k+1} - \rho_{k-1})^2 - 2l^2(1 - \cos \theta) = 0 \quad \text{for } k = 1, \ldots, N - 1, \quad (E.2) \]

where $\theta$ is the angle between two consecutive links. The secondary constraints $Y_\alpha$ are defined in
an analogous manner as in section 3

\[ Y_{2k} = \{ X_{2k}, H \} = \frac{2}{m} (r_k - r_{k-1}) \cdot (p_k - p_{k-1}) \quad \text{for } k = 1, \ldots, N, \quad (E.3) \]

\[ Y_{2k-1} = \{ X_{2k-1}, H \} = \frac{2}{m} (r_{k+1} - r_{k-1}) \cdot (p_{k+1} - p_{k-1}) \quad \text{for } k = 1, \ldots, N - 1. \quad (E.4) \]

In an analogous manner, as in section 3, one can define the Dirac–Poisson bracket, see equations
(12) and (13) but now with $2(2N - 1)$ constraints and the corresponding equations for the phase
space variables $r_k, p_k$ now read as

\[ \dot{r}_k = \frac{p_k}{m}, \quad (E.5) \]

\[ \dot{p}_k = -\frac{\partial U}{\partial r_k} - \lambda_k (r_k - r_{k-1}) + \lambda_{k+1} (r_{k+1} - r_k) - \mu_k (r_k - r_{k-2}) + \mu_{k+2} (r_{k+2} - r_k). \quad (E.6) \]
The velocities have to be consistent with these constraints, which in this case amounts to

\[(r_k - r_{k-1}) \cdot (p_k - p_{k-1}) = 0 \quad \text{for } k = 1, \ldots, N, \tag{E.7}\]

\[(r_{k+1} - r_{k-1}) \cdot (p_{k+1} - p_{k-1}) = 0 \quad \text{for } k = 1, \ldots, N - 1. \tag{E.8}\]

It is sufficient that these constraints on the ‘momenta’ are fulfilled at the initial time. The Dirac–Poisson dynamics conserves the constraints irrespective of the Hamiltonian, see section 3. The quantities \(\lambda_k\) (nonvanishing for \(k = 1, \ldots, N\) only) and \(\mu_k\) (nonvanishing for \(k = 1, \ldots, N - 1\) only) can be determined also by requiring that

\[
(\dot{r}_k - \dot{r}_{k-1})^2 + (\dot{r}_k - \dot{r}_{k-1}) \cdot (r_k - r_{k-1}) = 0 \quad \text{for } k = 1, \ldots, N, \tag{E.9}
\]

\[
(\dot{r}_{k+1} - \dot{r}_{k-1})^2 + (\dot{r}_{k+1} - \dot{r}_{k-1}) \cdot (r_{k+1} - r_{k-1}) = 0 \quad \text{for } k = 1, \ldots, N - 1. \tag{E.10}
\]

One thus obtains

\[
\lambda_{k+1}(r_{k+1} - r_k) \cdot (r_k - r_{k-1}) - 2\lambda_k(r_k - r_{k-1})^2 + \lambda_{k-1}(r_{k-1} - r_{k-2}) \cdot (r_k - r_{k-1})
\]

\[+ \mu_{k+2}(r_{k+2} - r_k) \cdot (r_k - r_{k-1}) - \mu_{k+1}(r_{k+1} - r_{k-1}) \cdot (r_k - r_{k-1})
\]

\[+ \mu_k(r_k - r_{k-2}) \cdot (r_k - r_{k-1}) + \mu_{k-1}(r_{k-1} - r_{k-3}) \cdot (r_k - r_{k-1})
\]

\[= -\frac{1}{m} (p_k - p_{k-1})^2 - (r_k - r_{k-1}) \cdot \left( \frac{\partial U}{\partial r_k} - \frac{\partial U}{\partial r_{k-1}} \right), \tag{E.11}\]

\[
\mu_{k+2}(r_{k+2} - r_k) \cdot (r_k - r_{k-2}) - 2\mu_k(r_k - r_{k-2})^2 + \mu_{k-2}(r_{k-2} - r_{k-4}) \cdot (r_k - r_{k-2})
\]

\[+ \lambda_{k+1}(r_{k+1} - r_k) \cdot (r_k - r_{k-2}) - \lambda_k(r_k - r_{k-1}) \cdot (r_k - r_{k-2})
\]

\[+ \lambda_{k-1}(r_{k-1} - r_{k-2}) \cdot (r_k - r_{k-2}) + \lambda_{k-2}(r_{k-2} - r_{k-3}) \cdot (r_k - r_{k-2})
\]

\[= -\frac{1}{m} (p_k - p_{k-1})^2 - (r_k - r_{k-2}) \cdot \left( \frac{\partial U}{\partial r_k} - \frac{\partial U}{\partial r_{k-2}} \right). \tag{E.12}\]

