Dynamical models of molecular chains and efficient integration algorithms

Armando Bazzani,1,2,3 Carlo Benedetti,2,3 Sandro Rambaldi,1,2,3 Luca Rossi,2,3 and Giorgio Turchetti1,2,3

1Centro Interdipartimentale Galvani, Università di Bologna
2Dipartimento di Fisica, Università di Bologna
3INFN, Sezione di Bologna
(Dated: 16 October 2007)

Chains of point masses and chains of rigid bodies are used to model biological polymers. To investigate their dynamics we propose a method which allows an efficient realization of the constraints jointly with a simple and accurate integration of the free rigid body motion. The method is quite effective to evolve the geodesic flow of a rigid body chain and the global performance depends on the computational complexity of the algorithms used to compute the interaction forces. Our approach is suitable to describe a chain of rigid bodies immersed in a thermal bath. In the method we propose, the constraints are realized by hard springs whose elastic constant is set to maximize the energy dissipation rate of a Runge-Kutta integrator scheme. Moreover the use of local Lagrangian coordinates is introduced using the possibility of a continuous change of chart, such that the distance from the coordinate singularities is the highest possible. For a chain of point masses the numerical results are checked with another method where the constraints are exactly realized by means of Lagrangian coordinates. When the chain is subject to regular interactions potentials plus a thermal bath the exact and approximate constraints realization provide comparable results.

PACS numbers: 45.40.-f, 36.20.Fz, 02.60.Lj, 07.05.Tp
Keywords: Rigidity constraints, hard springs, Runge-Kutta, flipping rule, molecular chains

I. INTRODUCTION

The remarkable progress of molecular biology has stimulated the development of the dynamical approach for the description of some basic processes such as the folding of proteins [1, 2].

During the last decade several models have been proposed to describe the evolution to the native state of a peptidic chain in an aqueous medium at a given temperature [3]. A funnel shaped landscape of the potential was suggested to cope with the observed folding times [4, 5]. In a suitable temperature range the trapping times on local valleys must be very short to allow a rapid relaxation towards the free energy minimum corresponding to the native state and the contribution of entropic factors is important [6]. By decreasing the temperature the trapping times on local minima increase and other metastable states are populated.

A very useful off-lattice model to study the folding dynamics is a three dimensional (3D) chain of labeled identical point masses with rigid massless links [7]. The label is needed to specify the interactions which depend on the hydrophilic, hydrophobic or neutral character of the aminoacid represented by the point mass. The effect of the aqueous medium can be described by an Ornstein-Uhlenbeck process or by colliding molecules in a thermostat. Using ad hoc phenomenological potentials (elastic, dihedral and Van der Waals) to describe the aminoacids interactions, the main secondary structures such as α-helix and β-sheets are obtained. However a study of the folding dynamics requires efficient numerical procedures since the number of degrees of freedom ranges from $10^2$ to $10^3$. A usual procedure realizes the constraints by rigid springs. Indeed the use of Cartesian coordinates simplifies the equations of motion and renders the integration procedures elementary. However hardening the springs decreases the oscillation period and requires a shorter time step. This realization of the constraints is computationally expensive and resolving the hard springs oscillations is of no interest. A widely used alternative (LINCS) consists in determining the constraint forces and integrating the equations of motion in Cartesian coordinates [8]. A system of linear equations must be solved at each time step and elastic constraint forces with unresolved oscillations are nevertheless introduced to avoid instabilities due to the accumulation of numerical errors. Other models for biopolymers substitute the chain of point-masses with a chain of rigid bodies. The constraints may be a single points (pivot) or two points links (hinge). Approximation methods to determine the constraint forces have been provided jointly with the use of quaternions (QSHAKE) to avoid the singularities occurring when local coordinates such as the Euler angles are used [9]. Realistic all atoms models [10] have a high number of constraints, and the computational cost is so high that the study of the folding process cannot be afforded. In this paper we discuss the possibility of realizing rigid constraints by using hard springs whose dynamics is not resolved by the integration scheme since, after a short transient, the energy exchange between the rigid springs and the other degrees of freedom is negligible. The time step is fixed by the smallest oscillation period we wish to resolve and the solution of linear systems to determine the constraints is avoided. The key point is the use of dissipative integration schemes, like Runge-Kutta of order $n = 3, 4, 5$, that act as high frequency filters. The
high frequency modes of the hard springs cannot be excited. Moreover the energy transfer from low frequency modes to the high frequency modes is so small so that the numerical scheme conserves energy with a good accuracy for long times. The elastic constant of the springs is computed as a function of the time step $\Delta t$, optimizing the energy dissipation rate of the constraints. The optimal ratio between the period $T$ of the spring dynamics and the time step turns out to be $T/\Delta t \sim 6$ for the Runge-Kutta scheme of order $n = 3$ and $T/\Delta t \sim 2$ for the schemes of order $n = 4, 5$. For $n = 4, 5$ the spring rigidity is considerably higher so that, taking into account the computational load and the accuracy, the fourth-order Runge-Kutta appears to be the best choice. The time step is chosen according to the desired accuracy.

In the case of a chain with $N$ point masses, we validate our procedure by a comparison with a numerical scheme that solves the equations of motion where the constraints forces are determined by solving a tridiagonal linear system as in LINCS. We integrate the equation for the $N$ spherical pendula using polar coordinates by changing the polar axis at each time step, if needed, to avoid at best the singularity. For a chain moving in a constant force field almost machine accuracy is reached, as far the conservation of first integrals is concerned, for a suitable time step. We have tested the accuracy of the Runge-Kutta integrator applied to the system with hard springs whose unresolved oscillations are optimally damped. For the same time step the accuracy with respect to a LINCS type algorithm is only slightly lower, but the algorithms are trivial to implement and considerably faster since the solution of the linear system to determine the constraints is avoided. The rigid body chain is realized by using rigid springs for the pivots together with local Lagrangian coordinates such as the Euler angles. The possibility of flipping to the best reference frame at any time step for every rigid body, jointly with the proper choice of the rigidity of the springs provides a numerically reliable evolution with a fourth-order Runge-Kutta integrator. An alternative numerical scheme of third order for the rigid body dynamics based on the quaternion algebra \cite{11} is discussed in Appendix C; in this case we have no singularity in the equations of motion. We have studied the dynamics of a rigid body chain in presence of external forces by the energy conservation to test the accuracy. The application to biopolymer dynamics may be considered if the computational load of the interactions can by made comparable with the load due the geodesic motion. This can be done by using hierarchical methods for the long range forces. The Runge-Kutta algorithms do not preserve the symplectic character of the dynamics, however the presence of a solvent within a thermostat makes the use of a symplectic integrator not really relevant.

