J. Phys.: Condens. Matter **20** (2008) 244133 (24pp) doi:10.1088/0953-8984/20/24/244133

Minimal distance transformations between links and polymers: principles and examples

Ali R Mohazab and Steven S Plotkin

Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T1Z1, Canada

E-mail: steve@physics.ubc.ca

Received 24 November 2007, in final form 25 February 2008
Published 29 May 2008
Online at stacks.iop.org/JPhysCM/20/244133

Abstract
The calculation of Euclidean distance between points is generalized to one-dimensional objects such as strings or polymers. Necessary and sufficient conditions for the minimal transformation between two polymer configurations are derived. Transformations consist of piecewise rotations and translations subject to Weierstrass–Erdmann corner conditions. Numerous examples are given for the special cases of one and two links. The transition to a large number of links is investigated, where the distance converges to the polymer length times the mean root square distance (MRSD) between polymer configurations, assuming that curvature and non-crossing constraints can be neglected. Applications of this metric to protein folding are investigated. Potential applications are also discussed for structural alignment problems such as pharmacophore identification, and inverse kinematic problems in motor learning and control.

1. Introduction

The standard variational definition of distance can be generalized to higher dimensional objects such as strings or membranes. In a previous paper [1], one of us has introduced the formalism for this calculation. Consider first zero-dimensional objects (points). The distance between two points A and B is defined through a transformation that takes A to B, an object of dimension one higher than the points themselves (here one dimensional). The transformation minimizing the arc-length travelled between A and B gives the scalar distance $\mathcal{D}^*$. The differential increment of arc-length may be defined as either $\sqrt{1+(\mathrm{d}y/\mathrm{d}x)^2+(\mathrm{d}z/\mathrm{d}x)^2}\,\mathrm{d}x$, or without the assumption that y, z are functions of x, parametrically. To be specific, introduce a 'time' parameter t such that $0 \leqslant t \leqslant T$, and $\mathbf{r}(0) = \mathbf{r}_\mathrm{A}$, $\mathbf{r}(T) = \mathbf{r}_\mathrm{B}$, and $\mathbf{r}(t) = (x(t), y(t), z(t))$. The distance between $\mathbf{r}_\mathrm{A}$ and $\mathbf{r}_\mathrm{B}$ can be found variationally [2]:

$$\mathcal{D}^* = \mathcal{D}\left[\mathbf{r}^*(t)\right] \qquad \text{where } \mathbf{r}^*(t) \text{ satisfies} \tag{1a}$$

$$\delta \int_0^T \mathrm{d}t \left(g_{\mu\nu}\dot{x}^\mu(t)\dot{x}^\nu(t)\right)^{1/2} = 0 \tag{1b}$$

$$\text{or} \qquad \delta \int_0^T \mathrm{d}t \sqrt{\dot{\mathbf{r}}^2} = 0 \text{ (Euclidean metric).} \tag{1c}$$

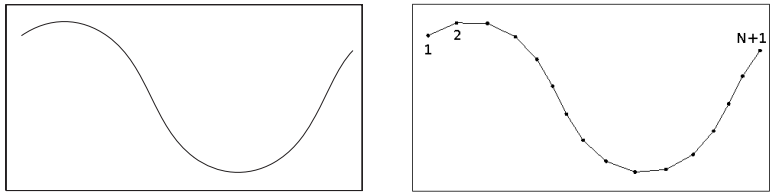
Here we have let $\dot{x} \equiv \mathrm{d}x/\mathrm{d}t$, and $\dot{\mathbf{r}} \equiv \mathrm{d}\mathbf{r}/\mathrm{d}t$. The boundary conditions on the extremal path are $\mathbf{r}^*(0) = \mathbf{r}_\mathrm{A}$ and $\mathbf{r}^*(T) = \mathbf{r}_\mathrm{B}$.

Taking the functional derivative in equation (1*c*) gives Euler–Lagrange (EL) equations for the Lagrangian $\mathcal{L} = \sqrt{\dot{\mathbf{r}}^2}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}}\right) = 0 \qquad \text{or} \qquad \dot{\hat{\mathbf{v}}} = 0 \tag{2}$$

with $\hat{\mathbf{v}}$ the unit vector in the direction of the velocity.

Since the derivative of a unit vector is always orthogonal to this vector, equation (2) says that the direction of the velocity cannot change, and therefore straight line motion results. Applying the boundary conditions gives $\hat{\mathbf{v}} = (\mathbf{r}_\mathrm{B} - \mathbf{r}_\mathrm{A})/|\mathbf{r}_\mathrm{B} - \mathbf{r}_\mathrm{A}|$. However, *any* function $\mathbf{v}(t) = |\nu_o(t)|\hat{\mathbf{v}}$ satisfying the boundary conditions is a solution, so long as $\int_0^T \mathrm{d}t|\nu_o(t)| = |\mathbf{r}_\mathrm{B} - \mathbf{r}_\mathrm{A}|$. The solution is reparametrization invariant. Then



a b

Figure 1. Continuum (a) and discretized (b) polymer chains. The EL equation for the continuum polymer is a nonlinear (vector) PDE, while the EL equations for the discretized polymer are a set of nonlinear ODEs.

the extremal functional $\mathbf{r}^*(t)$ is given by

$$\mathbf{r}^*(t) = \mathbf{r}_\mathrm{A} + \frac{\mathbf{r}_\mathrm{B} - \mathbf{r}_\mathrm{A}}{|\mathbf{r}_\mathrm{B} - \mathbf{r}_\mathrm{A}|} \int_0^t \mathrm{d}t\, |v_o(t)| \tag{3}$$

and the distance by

$$\mathcal{D}^* = \int_0^T \mathrm{d}t\, \sqrt{\dot{\mathbf{r}}^{*2}} = \int_0^T \mathrm{d}t\, |v_o(t)| = |\mathbf{r}_\mathrm{B} - \mathbf{r}_\mathrm{A}| \tag{4}$$

which represents the diagonal of a hypercube, as expected. At this point we could fix the parametrization by choosing $|v_o(t)| = |\mathbf{r}_\mathrm{B} - \mathbf{r}_\mathrm{A}|/T$ (constant speed), for example.

The extremal transformation (3) is also a minimum. In section 2.4 we will give the sufficient conditions for an extremum to be a (local) minimum, where we will return to this example.

The above idea can be generalized to space curves, surfaces, or higher dimensional manifolds [1]. The distance is defined through the transformation between the objects that minimizes the cumulative amount of arc-length travelled by all parts of the manifold.

2. Distance for polymers or strings

Describing the transformation $\mathbf{r}(s,t)$ between two space curves $\mathbf{r}_\mathrm{A}(s)$ and $\mathbf{r}_\mathrm{B}(s)$ requires two scalar parameters: s the arc-length *along the space curve*, and t the 'time' as in the above zero-dimensional case measuring progress during the transformation. The boundary conditions are then $\mathbf{r}(s,0) = \mathbf{r}_\mathrm{A}(s)$ and $\mathbf{r}(s,T) = \mathbf{r}_\mathrm{B}(s)$. The minimal transformation $\mathbf{r}^*(s,t)$ is an object of dimension one higher than A or B, i.e. it yields a distance that is two dimensional.

The distance $\mathcal{D}^* = \mathcal{D}[\mathbf{r}^*(s,t)]$, where the functional $\mathcal{D}[\mathbf{r}]$ is given by

$$\mathcal{D}[\mathbf{r}] = \int_0^L \mathrm{d}s \int_0^T \mathrm{d}t\, \sqrt{\dot{\mathbf{r}}^2}. \tag{5}$$

Here we have used the shorthand $\mathbf{r} \equiv \mathbf{r}(s,t) = (x(s,t), y(s,t), z(s,t))$ (a 3-vector), and $\dot{\mathbf{r}} \equiv \partial\mathbf{r}/\partial t$.

It has been shown previously that the problem of distance does not map to a simple soap film, nor to the minimal area of a world-sheet (which corresponds to the action of a classical relativistic string) [1].

Formulated as above, the string can contract and expand arbitrarily in order to minimize the distance travelled. The transforming object is akin to a rubber band, and all points on $\mathbf{r}_\mathrm{A}(s)$ will move in straight lines to their partner points on $\mathbf{r}_\mathrm{B}(s)$ to minimize the distance. It is worth mentioning that protein chains for example only change their length by about one per cent at biological temperatures. To accurately represent the transformation of a non-extensible string, a Lagrange multiplier $\lambda(s,t)$ must be introduced into the effective Lagrangian, weighting the constraint:

$$\sqrt{\mathbf{r}'^2} = 1, \tag{6}$$

where $\mathbf{r}' \equiv \partial\mathbf{r}/\partial s$.

Under this constraint, points along the string can no longer move independently of each other, but must always be a fixed (infinitesimal) distance apart. The tangent vector $\hat{\mathbf{t}} = \mathbf{r}'$ is now a unit vector, and the total length of the string is $L = \int_0^L \mathrm{d}s\, \sqrt{\mathbf{r}'^2} = \int_0^L \mathrm{d}s$.

Consider the minimal distance transformation between two configurations $\mathbf{r}_\mathrm{A}(s)$ and $\mathbf{r}_\mathrm{B}(s)$ of an ideal polymer of length L. Let us derive the EL equations for this case. From equations (5) and (6), the effective action is

$$\mathcal{D} = \int_0^L \int_0^T \mathrm{d}s\, \mathrm{d}t\, \mathcal{L}\left(\dot{\mathbf{r}}, \mathbf{r}'\right) \tag{7a}$$

$$\text{where } \mathcal{L} = \sqrt{\dot{\mathbf{r}}^2} - \lambda\left(\sqrt{\mathbf{r}'^2} - 1\right) \tag{7b}$$

and the Lagrange multiplier $\lambda \equiv \lambda(s,t)$ is a function of both s and t. The extrema of the distance functional $\mathcal{D}$ in (7*a*) are found from $\delta\mathcal{D} = 0$. Taking the functional derivative gives EL equations [1]:

$$\dot{\hat{\mathbf{v}}} = \lambda\boldsymbol{\kappa} + \lambda'\hat{\mathbf{t}} \tag{8}$$

where $\hat{\mathbf{v}}$ is the unit velocity vector, $\hat{\mathbf{t}}$ is the unit tangent vector, and $\boldsymbol{\kappa}$ is the curvature vector. In equation (8) we see explicitly that if the non-extensibility constraint is set to zero ($\lambda = 0$) all points on $\mathbf{r}_\mathrm{A}(s)$ move in straight lines to $\mathbf{r}_\mathrm{B}(s)$.