The underlined scalar products in these equations have fixed values, due to the constraints as follows:

\[(r_{k+1} - r_k) \cdot (r_k - r_{k-1}) = -l^2 \cos \theta, \tag{E.13}\]

\[(r_{k+1} - r_{k-1}) \cdot (r_k - r_{k-1}) = l^2(1 - \cos \theta), \tag{E.14}\]

\[(r_k - r_{k-1})^2 = l^2, \tag{E.15}\]

\[(r_k - r_{k-2})^2 = 2l^2(1 - \cos \theta). \tag{E.16}\]

If one forms the \((2N - 1)\)-vector \((\lambda_1, \mu_2, \lambda_2, \ldots, \mu_N, \lambda_N)\), then the above equations can be written as a system of inhomogeneous linear equations with nine diagonals (banded matrix). The coupling to a thermostat can be realized in a similar manner as for the case of only fixed bond lengths, see equations (C.15) and (C.16) in appendix C. The only difference is that the average value of the term \(\sum_{k=0}^{N} \{x_k, p_{sk}\}_D\) in equation (C.16) is now \((N + 1)/3\), where \(N + 1\) is the number of internal degrees of freedom in this case.
References

[1] Doi M and Edwards S F 1994 The Theory of Polymer Dynamics (Oxford: Clarendon)
[2] Binder K (ed) 1995 Monte Carlo and Molecular Dynamics Simulations in Polymer Science (Oxford: Oxford University Press)
[3] Öttinger H Ch 1995 Stochastic Processes in Polymeric Fluids: Tools and Examples for Developing Simulation Algorithms (Berlin: Springer)
[4] Berendsen H J C 1987 Comput. Phys. Commun. 44 233 and references therein
   Mitchels J P J 1987 Comput. Phys. Commun. 44 217
[5] Morris G P and Evans D J 1991 Comput. Phys. Commun. 62 267
[6] Ryckaert J P, Ciccotti G and Berendsen H J C 1977 J. Comput. Phys. 23 327
[7] Mazur A, Dorofeev V and Abagyan R 1991 J. Comput. Phys. 92 261
[8] Jain A, Vaideli N and Rodriguez G 1993 J. Comput. Phys. 106 258
[9] Kneller G R and Hinsen K 1994 Phys. Rev. E 50 1559
   Hinsen K and Kneller G R 1995 Phys. Rev. E 52 6868
   Kneller G R and Mülders T 1996 Phys. Rev. E 54 6825
   Mülders T, Toxvaerd S and Kneller G R 1997 Phys. Rev. E 58 6766
[10] Dirac P A M 1950 Can. J. Math. 2 147
   Dirac P A M 1964 Lectures on Quantum Mechanics (New York: Yeshiva University)
[11] Kusnezov D and Bulgac A 1992 Ann. Phys. (NY) 214 180
[12] Nosé S 1984 J. Chem. Phys. 81 511
   Nosé S 1984 Mol. Phys. 52 255
   Nosé S 1991 Prog. Theor. Phys. Suppl. 103 1
   Hoover W G 1985 Phys. Rev. A 31 1695
   Hoover W G 1986 Molecular Dynamics (Lecture Notes in Physics vol 258) (New York: Springer)
   Hoover W G 1991 Computational Statistical Mechanics (Amsterdam: Elsevier)
[13] Bulgac A and Kusnezov D 1990 Phys. Rev. A 42 5045
   Kusnezov D, Bulgac A and Bauer W 1990 Ann. Phys. (NY) 204 155
   Bulgac A and Kusnezov D 1990 Phys. Lett. A 151 122
   Kusnezov D, Bulgac A and Sloan J 1992 Nucl. Phys. A 545 537c
   Sloan J H, Kusnezov D and Bulgac A 1992 Computational Quantum Physics vol 260 (New York: AIP) p 23
   Sloan J, Kusnezov D and Bulgac A 1992 Nucl. Phys. B 26 623
   Sloan J, Kusnezov D and Bulgac A 1992 Phys. Lett. B 296 379
[14] Ju N and Bulgac A 1993 Phys. Rev. B 48 2721
   Bulgac A and Kusnezov D 1992 Phys. Rev. B 45 1988
[15] Gear C W 1971 Numerical Initial Value Problems for Ordinary Differential Equations (Englewood Cliffs, NJ: Prentice-Hall)
[16] Porto M, Bastolla U, Roman H E and Vendruscolo M 2004 Phys. Rev. Lett. 92 218101
[17] Vendruscolo M, Kussell E and Domany E 1997 Fold. Des. 2 295
[18] Lamosa P, Brennan L, Vis H, Turner D L and Santos H 2001 Extremophiles 5 303
[19] http://www.expasy.org/cgi-bin/get-pdb-entry.pl?1e8j
[20] Tachibana A and Iwai T 1986 Phys. Rev. A 33 2262 and earlier references therein
   Littlejohn R G 1993 Proc. Int. Symp. on Coherent States (Oak Ridge, TN, June 14–17) at press
   Littlejohn R G and Reinsch M 1997 Rev. Mod. Phys. 69 213 and references therein
[21] Evans D J 1977 Mol. Phys. 34 317
   Evans D J and Murad S 1977 Mol. Phys. 34 327
[22] Bulgac A and Trache M 1996 J. Chem. Phys. 105 1131