The paper is organized into four sections. In section 2 we analyze the Runge-Kutta integration of linear chains discussing the stability condition and highest energy dissipation of hard springs. The procedure is applied to a chain of point masses moving in a uniform field. The implementation of a stochastic thermal bath is also outlined. In section 3 we realize exactly the constraints and show how to avoid the coordinate singularity by integrating the equations of motion for a spherical pendulum. A comparison between the chain with exact constraints and with hard springs is presented. In section 4 we consider a chain of rigid bodies and we integrate the equations of motion by using hard springs for the constraints and Euler angles allowing a flip of the polar axis at every time step. The accuracy for the motion in a uniform force field is tested from the energy integral. Possible applications are outlined in the concluding remarks.

II. HARD SPRINGS

The simplest way to realize a constraint consists in replacing it with a hard spring. This is often used in molecular dynamics calculations of polymeric chains. The serious drawback is that a hard spring has a small oscillation period and a very small time step must be chosen to resolve it. The procedure we propose renders the realization of constraints faster and more accurate. It is based on the simple observation that most Runge-Kutta integrators are dissipative if the time step $\Delta t$ is chosen in suitable range and that the realization of the constraint with a hard spring is optimal if the dissipation has the highest possible value. We prove this property in the case of a single hard spring and of a hard spring coupled to a soft one. Having examined a large sample of systems we believe that the property is quite general.

A. The single oscillator

We consider a point of unit mass moving along a line subject to an elastic force, namely a harmonic oscillator whose equations of motion read

$$\ddot{x} = \omega p, \quad \dot{p} = -\omega x. \quad (1)$$

The exact solution of the system is given by the map

$$x(t + \Delta t) = C(\alpha) \cdot x(t) + S(\alpha) \cdot p(t),$$
$$p(t + \Delta t) = -S(\alpha) \cdot x(t) + C(\alpha) \cdot p(t), \quad (2)$$

where $\alpha = \omega \Delta t$, $C(\alpha) = \cos(\alpha)$ and $S(\alpha) = \sin(\alpha)$. It is a simple exercise to verify that the second order Runge-Kutta integrator corresponds to replacing the functions $C(\alpha)$ and $S(\alpha)$ with their Taylor series truncated at the second order namely $C(\alpha) = 1 - \alpha^2/2$ and $S(\alpha) = \alpha$. The determinant of the map is $D = 1 + \alpha^2/4$ and gives the increase of the phase space area at every time step. The map of the fourth order Runge-Kutta is still given by Eq. (2) where $C(\alpha)$ and $S(\alpha)$ are the the Taylor series of the $\cos(\alpha)$ and $\sin(\alpha)$ truncated at fourth order

$$C(\alpha) = 1 - \frac{\alpha^2}{2} + \frac{\alpha^4}{24}, \quad S(\alpha) = \alpha - \frac{\alpha^3}{3}. \quad (3)$$
In this case the determinant is given by

\[ S^2(\alpha) + C^2(\alpha) = 1 - \frac{\alpha^6}{72} \left( 1 - \frac{\alpha^2}{8} \right). \]  

(4)

The map is contracting if \( \alpha \leq 2\sqrt{2} \) and the maximum contraction rate, occurring for \( \alpha_+ = \sqrt{6} \), is equal to 1/4. In this case the amplitude of oscillation decreases by 1/2 at every time step. Defining \( \beta = T/\Delta t = 2\pi/\alpha \) (ratio between the period and the time step) the values of this parameter such that the stability condition is satisfied and the maximum dissipation occurs are given by

\[ \beta = \frac{T}{\Delta t} > \frac{2\pi}{2\sqrt{2}} = 2.22 \quad \beta_+ = \frac{2\pi}{\sqrt{6}} = 2.57. \]  

(5)

More generally the Runge-Kutta of order \( n \) is given by Eq. (2) where \( C(\alpha) \) and \( S(\alpha) \) are the order of truncations of the Taylor series of \( \cos(\alpha) \) and \( \sin(\alpha) \). In Fig. 1 we show the plot of the determinant \( D \) as a function of \( \beta \) for the Runge-Kutta integrators up to order six.

**B. Small oscillations of a planar heavy chain**

In order to show that the method works for a realistic problem we have considered a chain of \( N \) points of mass \( m \), joined by equal massless rods of length \( \ell \), which move in a vertical plane under a uniform gravity field. We denote with \( O \) the fixed point of the chain, where the origin of the coordinate system is chosen, and by \( P_1, P_2, \ldots, P_N \) the remaining points, see Fig. 2. The relative position and velocity vectors are denoted by

\[ \mathbf{r}_k = P_k - P_{k-1}, \quad \mathbf{v}_k = \frac{d\mathbf{r}_k}{dt}. \]  

(6)

As a consequence, the position and velocity of point \( k \) are \( P_k - O = r_1 + \ldots + r_k \) and \( dP_k/dt = v_1 + \ldots + v_k \), respectively. The Lagrangian, where the constraints are realized by hard springs (with elastic constant equal to \( k \)), reads

\[ L = \frac{m}{2} \left[ v_1^2 + (v_1 + v_2)^2 + \ldots + (v_1 + v_2 + \ldots + v_N)^2 \right] - \frac{k}{2} \left( r_1 - \ell \right)^2 - \frac{k}{2} \left( r_2 - \ell \right)^2 - \ldots - \frac{k}{2} \left( r_N - \ell \right)^2 \]  

where \( r_j = ||\mathbf{r}_j|| \). The numerical integration with a fourth order Runge-Kutta is worked out by using Cartesian coordinates. When the quadratic approximation of the Lagrangian near the equilibrium configuration is considered, the degrees of freedom corresponding to the radial and polar coordinates, which describe the rods elongations and oscillations respectively, decouple and we can write \( L = L_1 + L_2 \) where \( L_1 \) and \( L_2 \) are given in Appendix A. Letting \( Y_n = \sum_{k=1}^{n} r'_k \) where \( r'_k \) is the radial displacement with respect to the equilibrium, the Lagrangian \( L_1 \) reads

\[ L_1 = \frac{m}{2} \sum_{n=1}^{N} Y_n^2 - \frac{k}{2} \sum_{n=1}^{N} (Y_n - Y_{n-1})^2, \]  