2.1. Discrete chains

To make the problem more amenable to solution, we can discretize the spatial variables while letting the time variable remain continuous, i.e. we implement the method of lines to solve equation (8). Rather than directly discretizing equation (8), however, it is more natural to consider a discretized chain as shown in figure 1 from the outset, and to calculate the EL equations for this system. This recipe then gives the same result as properly discretizing equation (8). For

the discretized chain, the constraint in equation (6) becomes $|\Delta\mathbf{r}| = \Delta s = L/(N-1)$, giving the length of each link. As the number of beads $N \to \infty$ the system approaches a continuous chain. For finite N, the Lagrangian becomes a function of the positions and velocities $\{\mathbf{r}_i, \dot{\mathbf{r}}_i\}$ of all beads i, $1 \leqslant i \leqslant N+1$. We use the shorthand notation $\mathcal{L}(\mathbf{r}_i, \dot{\mathbf{r}}_i)$.

This recipe yields the distance metric for an ideal, freely jointed chain, which has no non-local interactions and no curvature constraints. While this approximation is often used as a first step, real chains may behave quite differently for several reasons. In many cases, the configuration which is an energetic minimum is a straight line, or a single conformation dictated by the chemistry of the polymeric bonds. At finite temperature, energy in the bath induces conformational fluctuations. Real polymers also cannot cross themselves, and because of their stereochemistry also take up volume. We leave these interesting features for later analysis.

Equation (6) for the discretized chain becomes N constraint equations added to the effective Lagrangian:

$$\sum_{i=1}^{N} \hat{\lambda}_{i,i+1} \left(\sqrt{(\mathbf{r}_{i+1} - \mathbf{r}_i)^2} - \Delta s \right)$$

where each $\hat{\lambda}_{i,i+1} \equiv \hat{\lambda}_{i,i+1}(t)$ is a function of t, and $\hat{\lambda}_{N,N+1} = 0$. Letting $\lambda \equiv 2\hat{\lambda}\,\Delta s$ and $\mathbf{r}_{i+1/i} \equiv \mathbf{r}_{i+1} - \mathbf{r}_i$ we rewrite this strictly for convenience as

$$\sum \frac{\lambda_{i,i+1}}{2} \left(\frac{\mathbf{r}^2_{i+1/i}}{\Delta s^2} - 1 \right).$$

We next convert to dimensionless variables by letting $\mathbf{r} = (\Delta s)\hat{\mathbf{r}}$. To simplify the notation, from here on we simply refer to $\hat{\mathbf{r}}$ as $\mathbf{r}$. The distance for the discretized chain becomes

$$\mathcal{D}[\mathbf{r}_i, \dot{\mathbf{r}}_i] = \Delta s^2 \int_0^T \mathrm{d}t\, \mathcal{L}\,(\mathbf{r}_i, \dot{\mathbf{r}}_i) \tag{9}$$

with effective Lagrangian

$$\mathcal{L}\,(\mathbf{r}_i, \dot{\mathbf{r}}_i) = \sum_{i=1}^{N} \left(\sqrt{\dot{\mathbf{r}}_i^2} - \frac{\lambda_{i,i+1}}{2} \left(\mathbf{r}^2_{i+1/i} - 1 \right) \right). \tag{10}$$

The derivatives $\dot{\mathbf{r}}$ and $\mathbf{r}_{i+1/i}$ are raised to different powers in (10); however, so long as $\mathbf{r}_{i+1/i}$ satisfies the constraint $|\mathbf{r}_{i+1/i}| = 1$, the EL equations for $\mathbf{r}_i(t)$ will be the same whether the constraint $\sqrt{\mathbf{r}^2_{i+1/i}} = 1$ or $\mathbf{r}^2_{i+1/i} = 1$ is used.

The reparametrization invariance present for point particles (cf. section 1) is still present for beads on the chain, but the parametrization of arc-length along the chain is taken to be fixed by the discretization.

2.2. General variation of the distance functional

For reasons that will become clear as we progress, we consider the general variation of the functional $\mathcal{D}$, allowing for *broken extremals*. That is, we allow the curves describing the particle trajectories to be non-smooth in principle at one or more points in time. Consider the case of one such point at time t_1. The distance can be written as

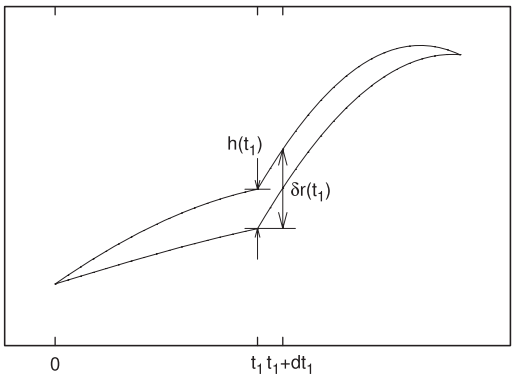


Figure 2. General variations of a functional with fixed end points allow for broken extremals. In the text we derive the extra 'corner' conditions for a piecewise continuous path to still be extremal for our distance functional.

$$\mathcal{D} = \int_0^{t_1} \mathrm{d}t\, \mathcal{L}(\mathbf{r}_i, \dot{\mathbf{r}}_i) + \int_{t_1}^{T} \mathrm{d}t\, \mathcal{L}(\mathbf{r}_i, \dot{\mathbf{r}}_i). \tag{11}$$

The space trajectories of the particles must be continuous at time t_1, so $\mathbf{r}_i(t_1 - \epsilon) = \mathbf{r}_i(t_1 + \epsilon)$, or in shorthand

$$\mathbf{r}_i\left(t_1^-\right) = \mathbf{r}_i\left(t_1^+\right). \tag{12}$$

Let $\mathbf{r}_i(t)$ and $\tilde{\mathbf{r}}_i(t)$ be two neighbouring trajectories from $\mathbf{r}_i(0) = \mathbf{r}_{\mathrm{A}i}$ to $\mathbf{r}_i(T) = \mathbf{r}_{\mathrm{B}i}$ (see figure 2). Neighbouring curves will differ by the first order quantity $\mathbf{h}_i(t) = \tilde{\mathbf{r}}_i(t) - \mathbf{r}_i(t)$. The fixed boundary conditions at $t = 0, T$ dictate that $\mathbf{h}_i(0) = \mathbf{h}_i(T) = 0$. The difference in distance between the two trajectories is

$$\begin{aligned}\Delta\mathcal{D} &= \mathcal{D}[\mathbf{r}_i + \mathbf{h}_i] - \mathcal{D}[\mathbf{r}_i] \\ &= \int_0^{t_1+\delta t_1} \mathrm{d}t\, \mathcal{L}(\mathbf{r}_i + \mathbf{h}_i, \dot{\mathbf{r}}_i + \dot{\mathbf{h}}_i) - \int_0^{t_1} \mathrm{d}t\, \mathcal{L}(\mathbf{r}_i, \dot{\mathbf{r}}_i) \\ &\quad + \int_{t_1+\delta t_1}^{T} \mathrm{d}t\, \mathcal{L}(\mathbf{r}_i + \mathbf{h}_i, \dot{\mathbf{r}}_i + \dot{\mathbf{h}}_i) - \int_{t_1}^{T} \mathrm{d}t\, \mathcal{L}(\mathbf{r}_i, \dot{\mathbf{r}}_i).\end{aligned} \tag{13}$$

Taylor expanding the Lagrangian to first order in $\mathbf{h}_i$,[2]

$$\mathcal{L} \approx \mathcal{L}(\mathbf{r}_i, \dot{\mathbf{r}}_i) + \sum_{i=1}^{N} \left(\mathcal{L}_{\mathbf{r}_i} \cdot \mathbf{h}_i + \mathcal{L}_{\dot{\mathbf{r}}_i} \cdot \dot{\mathbf{h}}_i \right)$$

and integrating by parts using the fixed boundary conditions at $t = 0, T$, the difference in distance up to first order in $\mathbf{h}_i$ is

$$\begin{aligned}\Delta\mathcal{D} \approx &\int_0^{t_1} \mathrm{d}t \sum_i \left(\mathcal{L}_{\mathbf{r}_i} - \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}_{\dot{\mathbf{r}}_i} \right) \cdot \mathbf{h}_i \\ &+ \int_{t_1}^{T} \mathrm{d}t \sum_i \left(\mathcal{L}_{\mathbf{r}_i} - \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}_{\dot{\mathbf{r}}_i} \right) \cdot \mathbf{h}_i + \mathcal{L}(t_1^-)\delta t_1 - \mathcal{L}(t_1^+)\delta t_1 \\ &+ \sum_i \mathcal{L}_{\dot{\mathbf{r}}_i} \cdot \mathbf{h}_i|_{t_1^-} - \sum_i \mathcal{L}_{\dot{\mathbf{r}}_i} \cdot \mathbf{h}_i|_{t_1^+}\end{aligned} \tag{14}$$

with the shorthand $\mathcal{L}(t) \equiv \mathcal{L}(\mathbf{r}_i(t), \dot{\mathbf{r}}_i(t))$.

[2] We use the notation $F_{\mathbf{r}} \equiv \partial F/\partial \mathbf{r}$, $F_{\dot{\mathbf{r}}} \equiv \partial F/\partial \dot{\mathbf{r}}$.

2.3. Conditions for an extremum

The variation $\delta\mathcal{D}$ differs from $\Delta\mathcal{D}$ above only by second order terms. Then for the transformation from $\{\mathbf{r}_{Ai}\}$ to $\{\mathbf{r}_{Bi}\}$ to be an extremum, $\delta\mathcal{D} = 0$. Thus, the EL equations (in the top line of equation (14)) must vanish in each regime $[0, t_1)$, $(t_1, T]$. Using the form of the Lagrangian in equation (10), the EL equations become

$$\dot{\hat{\mathbf{v}}}_1 + \lambda_{12}\,\mathbf{r}_{2/1} = 0 \tag{15a}$$

$$\dot{\hat{\mathbf{v}}}_2 - \lambda_{12}\,\mathbf{r}_{2/1} + \lambda_{23}\,\mathbf{r}_{3/2} = 0 \tag{15b}$$

$$\vdots$$

$$\dot{\hat{\mathbf{v}}}_N - \lambda_{N-1,N}\,\mathbf{r}_{N/(N-1)} = 0. \tag{15c}$$

According to equation (14) there are additional conditions for the transformation to be an extremum. To find these first note that up to first order (see figure 2)

$$\mathbf{h}_i(t_1) \approx \delta\mathbf{r}_i(t_1) - \dot{\mathbf{r}}_i(t_1)\,\delta t_1. \tag{16}$$

Then the first variation in the distance is

$$\delta\mathcal{D} = \left[\left(\mathcal{L} - \sum_i \dot{\mathbf{r}}_i\cdot\mathcal{L}_{\dot{\mathbf{r}}_i}\right)\Big|_{t_1^-} - \left(\mathcal{L} - \sum_i \dot{\mathbf{r}}_i\cdot\mathcal{L}_{\dot{\mathbf{r}}_i}\right)\Big|_{t_1^+}\right]\delta t_1 + \sum_i \left[\mathcal{L}_{\dot{\mathbf{r}}_i}|_{t_1^-} - \mathcal{L}_{\dot{\mathbf{r}}_i}|_{t_1^+}\right]\cdot\delta\mathbf{r}_i(t_1) \tag{17}$$

which must vanish at an extremum. Because the variations $\delta\mathbf{r}_i$ and δt_1 are all independent, the terms in square brackets in equation (17) must vanish. Writing these expressions in terms of the conjugate momenta $\mathbf{p}_i = \mathcal{L}_{\dot{\mathbf{r}}_i}$ and Hamiltonian $\mathcal{H} = \sum_i \dot{\mathbf{r}}_i\cdot\mathbf{p}_i - \mathcal{L}$ gives the conditions

$$\mathbf{p}_i|_{t_1^-} = \mathbf{p}_i|_{t_1^+} \tag{18a}$$

$$\mathcal{H}|_{t_1^-} = \mathcal{H}|_{t_1^+}. \tag{18b}$$

These conditions are called the *Weierstrass–Erdmann conditions* or *corner conditions* in the calculus of variations [2].

According to the Lagrangian in equation (10), the Hamiltonian is given by

$$\mathcal{H} = -\sum_{i=1}^{N} \frac{\lambda_{i,i+1}}{2}\left(\mathbf{r}_{i+1/i}^2 - 1\right)$$

which is identically zero, so corner condition (18*b*) provides no further information.

The conjugate momenta according to (10) are given by

$$\mathbf{p}_i = \frac{\dot{\mathbf{r}}_i}{|\dot{\mathbf{r}}_i|} = \hat{\mathbf{v}}_i\,. \tag{19}$$

Therefore, according to corner condition (18*a*), extremal trajectories cannot suddenly change direction: each $\mathbf{r}_i(t)$ follows a smooth path continuous up to first derivatives in the spatial coordinates.

The fact that one corner condition provided no information due to the vanishing of the Hamiltonian is related to our choice of parametrization in formulating the problem. For example, in the case of the distance of the single point particle mentioned in the introduction the Lagrangian may be defined either through independent variable x as $\mathcal{L}^{(x)} = \sqrt{1 + y'^2 + z'^2}$ (with e.g. $y' = \mathrm{d}y/\mathrm{d}x$) or parametrically through independent variable t as $\mathcal{L}^{(t)} = \sqrt{\dot{\mathbf{r}}^2}$. The conjugate momenta are then either $\mathcal{L}^{(x)}_{y'} = y'/\sqrt{1 + y'^2 + z'^2}$ and $\mathcal{L}^{(x)}_{z'} = z'/\sqrt{1 + y'^2 + z'^2}$, or $\mathcal{L}^{(t)}_{\dot{\mathbf{r}}} = \dot{\mathbf{r}}/|\dot{\mathbf{r}}| \equiv \hat{\mathbf{v}}$. The Hamiltonian are either $\mathcal{H}^{(x)} = 1/\sqrt{1 + y'^2 + z'^2}$ or $\mathcal{H}^{(t)} = \mathcal{L}^{(t)} - \dot{\mathbf{r}}\cdot(\dot{\mathbf{r}}/|\dot{\mathbf{r}}|) = 0$. The corner conditions can be shown to be equivalent for both choices of independent variable: for $\mathcal{L}^{(t)}$ they give $\hat{\mathbf{v}}(t_1^-) = \hat{\mathbf{v}}(t_1^+)$, so that the direction of the tangent to the curve cannot have a discontinuity. Together, the Hamiltonian and two conjugate momenta for $\mathcal{L}^{(x)}$ can be interpreted as components of the unit tangent vector to the curve, i.e. $\hat{\mathbf{t}}(x) = (\hat{\mathbf{i}} + y'\hat{\mathbf{j}} + z'\hat{\mathbf{k}})/\sqrt{1 + y'^2 + z'^2}$, and so once again the corner conditions enforce a continuous tangent vector, here $\hat{\mathbf{t}}(x_1^-) = \hat{\mathbf{t}}(x_1^+)$.

2.3.1. Boundary conditions.

In the continuum limit, the boundary conditions on $\mathbf{r}(s,t)$ are $\mathbf{r}(s,0) = \mathbf{r}_A(s)$, $\mathbf{r}(s,T) = \mathbf{r}_B(s)$, where $\mathbf{r}_A$ and $\mathbf{r}_B$ are the two configurations of the polymer. For discrete chains, these boundary conditions become

$$\{\mathbf{r}_i(0)\} = \{\mathbf{r}_i^{(A)}\} \tag{20a}$$

$$\{\mathbf{r}_i(T)\} = \{\mathbf{r}_i^{(B)}\}. \tag{20b}$$

There are also boundary conditions that hold for the *end points* of the chain at all times. From equations (15*a*) and (15*c*) we see that there are three solutions for the end points of the chain.

(1) If $\lambda \neq 0$, purely rotational motion results. This can be seen by taking the dot product of equation (15*a*) with $\mathbf{v}_1$, which yields $\lambda_{12}\mathbf{v}_1\cdot\mathbf{r}_{2/1} = 0$, so the velocity of the end point is orthogonal to the link. The rotation must be about a point that is internal to the link, i.e. on the line between points 1 and 2 for end point 1. This can be seen straightforwardly for the case of one link by removing point 3 from equations (15*a*) and (15*b*). Then the accelerations $\dot{\hat{\mathbf{v}}}_i$ must be in opposite directions. This can only occur if rotation is about a point on the line between points 1 and 2.

(2) If $\lambda = 0$, $\dot{\hat{\mathbf{v}}}_i = 0$, and straight line motion of the end point results.

(3) Writing out the time-derivative in (15*a*) yields

$$\mathbf{v}_1^2\dot{\mathbf{v}}_1 - (\mathbf{v}_1\cdot\dot{\mathbf{v}}_1)\,\mathbf{v}_1 = -\lambda_{12}\,|\mathbf{v}_1|^3\,\mathbf{r}_{2/1} \tag{21}$$

which has the trivial solution $\mathbf{v}_1 = 0$. The end point can be at rest, while other parts of the chain move.

2.4. Sufficient conditions for a minimum

For a transformation to be minimal, it is necessary, but not sufficient, that it be an extremum. We now derive the sufficient conditions for a given transformation to minimize the functional (9). We describe the formalism in some detail because it is not typically taught to physicists—for further reading see for example [2]. This section can be read

independently of the others, and might be skipped on first reading.

According to *Sylvester's criterion*, a quadratic form $\sum_{ij} A_{ij}x_i x_j$ is positive definite if and only if all descending principle minors of the matrix $\|A_{ij}\|$ are positive, i.e.

$$A_{11} > 0, \qquad \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} > 0,$$
$$\begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} > 0, \ldots, \det\|A_{ij}\| > 0, \tag{22}$$

and a function F of $\mathbf{x} \equiv (x_1, x_2, \ldots, x_n)$ has a minimum at $\mathbf{x}^\star$ if the Jacobian matrix $\|\partial^2 F/\partial x_i \partial x_j\|$ is positive definite at the position of the extremum (where $\partial F/\partial x_i = 0$).