(8)

which is the Lagrangian for \( N \) identical coupled oscillators with one fixed (\( y_0 \equiv 0 \)) and one free end. The elastic constant \( k \) of the springs is fixed by

\[ k = m\omega_0^2 = m \frac{\alpha^2}{(\Delta t)^2}, \]  

(9)

where \( \alpha = 2\pi/\beta \) is such that the Runge-Kutta integrator is stable. This condition is given by

\[ \alpha = \omega_{\text{max}} \Delta t \leq 2\sqrt{2}, \]  

(10)

just as for the single oscillator, where \( \omega_0 \) is replaced by the highest frequency of the normal modes \( \omega_{\text{max}} \). As a

---

**FIG. 1:** Determinant of the matrix defining the linear transformation of a Runge-Kutta scheme of order \( n \) as a function of the ratio \( \beta = T/\Delta t \): curve (a) \( n = 2 \), (b) \( n = 3 \), (c) \( n = 4 \), (d) \( n = 5 \) and (e) \( n = 6 \).

**FIG. 2:** Chain of four point masses with constraint forces.
consequence the range where $\omega_0$, and consequently $k$ for given $\Delta t$ can be chosen, is given by

$$\alpha \equiv \omega_0 \Delta t \leq \sqrt{2} \frac{2\omega_0}{\omega_{\text{max}}}. \quad (11)$$

The highest frequency $\omega_{\text{max}}$ of the normal modes, see Appendix A, is given by $\omega_{\text{max}} = 2\omega_0 \left(1 + O(N^{-2})\right)$ and the relative accuracy is precisely $N^{-2}$. As consequence for $N \geq 10$ the stability condition can be written

$$\alpha \leq \sqrt{2}. \quad (12)$$

We have verified the stability condition (within at least five decimal digits) for $2 \leq N \leq 1000$ by considering the ratio $E(n)/E(0)$ of the energy after $n$ steps to the initial energy. This quantity is easier to compute with respect to the determinant $D$ of the linear map defined by the fourth order Runge-Kutta algorithm; moreover for $N = 1$ we have $D^n = E(n)/E(0)$. In Fig. 3 we plot this ratio as a function of $\alpha = \omega_0 \Delta t$ for $N = 50$ uncoupled oscillators, and for a chain of identical coupled oscillators, whose Lagrangian is given by (8).

The value of $\alpha = \omega_0 \Delta t$, which gives the highest dissipation is not the same as for the single oscillator, because the energy ratio and the determinant of the linear map corresponding to the Runge-Kutta algorithms are no longer the same for $N \geq 2$. If $N = 1$ the minimum of $E(n)/E(0)$ is reached, for any $n$, at $\alpha = \sqrt{6}$. If $N = 2$ and $n = 1$ the minimum occurs at $\alpha = 1.52$ very close to $\sqrt{6}$ times the ratio $\omega_0/\omega_{\text{max}} = 0.618$. However for $n = 100$ the minimum is reached for $\alpha = 1.74$, which is close to stability threshold $\alpha = 2\sqrt{2} \omega_0/\omega_{\text{max}} = 1.748$. The minimum with respect to $\alpha$ reached by $E(n)/E(0)$ depends both on $n$ and $N$ (see Fig. 3). The dependence on $N$ is weak and the variation from $N = 10$ to $N = 100$, where the asymptotic value is basically reached, is less than 2%. The dependence on $n$ is also mild and is less than 4% from $n = 1$ to $n = 100$, where the asymptotic value is basically reached. The variation range of $\alpha$ for $N \geq 10$ and any $n \geq 1$ is within the interval $[1.3, 1.41]$ which becomes $[1.38, 1.41]$ for $n \geq 10$. As $n$ increases the value of $\alpha$ where the maximum dissipation occurs approaches the stability threshold given by (11), namely $\alpha = \sqrt{2}$ for $N$ large. A good compromise, valid for any $N \geq 10$ is $\alpha = 1.38$. This value is valid not only for the linear oscillations but also for the nonlinear oscillations of the heavy chain, up to initial elongation of 80° with zero initial velocity. A comparison with the heavy chain where the constraints are exactly realized shows that this method is fairly accurate (see Figs. 4, 5) so we believe it may be recommended to describe biopolymers.
Since for the Runge-Kutta algorithms, given by (10), reads

\[ \frac{\omega_1^2}{2} \left( 1 + \frac{\epsilon^2}{8} + O(\epsilon^4) \right), \]

and the corresponding eigenvectors are

\[ n_1 = (1, 1 + \epsilon^2/2 + O(\epsilon^4)), \]
\[ n_2 = (1, -1 + \epsilon^2/2 + O(\epsilon^4)). \]  