For a function to be a minimum of a given functional, it must satisfy similar sufficient conditions. Consider again the difference in distance between two trajectories in (9).[3] Taylor expanding the Lagrangian to second order in $\mathbf{h}_i$,

$$\begin{aligned}\Delta\mathcal{D} &= \mathcal{D}[\mathbf{r}_i + \mathbf{h}_i] - \mathcal{D}[\mathbf{r}_i] \\ &= \int_0^T \mathrm{d}t\,\mathcal{L}(\mathbf{r}_i + \mathbf{h}_i, \dot{\mathbf{r}}_i + \dot{\mathbf{h}}_i) - \int_0^T \mathrm{d}t\,\mathcal{L}(\mathbf{r}_i, \dot{\mathbf{r}}_i) \\ &\approx \int_0^T \mathrm{d}t \Bigg[\sum_{i=1}^{N} \left(\mathcal{L}_{\mathbf{r}_i} \cdot \mathbf{h}_i + \mathcal{L}_{\dot{\mathbf{r}}_i} \cdot \dot{\mathbf{h}}_i \right) \\ &\quad + \tfrac{1}{2} \sum_{i,j}^{3N} \left(\mathcal{L}_{x_i x_j} h_i h_j + 2\mathcal{L}_{x_i \dot{x}_j} h_i \dot{h}_j + \mathcal{L}_{\dot{x}_i \dot{x}_j} \dot{h}_i \dot{h}_j \right) \Bigg]. \end{aligned} \tag{23}$$

At an extremum, the first order term in (23) is zero, and $\Delta\mathcal{D} \approx \delta^2\mathcal{D}$, the second variation. For the extremum to be a minimum, $\delta^2\mathcal{D} > 0$. From equation (10), the matrix $\|\mathcal{L}_{x_i \dot{x}_j}\| = \|0\|$. Assuming $\|\mathcal{L}_{x_i \dot{x}_j}\|$ is in general a symmetric matrix, i.e. $\mathcal{L}_{x_i \dot{x}_j} = \mathcal{L}_{x_j \dot{x}_i}$, the second term in the quadratic form of (23) may be integrated by parts to give

$$\delta^2\mathcal{D} = \tfrac{1}{2} \int_0^T \mathrm{d}t \left[\langle \dot{h} | \mathrm{P} \dot{h} \rangle + \langle h | \mathrm{Q} h \rangle \right], \tag{24}$$

where we have let $|h\rangle$ denote the vector $(h_1, h_2, \ldots, h_{3N})$, and used the shorthand P and Q for the matrices:

$$\begin{aligned} \mathrm{P}(t) &= \|\mathrm{P}_{ij}\| = \|\mathcal{L}_{\dot{x}_i \dot{x}_j}\| \\ \mathrm{Q}(t) &= \|\mathrm{Q}_{ij}\| = \left(\|\mathcal{L}_{x_i x_j}\| - \frac{\mathrm{d}}{\mathrm{d}x} \|\mathcal{L}_{x_i \dot{x}_j}\| \right). \end{aligned} \tag{25}$$

From (10) the explicit form for these matrices may be calculated. P is block diagonal:

$$\mathrm{P} = \begin{bmatrix} I_{ij}^{(1)} & 0 & \cdots & 0 \\ 0 & I_{ij}^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_{ij}^{(N)} \end{bmatrix} \tag{26}$$

with each block matrix having elements

$$\begin{aligned} \|I_{ij}^{(J)}\| &= \frac{1}{|\dot{\mathbf{r}}^J|^3} \left(\delta_{ij} \dot{\mathbf{r}}^{(J)^2} - \dot{x}_i^{(J)} \dot{x}_j^{(J)} \right) \\ &= \frac{1}{|\dot{\mathbf{r}}^J|^3} \begin{bmatrix} \dot{y}^2 + \dot{z}^2 & -\dot{x}\dot{y} & -\dot{x}\dot{z} \\ -\dot{x}\dot{y} & \dot{x}^2 + \dot{z}^2 & -\dot{y}\dot{z} \\ -\dot{x}\dot{z} & -\dot{y}\dot{z} & \dot{x}^2 + \dot{y}^2 \end{bmatrix}_{(\text{particle } J)}. \end{aligned} \tag{27}$$

[3] We ignore corner conditions for purposes of the derivation. It can be shown that they do not modify the result.

Interestingly the numerator of (27) has the form of an inertia tensor for a point particle in velocity-space. The matrix Q is block tri-diagonal, because the spatial derivatives in (25) couple each bead to its two neighbours. Using indices I, J to enumerate beads and i, j to enumerate x, y, z components for each bead,

$$\begin{aligned} \|\mathrm{Q}_{IJ,ij}\| &= \delta_{ij} \big[\lambda_{J-1,J} \left(\delta_{IJ} - \delta_{I,J-1} \right) + \lambda_{J,J+1} \\ &\quad \times \left(\delta_{IJ} - \delta_{I,J+1} \right) \big] \qquad \text{or} \end{aligned}$$

$$\mathrm{Q} = \begin{bmatrix} \lambda_{12}\mathbf{1} & -\lambda_{12}\mathbf{1} & 0 & 0 & \cdots \\ -\lambda_{12}\mathbf{1} & (\lambda_{12}+\lambda_{23})\mathbf{1} & -\lambda_{23}\mathbf{1} & 0 & \cdots \\ 0 & -\lambda_{23}\mathbf{1} & (\lambda_{23}+\lambda_{34})\mathbf{1} & -\lambda_{34}\mathbf{1} & \\ \vdots & \ddots & \ddots & \ddots & \\ & & & -\lambda_{N-1,N}\mathbf{1} & \lambda_{N-1,N}\mathbf{1} \end{bmatrix}. \tag{28}$$

For the transformation $\mathbf{r}^*(t)$ to be a minimum of $\mathcal{D}[\mathbf{r}]$, the functional (24) must be positive definite for all $|h\rangle$. To derive the conditions for this, we can temporarily ignore the fact that (24) arose from the second variation of (9), and treat (24) as a new functional of the function $|h(t)\rangle = |h_1(t), \ldots, h_{3N}(t)\rangle$. We then ask what $|h(t)\rangle$ extremizes (24). If $\delta^2\mathcal{D} > 0$ we expect that the only extremal solution would be the trivial one: $|h(t)\rangle = |0\rangle$, at least for small variations of the $h_i(t)$. That is, changing the transformation $\{\mathbf{r}_i^*(t)\}$ from that which extremized (9) to a neighbouring transformation $\{\mathbf{r}_i^*(t) + \mathbf{h}_i(t)\}$ would increase the distance travelled.

The system of $3N$ EL equations for $|h\rangle$ from (24) is

$$-\frac{\mathrm{d}}{\mathrm{d}t} |\mathrm{P}\dot{h}\rangle + |\mathrm{Q}h\rangle = |0\rangle \tag{29}$$

with boundary conditions

$$|h(0)\rangle = |h(T)\rangle = |0\rangle. \tag{30}$$

Equation (29) is referred to as the *Jacobi equation* in the calculus of variations.

First note that if $|h\rangle$ satisfies the system of equations in (29) as well as the boundary conditions (30), then integration by parts gives

$$\begin{aligned} \delta^2\mathcal{D} &= \int_0^T \mathrm{d}t \left(\langle \dot{h} | \mathrm{P}\dot{h} \rangle + \langle h | \mathrm{Q}h \rangle \right) \\ &= \int_0^T \mathrm{d}t \langle h | - \frac{\mathrm{d}}{\mathrm{d}t} \left(\mathrm{P}\dot{h} \right) + \mathrm{Q}h \rangle = 0. \end{aligned} \tag{31}$$

This means that for $\delta^2\mathcal{D}$ to be >0, any nontrivial $|h(t)\rangle$ which satisfies the boundary conditions must not itself be an extremal solution of the Jacobi equation, otherwise solutions $|\mathbf{r}^*(t)\rangle$ perturbed by any constant times $|h(t)\rangle$ are themselves

extremals. One may think of this by analogy as the necessity for the absence of any 'Goldstone modes', where excitations by various $C|h(t)\rangle$ would lead to a family of curves with zero cost in action, and thus zero effective restoring force, between them.

Alternatively, we can ask what equation $\mathbf{h} \equiv |h\rangle$ must satisfy if the EL equations are satisfied for both $\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}})$ and the neighbouring extremal $\mathcal{L}(\mathbf{r}+\mathbf{h}, \dot{\mathbf{r}}+\dot{\mathbf{h}})$. Taylor expanding $\mathcal{L}(\mathbf{r}+\mathbf{h}, \dot{\mathbf{r}}+\dot{\mathbf{h}})$ in

$$\mathcal{L}_{\mathbf{r}}(\mathbf{r}+\mathbf{h}, \dot{\mathbf{r}}+\dot{\mathbf{h}}) - \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}_{\dot{\mathbf{r}}}(\mathbf{r}+\mathbf{h}, \dot{\mathbf{r}}+\dot{\mathbf{h}}) = 0$$

gives

$$-\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{L}_{\dot{\mathbf{r}}\dot{\mathbf{r}}} \cdot \dot{\mathbf{h}}\right) + \left(\mathcal{L}_{\mathbf{r}\mathbf{r}} - \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}_{\mathbf{r}\dot{\mathbf{r}}}\right) \cdot \mathbf{h} = 0$$

which is exactly Jacobi's equation (29) with definitions (25).

From here on, it is much simpler to elucidate the central concepts for sufficient conditions using the case of a single scalar function $h(t)$. The analysis can be generalized to the multi-dimensional case with a bit more effort, but the conclusions are essentially the same and so they will simply be stated along with the conclusions for the '1D' case. For further details see [2].

We write equation (24) in 1D as

$$\frac{1}{2}\int_0^T \mathrm{d}t \left(P\dot{h}^2 + Qh^2\right). \tag{32}$$

It was realized originally by Legendre that the integral could be brought to a simpler form by adding zero to it in the form of a total derivative. Since

$$\int_0^T \mathrm{d}t \frac{\mathrm{d}}{\mathrm{d}t}\left(w(t)h^2\right) = 0$$

for any $w(t)$ so long as $h(t)$ satisfies the boundary conditions (30), we can add it to the integral in (32) and seek a function $w(t)$ such that the expression

$$\delta^2\mathcal{D} = \frac{1}{2}\int_0^T \mathrm{d}t \left(P\dot{h}^2 + 2wh\dot{h} + (Q+\dot{w})\,h^2\right)$$

may be written as a perfect square. This yields the differential equation

$$P\,(Q+\dot{w}) = w^2 \tag{33}$$

for $w(t)$, and second variation

$$\delta^2\mathcal{D}[h] = \frac{1}{2}\int_0^T \mathrm{d}t\, P\left(\dot{h} + \frac{w}{P}h\right)^2. \tag{34}$$

Therefore, a necessary condition for a minimum is for $P > 0$. The analogous condition in the multi-dimensional case is for the matrix $\|\mathrm{P}\|$ to be positive definite.