Since \( \Omega_2 \) is the highest frequency the stability condition for the Runge-Kutta algorithms, given by (10), reads \( \Omega_2 \Delta t \leq 2\sqrt{2} \) and becomes

\[ \alpha = \omega_2 \Delta t \leq 2, \]

for \( \epsilon \to 0 \). The highest dissipation, namely the minimum of \( E(n)/E(0) \), occurs for a value of \( \alpha \), which approaches the stability threshold for \( n \) large. The two masses undergo harmonic oscillations with frequency \( \Omega_1 \) and amplitude of order 1, whereas their distance is of order \( \epsilon^2 \). Their energy is constant up to variation of order \( \epsilon^2 \), indeed for \( t \gg \Delta t \), when the oscillations of high frequency have died out, the solution becomes \( x(t) = n_1 C \cos(\Omega_1 t + \gamma) \) namely \( x_1(t) = C \cos(\Omega_1 t + \gamma) \) and \( x_2(t) - x_1(t) = \epsilon^2/2 C \cos(\Omega_1 t + \gamma) \). Both points oscillate with an amplitude \( C \) whereas their distance oscillates with amplitude \( \epsilon^2/2 \). The energy of the system can be written as \( H = H_1 + H_2 \) where

\[ \begin{align*}
H_1 &= m \left( \frac{\dot{x}_1^2}{2} + \frac{\dot{x}_2^2}{2} \right) + \frac{k_1}{2} x_1^2, \\
H_2 &= \frac{k_2}{2} (x_2 - x_1)^2,
\end{align*} \]

and a simple calculation shows that

\[ H_1 = m (\dot{x}_1^2 + \Omega_1^2 x_1^2) + \frac{m}{4} \dot{x}_1^2 \epsilon^2 + O(\epsilon^4), \]

where the last term gives the energy exchange with the hard spring. Its sum with \( H_2 \) remains constant in time up to terms of order \( \epsilon^4 \). The energy of the system \( H_1 = \frac{1}{4} m \omega_1^2 C^2 \) corresponds to the collapse of two points into one of mass \( 2m \), the time dependent correction term of order \( \epsilon^2 \) represents the energy exchange with the hard spring and its average is equal to

\[ \langle H_2 \rangle = \frac{1}{16} m \omega_1^2 C^2 \epsilon^2 + O(\epsilon^4) = \frac{\epsilon^2}{8} H_1 + O(\epsilon^4). \]

As a consequence not only the oscillation amplitude of the hard spring is of order \( \epsilon^2 \), but also the energy exchange ratio \( \langle H_2 \rangle/H_1 \) is of order \( \epsilon^2 \). This argument extends to \( N > 2 \), where the low and high frequency spectra are sharply separated. The hard springs oscillation amplitude is proportional to \( \epsilon^2 \) and the ratio of their average energy with respect to the soft springs is also of order \( \epsilon^2 \). It appears therefore that the fourth-order Runge-Kutta scheme with a choice \( \alpha = \omega_2 \Delta t \) close to the stability threshold given by (10) is adequate to integrate linear systems in which the hard springs are coupled to soft ones, providing an efficient realization of constraints. We have checked numerically that a weakly nonlinear oscillator coupled by a hard spring has the same properties, and that the energy transfer between the slow and fast modes appears to be negligible. The optimal choice of \( \alpha \) is the same as for the linear case. This result reminds the Nekhoroshev’s like estimate for Hamiltonian systems.

### D. Thermal bath

We briefly discuss the coupling of the heavy chain (biopolymer) to a stochastic thermal bath (solvent). We have checked that the thermal bath can be introduced...
using the Langevin approach without spoiling the realization of the constraints. In particular we have verified that when the equilibrium is reached the equipartition of energy between the effective degrees of freedom \((N \text{ for a planar } 2N \text{ for a non planar chain})\) occurs with statistical fluctuations of order \(N^{-1/2}\). Let’s consider the heavy chain Lagrangian \((7)\), by denoting
\[
\mathbf{R}_k = P_k - O = r_1 + r_2 + \ldots + r_k \quad \mathbf{V}_k = \frac{d\mathbf{R}_k}{dt},
\]
we have
\[
\mathcal{L} = \frac{m}{2} (\mathbf{V}_1^2 + \ldots + \mathbf{V}_N^2) - \frac{k}{2} (\|\mathbf{R}_1\| - \ell)^2 - \frac{k}{2} (\|\mathbf{R}_2 - \mathbf{R}_1\| - \ell)^2 - \ldots - \frac{k}{2} (\|\mathbf{R}_N - \mathbf{R}_{N-1}\| - \ell)^2
\]- \mg Y_1 - \ldots - \mg Y_N.
\]
The equations of motion introducing the Wiener noise and the damping become
\[
\frac{d\mathbf{R}_k}{dt} = \mathbf{V}_k \quad \frac{d\mathbf{V}_k}{dt} = \mathbf{F}_k - \gamma \mathbf{V}_k + \epsilon \frac{d\mathbf{w}_k}{dt}, \quad \mathbf{F}_k = \frac{\partial \mathcal{L}}{\partial \mathbf{r}_k},
\]
where \(\gamma\) is the damping coefficient and \(\mathbf{w}_k(t)\) denotes a Wiener process acting on particle \(k\). The integration is performed by using a splitting method. The deterministic evolution on a time step \(\Delta t\) is carried out with a fourth order Runge-Kutta. From \(\mathbf{R}_k(t), \mathbf{V}_k(t)\) we obtain \(\mathbf{R}_k(t + \Delta t), \mathbf{V}_k(t + \Delta t)\). The Langevin process changes only the speed and gives
\[
\mathbf{V}_k(t + \Delta t) = \mathbf{V}_k(t + \Delta t) - \gamma \mathbf{V}_k(t) \Delta t + \epsilon \xi_k \sqrt{\Delta t},
\]
where
\[
k_B T = \frac{\epsilon^2}{2m\gamma},
\]
\(T\) being the temperature of the bath. We have verified in several simulations (changing \(\gamma\) and the temperature) that the system reaches its equilibrium and the average kinetic energy of every degree of freedom is \(k_B T/2\). In order to improve the algorithm we finally notice that we can chose two different integration steps for the deterministic (Runge-Kutta) and the stochastic (Langevin) part of the dynamics, \(\Delta t_{RK} < \Delta t_{Langevin} = n \Delta t_{RK}\). This choice helps the reduction of the residual energy in the hard springs. Taking for instance \(n \sim 10\) the equipartition at equilibrium is still well verified with a very small amount of residual energy in the springs.