If the differential term $\dot{h} + \frac{w}{P}h$ in (34) were equal to zero for some $h(t)$, the boundary condition $h(0) = 0$ would then imply $\dot{h}(0) = 0$ and thus $h(t) = 0$ for all t by the uniqueness theorem, as applied to this first order differential equation.

Therefore the functional (34) is positive definite if, and only if,

(1) $P > 0$,
(2) a solution for equation (33) exists for the whole interval $[0, T]$.

In general, there is no guarantee of condition (2) even if condition (1) is valid. For example if $P = 1$, $Q = -1$, (33) has solution $w(t) = \tan(t + c)$, which has no finite solution if $|T| > \pi$.[4]

If (33) has a pole at say $\tilde{t}$, then for the integral (34) to remain finite $h(\tilde{t}) \to 0$. This point is said to be conjugate to the point $t_o = 0$, i.e. it is a *conjugate point*.

Moreover, equation (33) is a Riccati equation, which may be brought to linear form by the transformation $w(t) = -P\dot{H}/H$, with $H(t)$ an unknown function. Substitution in (33) gives

$$-\frac{\mathrm{d}}{\mathrm{d}t}\left(P\dot{H}\right) + QH = 0 \tag{35}$$

which is precisely equation (29)—the Jacobi equation for $h(t)$.

This means that, for equation (33) to have a solution on $[0, T]$, $H(t)$, as given by the solution to (35), must have no roots on $[0, T]$. But because equation (35) holds for $h(t)$ as well, $h(t)$ must have no roots (conjugate points) on $[0, T]$. Because $h(0) = h(T) = 0$, the only way to extremize (32) is to satisfy equation (35) with the trivial solution $h(t) = 0$. If $h(t) \neq 0$ for $0 < t < T$ then it would mean that there was a conjugate point at $\tilde{t} = T$.

In the multi-dimensional case an extremal $|h\rangle$ is one of $3N$ vectors satisfying equation (29), i.e. $|h^{(\alpha)}\rangle = |h_1^{(\alpha)} \ldots h_{3N}^{(\alpha)}\rangle$, $1 \leqslant \alpha \leqslant 3N$. A conjugate point is defined as a point where the determinant vanishes:

$$\det \begin{vmatrix} h_1^{(1)}(t) & \cdots & h_1^{(3N)}(t) \\ \vdots & & \vdots \\ h_1^{(3N)}(t) & \cdots & h_{3N}^{(3N)}(t) \end{vmatrix} = 0.$$

The conditions for a transformation to be minimal are then the following:

(1) the transformation $|\mathbf{r}^*(t)\rangle = \{\mathbf{r}_i^*(t)\}$ is extremal,
(2) along $|\mathbf{r}^*(t)\rangle$, the matrix $\mathrm{P}(t) = \mathcal{L}_{\dot{x}_i\dot{x}_j}$ is positive definite, and
(3) the interval $[0, T]$ contains no conjugate points to $t = 0$.

The above ideas can be made clear with a few examples below.

2.4.1. Distance between points. From the effective Lagrangian $\mathcal{L} = \sqrt{\dot{\mathbf{r}}^2}$, $\mathrm{P} = \|\mathcal{L}_{\dot{x}_i\dot{x}_j}\|$ is given in equation (27), which has determinant $\det \mathrm{P} = 0$, and so is not positive definite. This is due to our choice of parametrization. If we break symmetry by choosing one spatial direction as the independent variable, $\mathcal{L}(x, y', z') = \sqrt{1 + y'^2 + z'^2}$ (with e.g. $y' \equiv \mathrm{d}y/\mathrm{d}x$ and $x_0 \leqslant x \leqslant x_1$). Then

$$\mathrm{P} = \frac{1}{\left(1 + y'^2 + z'^2\right)^{3/2}} \begin{pmatrix} 1 + z'^2 & -y'z' \\ -y'z' & 1 + y'^2 \end{pmatrix}$$

[4] Because of reparametrization invariance in our problem, the value of T is adjustable; however, precisely because of this invariance, $\det\|\mathrm{P}\| = 0$ and so is no longer positive definite. We discuss this problem and its resolution below.

with positive definite determinant $\det\|\mathrm{P}\| = (1 + y'^2 + z'^2)^{-1/2} > 0$ for any trajectory. From equation (25), $\|\mathrm{Q}(t)\| = \|0\|$. Along the extremal, where $y(x) = ax + y_0$, $z(x) = bx + z_0$, equation (29) gives $\mathrm{P} \cdot \mathbf{h}' = \mathbf{c}$, with $\mathbf{c}$ a constant vector and P a positive definite matrix of constant values with respect to x. Solving this first order equation gives straight line solutions for $\mathbf{h}(x)$. Because $\mathbf{h}(x_0) = 0$, there can be no conjugate points, and because $\mathbf{h}(x_1) = 0$, the only solution to (29) is the trivial one, and the extremum is a minimum.

2.4.2. Geodesics on the surface of a sphere. Taking the azimuthal angle ϕ as the independent variable, and polar angle $\theta(\phi)$ as the dependent variable, the arc-length on the surface of a unit sphere may be written as

$$\mathcal{D}[\theta] = \int_{\phi_0}^{\phi_1} \mathrm{d}\phi \, \sqrt{\theta'^2 + \sin^2\theta}. \tag{36}$$

The EL equations give the extremal trajectory as $\cos\theta = A\sin\theta\cos\phi + B\sin\theta\sin\phi$ with A, B constants. This is the equation of a plane $z = Ax + By$, which intersects the surface of the sphere to make a great circle. The scalar $P = \mathcal{L}_{\theta'\theta'} = \sin^2\theta/(\theta'^2 + \sin^2\theta)^{3/2}$ which is always positive. To simplify the problem, let $\phi_0 = 0$, and $\theta(\phi_0) = \theta(\phi_1) = \pi/2$, so the great circle lies in the $z = 0$ plane. Along this extremal P is constant and equal to 1, while $Q = -1$. The second variation, equation (32), is then $(1/2)\int_0^{\phi_1}\mathrm{d}\phi\ (h'^2 - h^2)$. The corresponding Jacobi equation, $h'' + h = 0$, must not have a root between $[0, \phi_1]$. The nontrivial solution to the Jacobi equation satisfying the initial condition $h(0) = 0$ is $h(\phi) = C\sin\phi$, which has a conjugate point at $\phi = \pi$. Thus for the extremal curve to be minimal, ϕ_1 must be $< \pi$, the location of the opposite pole on the sphere. If $\phi_1 < \pi$, there is no extremal solution for $h(\phi)$ other than the trivial one which satisfies the boundary conditions. It is instructive to look at the arc-length under sinusoidal variations around the extremal path which satisfy the boundary conditions $h(0) = h(\phi_1) = 0$, so that $\theta(\phi) = \pi/2 + h(\phi) = \pi/2 + \epsilon\sin(\pi\phi/\phi_1)$. Inserting this into equation (36) above and expanding to second order in ϵ, we see that first order terms in ϵ vanish, and the difference in distance from the extremal path is $\Delta\mathcal{D} = (\epsilon^2/4\phi_1)(\pi^2 - \phi_1^2)$. For $\phi_1 < \pi$ this is always greater than zero, indicating the extremal is a minimum. For $\phi_1 > \pi$ this is always less than zero, indicating the extremal is a maximum with respect to these perturbations: the length may be shortened. When $\phi_1 = \pi$, $\Delta\mathcal{D} = 0$ to second order. When $h(\phi)$ represents the difference between great circles $\Delta\mathcal{D}$ is precisely zero.

2.4.3. Harmonic oscillator. It is not widely appreciated that the classical action for a simple harmonic oscillator is not always a minimum, and indeed in many cases can be a maximum with respect to some perturbations. The action for a harmonic oscillator with given spring constant is proportional to $S[x] = \int_0^T \mathrm{d}t\,\frac{1}{2}(\dot{x}^2 - x^2)$, which has EL equation $\ddot{x} + x = 0$. Taking the specific initial conditions $x(0) = 1$, $\dot{x}(0) = 0$, the extremal solution is $x(t) = \cos t$. The scalar $P(t) = \mathcal{L}_{\dot{x}\dot{x}} = 1$, which is always positive and satisfies the necessary conditions for a minimum. The scalar $Q = \mathcal{L}_{xx} - \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}_{x\dot{x}} = -1$. The second variation $\delta^2 S[h] = \frac{1}{2}\int_0^T \mathrm{d}t\,(\dot{h}^2 - h^2)$, which has Jacobi equation $\ddot{h} + h = 0$. This is the same Jacobi equation as that for geodesics on a sphere, so the sufficient conditions will parallel those above. The boundary condition $h(0) = 0$ gives $h(t) = A\sin t$, with conjugate points at $t = n\pi$, $n = 1, 2, \ldots$. This means that the action is a minimum only so long as $T < \pi$, i.e. a half-period. If we let $x(t)$ be the extremal solution plus a sin perturbation satisfying the Jacobi equation at the conjugate points, $x(t) = \cos t + \epsilon\sin t$, then the difference in action from the extremal path becomes $\Delta S = (\epsilon^2/4T)(\pi^2 - T^2)$. This result is exact because the action for the oscillator is quadratic (as opposed to the action for geodesics). When $T < \pi$, $\Delta S > 0$, indicating the extremal is a minimum. When T is larger than a half-period $\Delta S < 0$ and the extremal trajectory is a maximum (with respect to half-wavelength sinusoidal perturbations), and when $T = \pi$ the end point is the conjugate point and $\Delta S = 0$.

We discuss sufficient conditions further below in the context of minimal transformations for links.