III. EXACT CONSTRAINTS

A chain of \(N\) point masses joined by \(N - 1\) identical massless rods has the same geodesic flow as a system of \(N\) spherical pendula coupled by the generalized forces, obtained by projecting on the corresponding tangent spaces the constraint forces. These forces are directed along the rods and their intensity, depending on the dynamical state of the system, are determined by solving a tridigonal system. This is a special case of linear system which determines the rigid constraints forces \([8]\). The equations of motion of each spherical pendulum are solved by using local coordinates on two different charts, in order to avoid the singularity. Denoting the positions and velocities of the particles according to \((6)\) the Lagrangian of the chain is given by
\[
L = \frac{1}{2} \left[ v_1^2 + (v_1 + v_2)^2 + \ldots + (v_1 + v_2 + \ldots + v_N)^2 \right] - V(r_1, \ldots, r_N),
\]
where \(V\) is the potential of the active forces, internal and external. After the introduction of polar coordinates according to
\[
\mathbf{r}_k = \sin \theta_k \mathbf{e}_x + \sin \phi_k \mathbf{e}_z, \quad \mathbf{v}_k = \dot{\theta}_k \mathbf{e}_x + \sin \phi_k \mathbf{e}_y - \sin \theta_k \mathbf{e}_z, \quad \mathbf{F}_k = \frac{\partial L}{\partial \mathbf{r}_k},
\]
the writing of Lagrange’s equations appears to be a cumbersome task also with the help of an algebraic manipulator, and becomes practically prohibitive for \(N \geq 5\). Having chosen \(m = 1\), the matrix \(T\) of the quadratic form defining the kinetic energy is the metric tensor. Its determinant vanishes when any of the points crosses the \(z\) axis \((\theta_k = 0, \pi)\). For \(N \leq 4\) we have verified that
\[
\det(T) = \sin^2 \theta_1 \ldots \sin^2 \theta_N f(\phi, \theta).
\]
This singularity appears when Lagrange’s equations are written in normal form.

A. The spherical pendulum

In this case the Lagrangian reads
\[
L = \frac{1}{2} (\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta) - V(\theta, \phi),
\]
and the equations of motion, written in normal form, are
\[
\ddot{\theta} = \sin \theta \cos \theta \dot{\phi}^2 - \frac{\partial V}{\partial \theta}, \quad \ddot{\phi} = \frac{1}{\sin^2 \theta} \left( 2 \dot{\theta} \dot{\phi} \sin \theta \cos \theta - \frac{\partial V}{\partial \phi} \right).
\]

B. The flipping rule

The strategy to integrate Lagrange’s equation is to define two frames \(x, y, z\) and \(x', y', z'\) obtained from the first one is obtained by two subsequent rotations of \(\pi/2\)
C. Intrinsic method

A geodesic arc on the sphere belongs to a maximum circle obtained by intersecting the sphere with the plane to which the vectors \( r(t) \) and \( v(t) \) belong. As a consequence, the vectors \( r(t + \Delta t) \) and \( v(t + \Delta t) \) are obtained by rotating counterclockwise around the normal \( n = (r \times v) ||r \times v||^{-1} \):

\[
\begin{align*}
    r' &= r \cos \alpha + v \sin \alpha, \\
    v' &= -r \sin \alpha + v \cos \alpha.
\end{align*}
\]  

(35)  

(36)

When a potential is present, the evolution in a time interval \( \Delta t \) is obtained by composing the geodesic flow with the flow determined by the potential only (splitting method).

D. The constraints

It is possible to determine the constraints by solving a linear system as shown in [8]. For the spherical pendulum the equation of motion is

\[
\dot{v} = F + \Phi, \quad \Phi = \lambda r,
\]  

(37)

where \( \Phi \) is the constraint force. The projection into the tangent space gives Lagrange’s equation freed from the constraint force. Nevertheless we can determine the Lagrange multiplier \( \lambda \) by multiplying Eq. (37) by \( r \) and taking into account that from \( r \cdot v = 0 \) follows \( r \cdot dv/dt + v^2 = 0 \) which implies \( \lambda = -v^2 - r \cdot F \). Replacing into Eq. (37) and denoting with \( n = -r \) the normal we have

\[
\dot{v} = F - n (n \cdot F) + v^2 n,
\]  

(38)

which expresses the equality of the force and acceleration in the tangent plane. In the case of a chain with \( N \geq 2 \) points of unit mass, linked by massless rods of unit length, we have (Fig. 2)

\[
\frac{d}{dt} v_1 = F_1 + \Phi_1 - \Phi_2,
\]  

\[
\frac{d}{dt} (v_1 + v_2) = F_2 + \Phi_2 - \Phi_3,
\]  

\[
\vdots
\]  

\[
\frac{d}{dt} \left( v_1 + v_2 + \ldots + v_{N-1} \right) = F_{N-1} + \Phi_{N-1} - \Phi_N,
\]  

\[
\frac{d}{dt} \left( v_1 + v_2 + \ldots + v_N \right) = F_N + \Phi_N,
\]  

(39)

where \( \Phi_k = \lambda_k r_k \) for \( k = 1, \ldots, N \). Subtracting to each equation the previous one we obtain

\[
\dot{v}_k = F_k - F_{k-1} - \lambda_{k-1} r_{k-1} + 2\lambda_k r_k - \lambda_{k+1} r_{k+1},
\]  

(40)
with the exact constraint force and a hard spring, the blue curves to the equations of motion with the exact constraint force, the red curves to the equations with the exact constraint force and a hard spring, the blue curves to the hard spring only. The elastic constant of the hard spring is chosen according to (5).

\[\dot{\theta}_k = \frac{\theta_k^2}{2} \sin \theta_k \cos \theta_k - \frac{\partial r_k}{\partial \theta_k} \cdot (r_{k-1} \lambda_{k-1} + r_{k+1} \lambda_{k+1}) + \frac{\partial r_k}{\partial \theta_k} \cdot (F_k - F_{k-1})\]

This is a tridiagonal system determining the constraints \(\lambda_k\). We use polar coordinates and project equation (40) into the tangent space of the \(k\)-th pendulum scalar multiplying it by \(\partial r_k/\partial \theta_k\) and \(\partial r_k/\partial \phi_k\). The scalar product of the constraint forces is non zero except for \(\lambda_k r_k\). We have verified for \(N = 2, 3\) that the result is the same as Lagrange’s equations obtained from Eq. (25). The equations of motion finally read

In this case, to avoid the singularities, we must choose one among \(2^N\) charts at every time step. The choice for the \(k\)-th pendulum is given by the initial or the flipped coordinates depending on \(\sin \theta_k > 1/\sqrt{2}\) or \(\sin \theta_k' > 1/\sqrt{2}\).

E. The heavy chain

We have first compared the results obtained by using a fourth order Runge-Kutta to integrate the equations for the spherical pendulum taking into account the constraint force, see equation (38) and the hard spring. In the absence of the spring the error rises exponentially.

where \(\lambda_0 = \lambda_{N+1} = 0\). Scalar multiplying each equation by \(r_k\) and recalling that \(r_k \cdot v_k = -v_k^2\) we obtain

\[mu_k^2 + r_k \cdot (F_k - F_{k-1}) = \lambda_{k-1} r_k \cdot r_{k-1} - 2\lambda_k + \lambda_{k+1} r_k \cdot r_{k+1}.\]  

(41)

After inserting the hard spring and choosing the elastic constant as discussed in section II, the accuracy with or without the constraint forces is about the same, see Fig. 8. Similar plots are obtained by considering the errors on the pendulum length or on the orthogonality condition \(r \cdot v = 0\). For the heavy chain this is still true and in Fig. 4 we plot the energy fluctuations for \(N = 30\) point masses. The integration of Lagrange’s equations using spherical coordinates with respect to two different frames is much more accurate for the same time step as shown in Fig. 9 (to compare with Fig. 4), but requires the determination of the constraint forces. We have also evaluated the growth of the distance between these orbits and the orbits obtained with the hard springs but due to the presence of positive Lyapunov exponents this information is not very significant except for small oscillation amplitudes.

IV. THE CHAIN OF RIGID BODIES

We assume that any of the \(N\) rigid bodies forming the chain is a set of points, whose distances are fixed. The rigid body may be considered as a set of \(n_B\) point masses with rigidity constraints, which can be realized by hard springs. However it is convenient to treat it as a dynamical system with 6 degrees of freedom, as long as \(n_B \geq 3\). The Lagrangian parameters, which determine the body frame, are usually the Euler angles and the and center of mass coordinates, even through other choices are possible (Deprit’s rather than Euler angles). Since the determination of the constraint force corresponding to the pivot requires an approximation procedure [9], we realize this constraint by a hard spring. If the first body has a fixed
However, the melting temperatures of the resulting two kinds of bond are different; as a consequence, when the rigidity constraints, to simulate an ensemble of interacting parallelogram-shaped nanostructures composed of DNA strands [12]. These parallelograms are designed in such a way that corresponding ends may bind together. The proposed method has been tested in the case of a heavy chain of 30 rigid bodies. In Fig. 10 the energy fluctuations for a chain with \( N = 30 \) are shown and the typical structure of the rigid bodies is shown in Fig. 11 for \( N = 5 \). The method appears to be accurate and robust and an application to peptidic chains is possible.

A. The heavy chain

The proposed method has been tested in the case of a heavy chain of \( N \) rigid bodies. In Fig. 10 the energy fluctuations for a chain with \( N = 30 \) are shown and the typical structure of the rigid bodies is shown in Fig. 11 for \( N = 5 \). The method appears to be accurate and robust and an application to peptidic chains is possible.

B. A model of interacting DNA nanostructures

We applied our method, concerning the realization of rigidity constraints, to simulate an ensemble of interacting parallelogram-shaped nanostructures composed of DNA strands [12]. These parallelograms are designed in such a way that corresponding ends may bind together. However, the melting temperatures of the resulting two kinds of bond are different; as a consequence, when the system is cooled in an appropriate way, this allows the formation of two types of assembly, that is rods and donuts (see Fig. 12).

In this dynamical model each parallelogram is represented by a collection of point masses connected by hard springs. In a similar way, even the bonds among the various parallelograms are realized by means of hard springs; what is interesting in this case is that such bonds may be formed and destroyed during the simulation, according to the environmental conditions. To each mass point is given a shielded Coulomb potential which reads

\[
V(r) = \exp\left(\frac{(r_0 - r)}{a}\right) \left( r - r_0 \right) \quad (a \ll r_0),
\]

so that it behaves like a sort of hard sphere of hard core \( r_0 \) with respect to all the other points and to the container. Of course, in order to avoid stability issues, the time step \( \Delta t \) has to verify the condition: \( \langle v \rangle \Delta t \ll a \), where \( \langle v \rangle \) represents the average speed. A thermal bath is introduced in order to control the temperature of the
nanostructures) and of the thermal bath. As a final remark, by “switching off” the dissipative effects due to the parallelograms, it is possible to analyse the conservation of the energy. Generally this is really good; for example, using a time step $\Delta t = 10^{-3}$, the relative error in the conservation of the energy is about $10^{-4}$.

V. CONCLUSIONS

We have proposed a numerical scheme to integrate the equations of motion of a chain of point masses and rigid bodies. The key point is the realization of the constraints with rigid springs, whose energy is rapidly dissipated, with a suitable choice of the time step of the fourth order Runge-Kutta integrator. The time step is such that the rapid oscillations are not resolved and its optimal value is very close to the instability threshold of the Runge-Kutta integrator for the linearized equations of motion. The condition on the time step is given by $\Delta t \omega_{\text{max}} \leq 2\sqrt{2}$ where $\omega_{\text{max}}$ is the highest frequency of the normal modes. By allowing a flip of the reference frame at each time step we avoid in an optimal way the singularity associated to the use of local coordinates. For the chain of point masses we have compared the hard springs solution with the solution of Lagrange’s equation in spherical coordinates for two flipping reference frames. Compared with the standard procedures where the constraints are determined by solving a linear system a comparable accuracy is reached avoiding the determination of the constraints. The method has been used in a model of interacting DNA fragments, where the links can be dynamically generated and destroyed.