3. Single links

In the limit of one link, equations (15*a*)–(15*c*) reduce to

$$\begin{aligned} \dot{\hat{\mathbf{v}}}_\mathrm{A} + \lambda\,\mathbf{r}_{\mathrm{B/A}} &= 0 \\ \dot{\hat{\mathbf{v}}}_\mathrm{B} - \lambda\,\mathbf{r}_{\mathrm{B/A}} &= 0 \end{aligned} \tag{37}$$

where we have let A represent point 1, B point 2, and $\lambda \equiv \lambda_{12}$. The link has length 1 in our dimensionless formulation, so the vector $\mathbf{r}_{\mathrm{B/A}}$ could also have been written as a unit vector $\hat{\mathbf{r}}_{\mathrm{B/A}}$.

Both points A and B are end points and satisfy the boundary conditions of section 2.3.1. This means that points A and B move by either pure rotation or straight line translation, or remain at rest. The initial and final conditions may be written $\mathbf{r}_\mathrm{A}(0) = \mathbf{A}$, $\mathbf{r}_\mathrm{B}(0) = \mathbf{B}$, $\mathbf{r}_\mathrm{A}(T) = \mathbf{A}'$, $\mathbf{r}_\mathrm{B}(T) = \mathbf{B}'$.

The link in our problem has direction, so A must transform to A′ and B to B′. We will often use arrowheads in figures to denote this direction.

3.1. Straight line transformations

As a first example, consider the two links shown in figure 3(a). The four points A, B, A′, B′ need not lie in a plane (see for example figure 3(b)). Let angle $\angle\mathrm{BAA}' \equiv a$ be obtuse. We draw straight lines from A to A′ and B to B′, and ask whether such a transformation is possible. We can thus derive the following rule:

- For a straight line transformation to exist between two links, opposite angles of the quadrilateral made by $\overline{\mathrm{AB}}$, $\overline{\mathrm{A'B'}}$, $\overline{\mathrm{AA'}}$, $\overline{\mathrm{BB'}}$ must be obtuse.

Let the length that point A travels be x_A, i.e. we imagine the point A′ and the distance $x_\mathrm{A} = |\mathrm{AA}'|$ to be variable. The length r_B that point B travels is then a function of x_A and the original angle a, $r_B(x_\mathrm{A}, a)$. We can now find conditions on the angle $b \equiv \angle\mathrm{BB'A'}$ such that the transformation is possible.

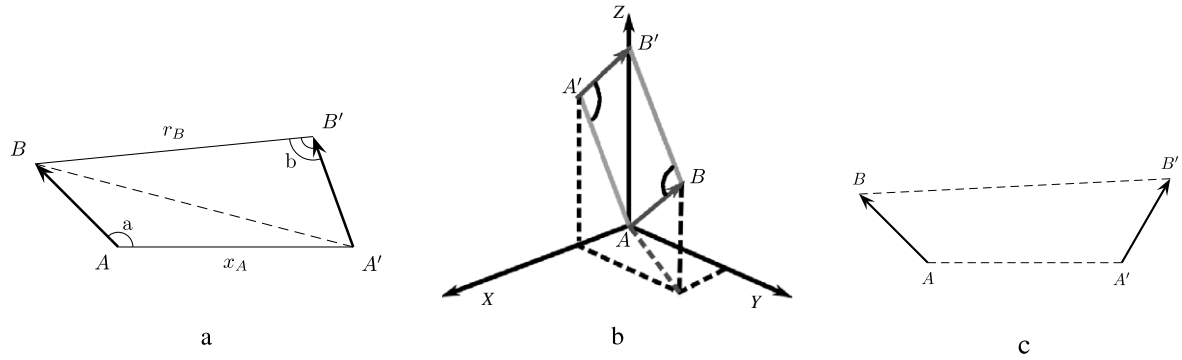


Figure 3. Possible ((a), (b)) and impossible (c) straight line transformations between links AB and A′B′. (b) A straight line transformation where the initial and final states do not lie in the same plane. In the text we derive the conditions for the possibility of a straight line transformation between links.

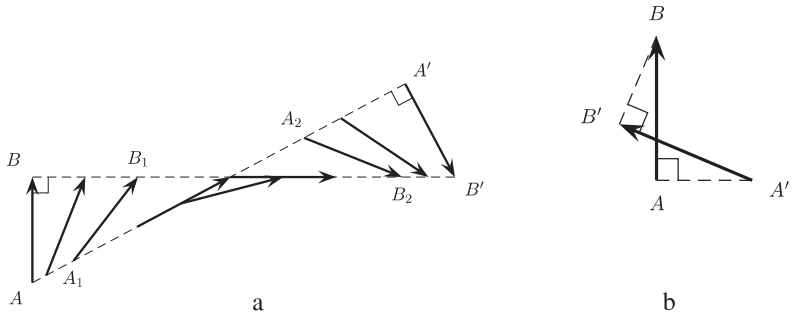


Figure 4. (a) An example of a set of link configurations connected by a straight line transformation. The link rotates clockwise as it translates to allow the end points to move in straight lines. The translation can proceed no farther than the end points AB and A′B′, which have link vectors $\overrightarrow{AB}$ or $\overrightarrow{A'B'}$ that are perpendicular to one or other of the vectors $\hat{\mathbf{v}}_A$ or $\hat{\mathbf{v}}_B$. The totality of states thus connected forms a 'bow-tie'. (b) A bow-tie where the terminal states AB and A′B′ happen to cross each other.

After some distance x_A travelled by point A, the length of the line from B to A′ is

$$\overline{BA'} = x_A^2 + 1 - 2x_A \cos a$$
$$= r_B^2 + 1 - 2r_B \cos b$$

so that

$$r_B(x_A, a) = \cos b \pm \sqrt{\cos^2 b + f(x_A, a)}$$

with $f(x_A, a) = x_A^2 - 2x_A \cos a$. Since a is obtuse, $f > 0$ when $x_A > 0$, and so the positive root must be taken for r_B to positive. When $x_A = 0$, $f(0, a) = 0$, and

$$r_B(0, a) = \cos b + |\cos b| = 0.$$

Therefore b must also be an obtuse angle. If two opposite angles are obtuse, then the other two angles must be acute. This concludes the proof that the above conditions are sufficient. An additional proof that they are necessary is given in appendix A.

We readily see that figure 3(a) is one pair of a larger set of straight line transformations that can continue until one or both of the obtuse angles reaches 90°. This collection forms a 'bow-tie' of admissible configurations, as in figure 4. Note that straight lines in the quadrilateral may cross as in the transformation from A, B to A′, B′ in figure 4. Trivial translations of the link without any concurrent rotation are a special case of general straight line transformations.

3.2. Piecewise extremal transformations: transformations with rotations

An immediate question is the nature of the transformation between AB and A′B′ in figure 3(c), where opposite angles of the quadrilateral are not obtuse. Recall our link has direction so A cannot transform to B′. Then a direct straight line solution is not possible due to the constraint of constant link length.

The only remaining solution is for the link to rotate as part of the transformation. Consider first the rotation of link AB. The EL equation (37) allow for pure rotations about A, B, or a common centre along the link. Likewise for link A′B′.

The rotation can occur from either link AB (figure 5(a)) or link A′B′ (figure 5(b)). After the link rotates to a critical angle, it can then travel in a straight line. The extremals are broken in that they involve matching up a piece consisting of pure rotation with a piece consisting of pure translation of the end points of the link. Where the pieces match they must satisfy the corner conditions (18*a*) and (18*b*). This means that the end points cannot suddenly change direction, a situation which is only satisfied by a straight line trajectory that lies tangent to the circle of rotation.

From figure 4, we see that a straight line transformation exists only when an angle between a link and one of the straight line trajectories reaches $\pi/2$.

The critical angle that link AB must rotate is then determined by the point where a line drawn from B′ is just

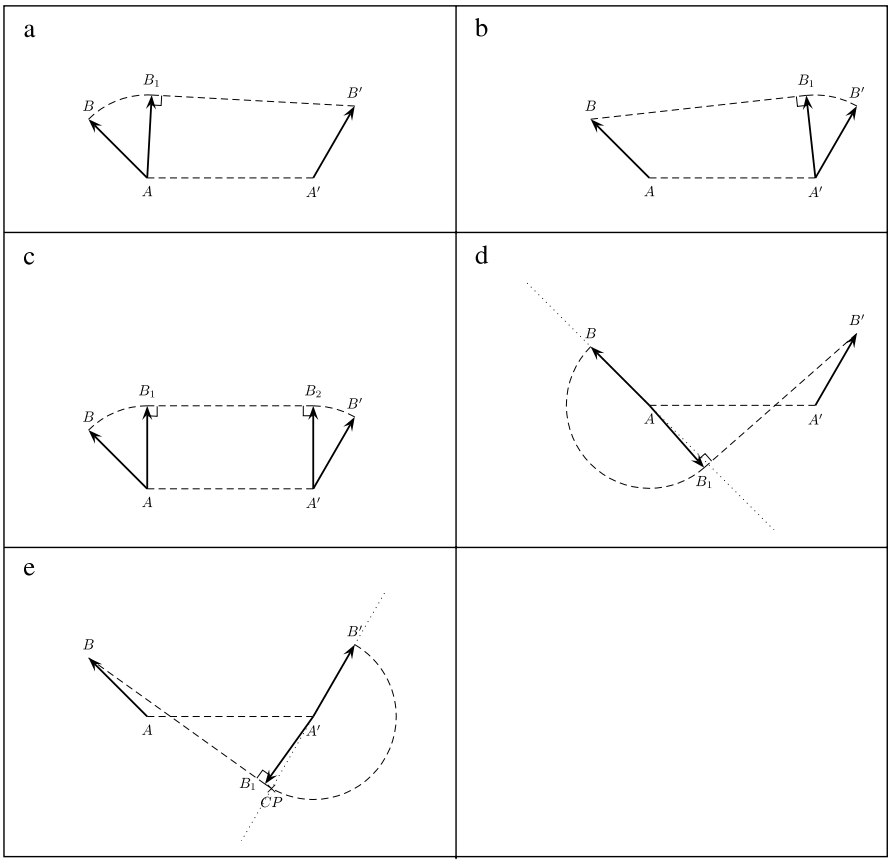


Figure 5. Transformations between two links involving broken extremals consisting of rotation and translation. (b) The global minimum, with shortest distance travelled during the transformation. (a), (c), (d) Local minima. (e) is extremal, but not minimal as the trajectory of arc $\mathrm{B'\widehat{B}_1}$ passes through a conjugate point.

tangent to the unit sphere centred at point A, point B_1 in figure 5(a). There is generally a different critical angle if the rotation occurs at link $\mathrm{A'B'}$ as in figure 5(b). It is shown in appendix B that in general the critical angle is determined by drawing the tangent to a circle or sphere about one of the link ends.