APPENDIX A: SMALL OSCILLATIONS FOR THE PLANAR CHAIN

Using polar coordinates $x_n = r_n \sin \phi_n$, and $y_n = -r_n \cos \phi_n$ the stable equilibrium conditions are

$$\phi_n = 0 \quad r_n = r^*_n \equiv \ell + (N+1-n)\frac{q}{k} \quad n = 1, \ldots, N.$$  

(A1)

Denoting by $r'_n = r_n - r^*_n$ the displacement from the equilibrium distance we have $x_n = r'_n \phi_n + r^*_n \phi_n$ and $y_n = -r'_n - r^*_n + \frac{1}{2} r^*_n^2 \phi^2_n$ up to third order terms and the small oscillation Lagrangian becomes $L_{\text{sm.osc.}} = L_1 + L_2$ where

$$L_1 = \frac{m}{2} \sum_{n=1}^{N} \left( \sum_{j=1}^{n} r'_j \right)^2 - \frac{k}{2} \sum_{n=1}^{N} r_n^2.$$  

(A2)

and

$$L_2 = \frac{m}{2} \sum_{n=1}^{N} \left( \sum_{j=1}^{n} r'_j \phi_j \right)^2 - \frac{mg}{2} \sum_{n=1}^{N} \sum_{j=1}^{n} r'_j \phi^2_j.$$  

(A3)

The Lagrangian for the radial displacements is the same as for a one dimensional chain of points of mass $m$ connected by identical springs with elastic constant $k$. Using
function of $N$ red, the eigenvalues of the matrix $T$ plotted for different values of $N = 10$ red, $N = 25$ blue, $N = 50$ green, $N = 100$ purple as a function of $n/N$.

\[ L_2 = \frac{m}{2} \dot{r}^2 + \frac{k}{2} r^2 - \frac{k}{2} r^2 \quad T_{ij} = N+1-i \quad j \leq i \quad T_{ji} = T_{ij}. \tag{A4} \]

Denoting by $\lambda_n$ the eigenvalues of the matrix $T$ the frequencies $\omega_n$ for the radial normal modes are

\[ \omega_n = \left( \frac{k}{m} \right)^{1/2} \frac{1}{\sqrt{\lambda_n}}. \tag{A5} \]

As shown in the figure 14 the $\lambda_n^{-1/2} \leq 2$ and the accuracy with which the upper limit is reached is $N^{-2}$. The Lagrangian for the transverse oscillations, is the same one obtained when the the constraints $r_j = \ell$ are exactly realized if we set $r_n = \ell$, which is exact in the limit $k \to \infty$ and reads

\[ L_1 = \frac{m}{} \dot{\theta} \cdot T \dot{\phi} - \frac{mg}{2} \phi \cdot V \phi \quad V_{ij} = (N + 1 - i) \delta_{ij}. \tag{A6} \]

The frequencies are given by $\omega_n = (g/\ell)^{1/2} \lambda_n^{-1/2}$ where $\lambda_n$ are the eigenvalues of the matrix $L = V^{-1/2} TV^{-1/2}$ whose elements are

\[ L_{ij} = (N + 1 - i)^{1/2} (N + 1 - j)^{-1/2} \quad j \leq i \quad L_{ji} = L_{ij}. \tag{A7} \]

\section*{APPENDIX B: REFERENCE FRAMES FOR THE RIGID BODY}

Letting $(e_x, e_y, e_z)$ and $(e_{x}^{body}, e_{y}^{body}, e_{z}^{body})$ be the unit vectors along the axis of the fixed frame and body frame, chosen along the principal axis of inertia, Euler angles, see Fig. 15, are defined by

\[ \begin{align*}
\cos \theta &= e_z \cdot e_{z}^{body}, \\
\cos \phi &= e_x \cdot e_{y}^{body}, \\
\cos \psi &= e_x \cdot e_{z}^{body},
\end{align*} \tag{B1} \]

where

\[ e_L = \frac{e_x \times e_{z}^{body}}{\|e_x \times e_{z}^{body}\|}. \tag{B2} \]

The inverse rotation matrix $R^{-1}$, whose columns are the components of body frames unit vectors, is defined by

\[ R = R_x(\psi) R_y(\theta) R_z(\phi), \quad R^{-1} = (e_{x}^{body}, e_{y}^{body}, e_{z}^{body}), \tag{B3} \]

where $R_x(\alpha), R_y(\alpha), R_z(\alpha)$ denote the clockwise rotation matrices of an angle $\alpha$ around the coordinate axis. The velocity of a point $P$ of the body identified by the vector $r = P - O$ is $v = \omega \times r$ where the components $(\omega_{x}^{body}, \omega_{y}^{body}, \omega_{z}^{body})$ of the angular velocity $\omega$ on the body frame are

\[ \omega = \dot{\phi} e_z + \psi e_{z}^{body} + \dot{\theta} e_L = R_z(\psi) \begin{pmatrix} \dot{\phi} \\ \phi \sin \theta \\ \dot{\psi} + \dot{\phi} \cos \theta \end{pmatrix}. \tag{B4} \]

In order to obtain the Euler angles from the components of the unit vectors in the body frame, namely from the matrix $R_{ij} = e_{j} \cdot e_{i}^{body}$, we observe that

\[ \begin{align*}
e_{z}^{body} \cdot e_x &= \sin \theta \sin \phi, \\
e_{z}^{body} \cdot e_y &= -\sin \theta \cos \phi, \\
e_{z}^{body} \cdot e_z &= \cos \theta,
\end{align*} \tag{B5} \]
\[ \mathbf{e}_z \cdot \mathbf{e}_y^{\text{body}} = \sin \theta \sin \psi, \]
\[ \mathbf{e}_z \cdot \mathbf{e}_y^{\text{body}} = \sin \theta \cos \phi, \]
\[ \mathbf{e}_z \cdot \mathbf{e}_x^{\text{body}} = \cos \theta. \]

The inverse relations to obtain the Euler angles are given by
\[ \theta = \arccos(\mathbf{e}_z \cdot \mathbf{e}_x^{\text{body}}) = \arccos R_{31}^{-1}, \]
\[ \phi = \text{phase} \left( \frac{\mathbf{e}_z \cdot \mathbf{e}_y^{\text{body}}}{\sin \theta}, \frac{\mathbf{e}_z \cdot \mathbf{e}_x^{\text{body}}}{\sin \theta} \right) \equiv \text{phase} \left( \frac{R_{32}^{-1} R_{31}^{-1}}{\sin \theta}, \frac{R_{31}^{-1}}{\sin \theta} \right), \]
\[ \psi = \text{phase} \left( \frac{\mathbf{e}_z \cdot \mathbf{e}_y^{\text{body}}}{\sin \theta}, \frac{\mathbf{e}_z \cdot \mathbf{e}_x^{\text{body}}}{\sin \theta} \right) \equiv \text{phase} \left( \frac{R_{32}^{-1}}{\sin \theta}, \frac{R_{31}^{-1}}{\sin \theta} \right). \]

1. The flipping rule

As for the spherical pendulum, this operation consists in exchanging the fixed frame axis according to Eq. (30), namely
\[ \mathbf{e}_x^{\text{body}}' = \mathbf{e}_z^{\text{body}}, \]
\[ \mathbf{e}_y^{\text{body}}' = \mathbf{e}_y^{\text{body}}, \]
\[ \mathbf{e}_z^{\text{body}}' = \mathbf{e}_z^{\text{body}}. \]

As a consequence the rotation matrix \( R'^{-1} \) is obtained from \( R^{-1} \) by interchanging the rows
\[ R'^{-1}_{1k} = R_{2k}^{-1}, \]
\[ R'^{-1}_{2k} = R_{3k}^{-1}, \]
\[ R'^{-1}_{3k} = R_{1k}^{-1}. \]

The new Euler angles \( (\phi', \theta', \psi') \) are then obtained from \( R'^{-1} \) by the inversion formulas (B7) written for the primed quantities. The same permutation is performed on the components of angular velocity, of forces and of their momentum, written in the fixed frame. We first transform the vectors written in the body frame to the fixed frame by applying \( R^{-1} \), perform the permutation and transform them back to the body frame.

2. Intrinsic methods

Intrinsic methods based on the quaternionic representation of the rotation of a Cartesian frame have been extensively considered in the literature, to which we refer. The accuracy and reliability of our method are quite satisfactory.

**APPENDIX C: RIGID BODY DYNAMICS AND QUATERNIONS**

We construct an explicit numerical scheme of order three for the integration of the rigid body dynamics with a fixed point based on the quaternion algebra [11]. As it is well known, the rigid body configuration space is the rotation group \( SO(3) \), that can be represented by the group of unitary quaternions. Moreover we take advantage of the isomorphism between the purely imaginary quaternion and the 3-dimensional vectors. Let \( \mathcal{R}(\hat{\omega}, \alpha) \in SO(3) \) a rotation around the \( \omega \) axis of an angle \( \alpha \), we have the homomorphism
\[ \mathcal{R}(\hat{\omega}, \alpha) \mathbf{v} \to e^{\omega \alpha/2} \mathbf{v} e^{-\omega \alpha/2}, \]
where \( \mathbf{v} \) and \( \omega \) are, at the same time, vectors in \( \mathbb{R}^3 \) and purely imaginary quaternions. It is straightforward to see that
\[ e^{\omega \alpha/2} = \cos \alpha/2 + \omega \sin \alpha/2, \]
due to the identity \( \omega^2 = -1 \). The rigid body configuration at any time \( t \) is defined by a unitary quaternion \( w(t) \) that is associated to the rotation from the fixed frame to the body frame
\[ \mathbf{r}(t) = \overline{w}(t) \mathbf{r}_0 w(t), \]
where \( \mathbf{r}_0 \) is the initial position of generic point of the body and \( \overline{w} \) is the conjugated quaternion. According to the correspondence (C1), by deriving \( w \) we get
\[ \dot{w} = -\frac{\omega}{2} w, \]
where \( \omega \) is the angular velocity in the body system. In order to write the equations of motion we consider the Euler equations for the angular momentum \( \mathbf{L} \) in the body systems of principal inertial axes:
\[ \dot{\mathbf{L}} = -\mathbf{\Omega} \times \mathbf{L} + \mathbf{\Omega}, \]
where \( \mathbf{\Omega} \) is the momentum of the external forces in the body system. The equations of motion (C4) and (C5) are completed by the relation \( \mathbf{L} = J \omega \) where \( J \) is the inertia tensor. The external forces momentum \( \mathbf{\Omega} \) in the body system can be easily computed by using the relation (C3) from the momentum in the space system. Finally the equations of motion can be written in the generic form
\[ \dot{\mathbf{w}} = -\frac{\omega}{2} \mathbf{w}, \]
\[ \dot{\omega} = \mathbf{f}(\mathbf{w}, \overline{\mathbf{w}}, \omega). \]

We look for a simple integration scheme of Eq. (C6) that preserves the Lie group structure \( \omega \overline{\mathbf{w}} = 1 \). A scheme of
order $O(\Delta t^3)$ can be computed in the form

\[ w(t + \Delta t) = \exp \left[ -\frac{\Delta t}{6} \omega(t + \Delta t) \right] \times \exp \left[ -\frac{\Delta t}{3} \omega(t + \Delta t/4) \right] w(t), \]

\[ \omega(t + \Delta t) = \omega(t) + \frac{3}{4} f \left( t + \frac{2}{3} \Delta t \right) \Delta t + \frac{1}{4} f(t) \Delta t, \quad (C7) \]

where we use the notation

\[ f(s) = f(w(s), \overline{w}(s) \omega(s)), \quad (C8) \]

and we compute the various quantities according to the schemes

\[ w \left( t + \frac{2}{3} \Delta t \right) = \exp \left[ -\frac{1}{3} \omega \left( t + \frac{\Delta t}{3} \right) \Delta t \right] w(t), \]

\[ \omega \left( t + \frac{2}{3} \Delta t \right) = \omega(t) + \frac{2}{3} f \left( t + \frac{\Delta t}{3} \right) \Delta t, \quad (C9) \]

A direct calculation that compares the integration scheme (C8) with the Taylor expansion of the exact solution, gives an error $O(\Delta t^4)$ for a single time step. Each time step requires three evaluations of the force field, that determines the computational complexity of the algorithm. Finally we remark that the previous procedure cannot be generalized to get an explicit integration scheme of fourth-th order.