If the rotation was about a common centre, we see that one or another of the link ends would violate a corner condition, so the rotation must be about one of the link ends.

According to equations (26) and (27), the matrix P has a determinant of zero due to the parametric formulation in the problem and so is not positive definite. To show that the transformations in figures 5(a) and (b) are indeed minimal, we need to then express the problem in non-parametric form. To do this, let the independent variable be the angle θ of the link with the vertical. Then the displacement x along the line $\overline{\mathrm{AA'}}$ is the unknown function of θ to be determined by minimizing the total arc-length travelled. This distance can be written as

$$\mathcal{D}[x] = \int_{\theta_0}^{\theta_1} \mathrm{d}\theta \left(\sqrt{x'^2 + 2x' \cos\theta + 1} + \sqrt{x'^2} \right).$$

In this formulation, the scalar quantity $\mathrm{P}(\theta) = \mathcal{L}_{x'x'}$ becomes

$$\mathrm{P}(\theta) = \frac{\sin^2\theta}{\left(x'^2 + 2x' \cos\theta + 1\right)^{3/2}}$$

which is always >0 except for the isolated point $\theta = 0$; in particular, it is positive along the extremal trajectory, which is necessary for a minimum. So we conclude that the transformation with the smaller angle of rotation in figure 5(b) is here the global minimum, and the other transformation (figure 5(a)) is a local minimum.

Figure 5(e) is also an extremal trajectory, satisfying corner conditions, and with positive definite P. However, it is not a local minimum because the trajectory passes through a conjugate point (denoted by point CP, where the dotted line along $\overline{\mathrm{A'B'}}$ meets the great circle about $\mathrm{A'}$). According to the results in section 2.4.2, if the extremal trajectory (a great circle) traverses an angle larger than π radians, it passes through a conjugate point and thus becomes unstable to long-wavelength perturbations. Transformations involving rotations about points B or $\mathrm{B'}$ in figure 5 both have conjugate points and so are not minimal.

The transformation in figure 5(c) does not pass through a conjugate point and so is in fact another local minimum. The part of the extremum along the straight line section of the trajectory has no conjugate points as discussed above.

3.3. Systematically exploring transformations by varying link positions

We can investigate what happens to the minimal transformation when one of the link positions or angles is varied with respect

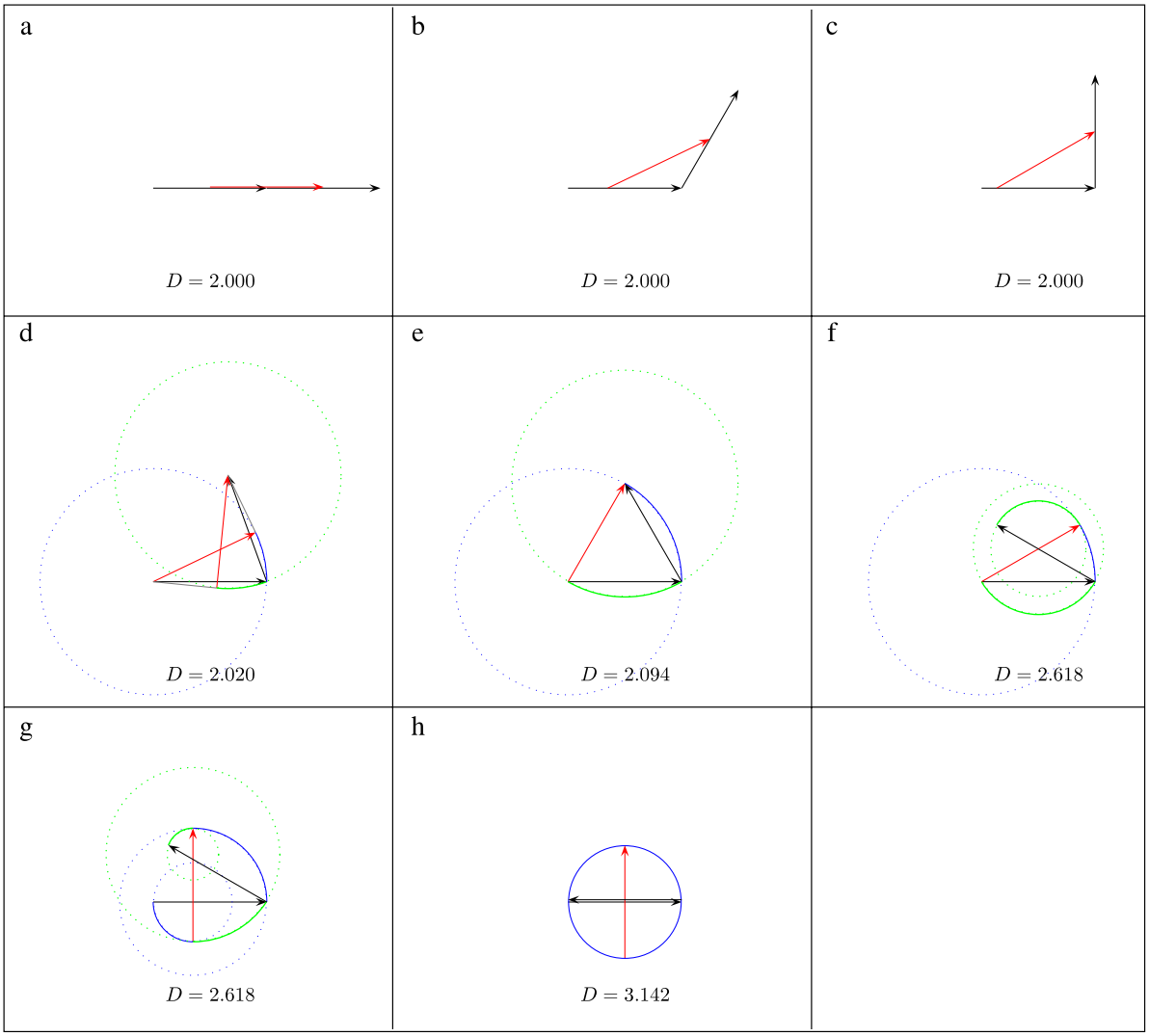


Figure 6. Successive transformations between two links made by rotating a link so that there is a progressively larger angle between the links as vectors (or smaller angle made between them as lines). The two boundary conditions (the initial and final conditions) are shown as black links, and an intermediate state is shown as a red link or links. The arcs traced out by the end points are shown in blue or green, while straight line motions when they are not along the links themselves are shown in grey. The distance travelled over the course of the transformation is given below each figure.

to the other. Let us start by putting the two links head to tail as shown in figure 6(a). The distance between them is 2 by simple translation of link end points.

We can now increase the angle between the two vectors by rotating the right link for example, as in figures 6(b)–(h). So long as the angle between the two vectors is less than 90°, one link may slide along another and the distance is unchanged (figures 6(a)–(c)). This is a special case of the transformations shown in figure 4 (compare for example figure 6(b) with the middle three unlabelled links in this figure).

Beyond 90°, however, the transformation must include rotation. Figure 6(d) has an angle of 150°. The minimal transformation first rotates, for example with the tail of the horizontal black arrow fixed, and the head tracing out the blue arc, until the critical angle is reached, where a straight line made from the final arrowhead (at the top of the figure) is just tangent to the circle made by the blue arc. This state is indicated by a red link in figure 6(d). The link then translates to its reciprocal position at the opposite end of the bow-tie, denoted by a second red link (cf. also figure 4(b)). At this point the arrowhead has completed the transformation. Finally, the tail rotates into its final position. The total distance travelled is slightly larger than 2.

When the angle between the vectors is 120° as shown in figure 6(e), the transformation consists of pure rotations. Taking the initial state to be the horizontal black vector, the link first rotates about its fixed tail, the head tracing out the blue arc, until the link reaches the state shown in red, where the position of the arrowhead has reached its final end point. Then the link rotates about its head until the position of the tail reaches the final state.

When the angle between the links is larger than 120°, as shown in figures 6(f) and (g), the transformation must involve rotation about an internal point along the link. Let points A

and B denote the tail and head of the link respectively. If an infinitesimal rotation $\Delta\theta$ occurs about an internal point P, the increment in distance travelled is

$$\Delta\mathcal{D} = |\mathbf{r}_{B/P}|\Delta\theta + |\mathbf{r}_{B/A}|\Delta\theta = \Delta\theta$$

which is *independent* of the position of the instantaneous centre of rotation (ICR). This means that there is an infinity of transformations all giving the same distance, depending on the time-dependence of the ICR. Two simple alternatives with only two discrete positions of ICR are shown in figures 6(f) and (g). Specifically, in figure 6(f), the horizontal black vector first rotates about its tail to the red configuration, which is a mirror image of the final black vector. Then rotation is about an internal point determined by the intercept of the red vector with the final black vector, with end points tracing out the green arcs. In figure 6(g) the two ICRs are both internal and determined by the intercepts of the initial and final states with the red vector shown.

Figure 6(h) depicts the transformation for overlapping, opposite pointing vectors. Rotation can now only occur about one point in the centre of the vectors.

Figure 7 illustrates what happens when one of the links is translated with respect to another, starting from two different scenarios shown in figures 7(a) and (e). In 7(a), the tail of the vertical link is displaced $(1/3, -1/3)$ with respect to the tail of the horizontal link. The minimal transformation is a pure rotation by $\pi/2$.

In figure 7(b), the tail of the vertical link is now displaced to $(2/3, -1/3)$. Pure rotations again give a distance of $\pi/2$. Rotation about a point on the horizontal link that is equidistant from both arrowheads transforms the initial arrowhead to the final (red intermediate state). Then rotation of the tail about the arrowhead transforms to the final state.

In figure 7(c), the minimal transformation first involves a translation by sliding the arrowhead along the vertical, until the arrowheads overlap (red intermediate state). The tail end of the link then rotates into place.

In figure 7(d), straight lines from the end points will not satisfy the obtuse condition in section 3.2, so the transformation must involve rotations. Here a straight line transformation takes the link almost to the final state. It then must undergo a small rotation to complete the transformation. Seen in reverse, the vertical arrow must rotate to a critical angle determined by the criterion in section 3.2, before the link can finish the transformation by pure translation.

Figure 7(e) is figure 6(f) once again. The final condition (the tilted link) will be systematically changed by translating it vertically away from the horizontal link (which we choose arbitrarily as the initial configuration).

In figure 7(f) the tilted link is translated a distance $1/3$ vertically. The transformation can be achieved by rotating the horizontal link about a point equidistant from both arrowheads to the red intermediate configuration. The link then rotates about the arrowhead into the final configuration. The distance is still the angle rotated for the reasons mentioned above in the context of figures 6(f) and (g), $\theta = (150/180)\pi$, which is unchanged from 7(e). In fact, so long as the arrowhead can be reached by rotation (the translated distance is less than d where d is the solution to $d^2+d+1-\sqrt{3} = 0$ for this angle), then the distance will be unchanged. The transformation at the critical distance is shown in figure 7(g). The rotations now occur about the end points: the tail and head of the link.

In figure 7(h) the translated distance is now equal to 1. The transformation first consists of a rotation about the tail to a critical angle (blue arc and red intermediate state), then a translation much like that in figure 4 (grey straight lines between red intermediate states), and finally a rotation about the head (green arc) to the final configuration.

4. Two-link chains

We now consider the next simplest case of two links (three beads). The Lagrangian now reads

$$\mathcal{L}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dot{\mathbf{r}}_1, \dot{\mathbf{r}}_2, \dot{\mathbf{r}}_3) = \sqrt{\dot{\mathbf{r}}_1^2} + \sqrt{\dot{\mathbf{r}}_2^2} + \sqrt{\dot{\mathbf{r}}_3^2} - \tfrac{1}{2}\lambda_{12} \times \left((\mathbf{r}_2 - \mathbf{r}_1)^2 - 1\right) - \tfrac{1}{2}\lambda_{23}\left((\mathbf{r}_3 - \mathbf{r}_2)^2 - 1\right) \quad (38)$$

which has EL equations (cf equations (15*a*)–(15*c*))[5]:

$$\dot{\hat{\mathbf{v}}}_A + \lambda_{AB}\mathbf{r}_{B/A} = 0 \quad (39a)$$

$$\dot{\hat{\mathbf{v}}}_B - \lambda_{AB}\mathbf{r}_{B/A} + \lambda_{BC}\mathbf{r}_{C/B} = 0 \quad (39b)$$

$$\dot{\hat{\mathbf{v}}}_C - \lambda_{BC}\mathbf{r}_{C/B} = 0. \quad (39c)$$

The corner conditions (18*a*), (19) imply

$$\hat{\mathbf{v}}_i\left(t^-\right) = \hat{\mathbf{v}}_i\left(t^+\right)$$

so the direction of motion cannot suddenly change, unless along one part of the extremal the velocity of point i is zero (the point is at rest), where its direction $\hat{\mathbf{v}}$ is then undefined.

The boundary conditions described in section 2.3.1 hold as well, so the end points can either be at rest, move in straight lines, or purely rotate. This gives $3 \times 3 = 9$ possible scenarios to investigate here, many of which can readily be ruled out. For example consider the states in figure 8(a). Because A and A′ are in the same position, rotation and translation of A are ruled out and point A remains at rest, leaving three scenarios for the other end point C. However since C and C′ are at different positions and ABC are along a straight line, C cannot remain at rest initially, leaving either translation or rotation for point C.

Suppose C translates towards C′ as in figure 8(b). Then $\dot{\hat{\mathbf{v}}}_C = 0$ and from (39*c*), (39*b*) $\lambda_{BC} = 0$ and $\dot{\hat{\mathbf{v}}}_B = \lambda_{AB}\mathbf{r}_{B/A}$. B cannot move in a straight line without moving point A, so $\lambda_{AB} \neq 0$ and thus B must rotate about point A. The transformation then proceeds as in figure 8(b) until B reaches B′ and C reaches C″. Then however if C″ were to rotate to C′, the trajectory would violate corner conditions at point C″. Therefore, the direction of translation of C must not be directly to C′ but must be tangential to the arc $\overset{\frown}{\mathrm{C'C''}}$ as in figure 8(c).

The reverse of this transformation is allowable as well, as can be seen by swapping the labels ABC → A′B′C′. Here C

[5] The links have length 1 in our dimensionless formulation, so the vectors $\mathbf{r}_{B/A}$ and $\mathbf{r}_{C/B}$ could also have been written as unit vectors $\hat{\mathbf{r}}_{B/A}$ and $\hat{\mathbf{r}}_{C/B}$.

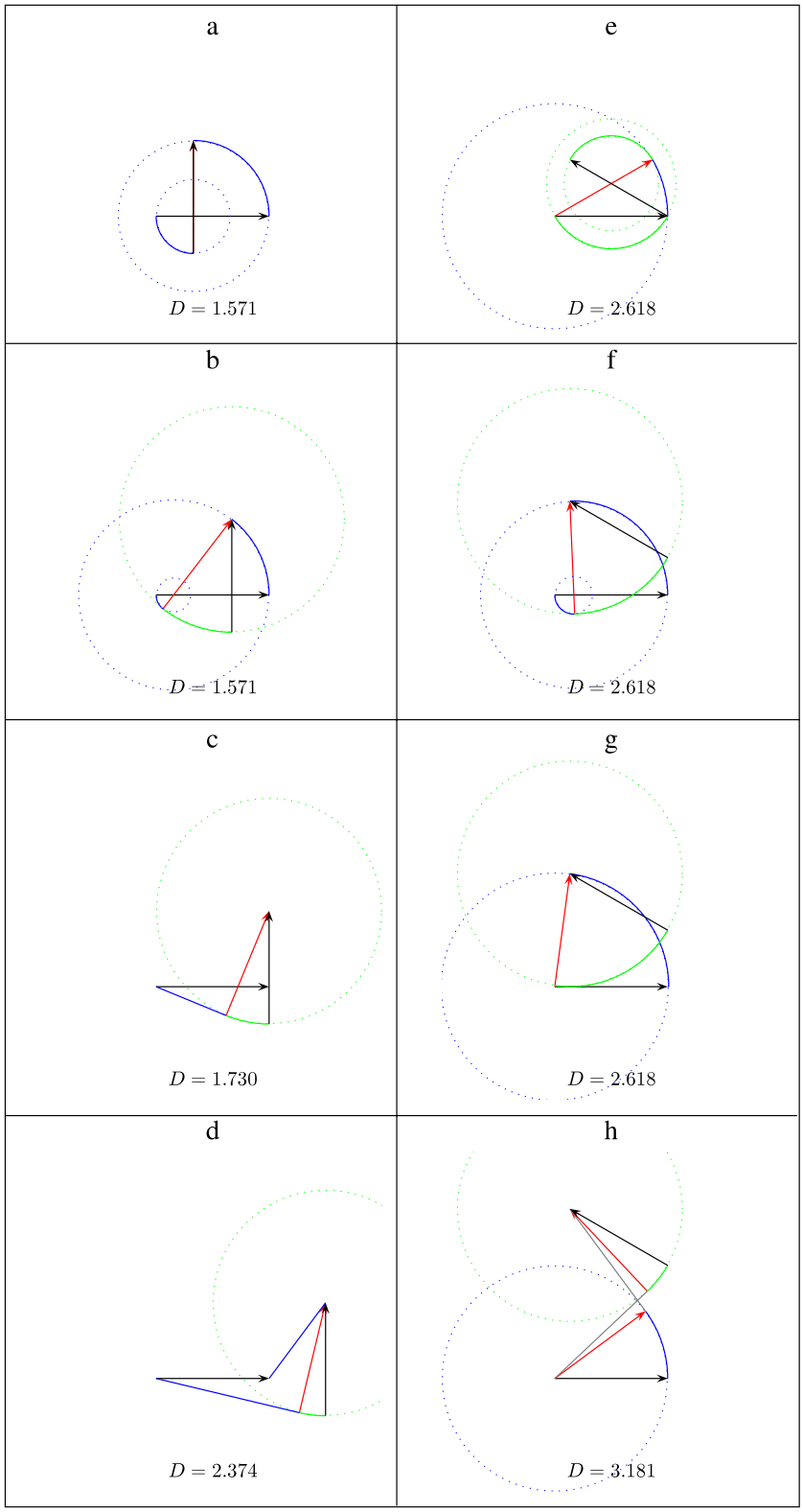


Figure 7. Successive transformations between two links made by translating one link with respect to the other. In (a)–(d) the initial and final configurations are perpendicular, while in (e)–(h) they are at an angle of 150° to each other. Note that the distances in (e)–(g) are all the same, even though the end points of the links are at varying distances from each other.

first rotates to the critical angle θ shown in figure 8(d) and then translates to C′.

In fact one can see that links BC and B′C′ along with lines $\overline{\mathrm{BB}'}$ and $\overline{\mathrm{CC}'}$ form a quadrilateral as in figure 5, with the same consequences for rotation to a critical angle. For the links in figure 8 the situation is symmetric, so rotation can occur at the beginning or end of the transformation. Figure 9(a) shows an example with this symmetry broken, so that the

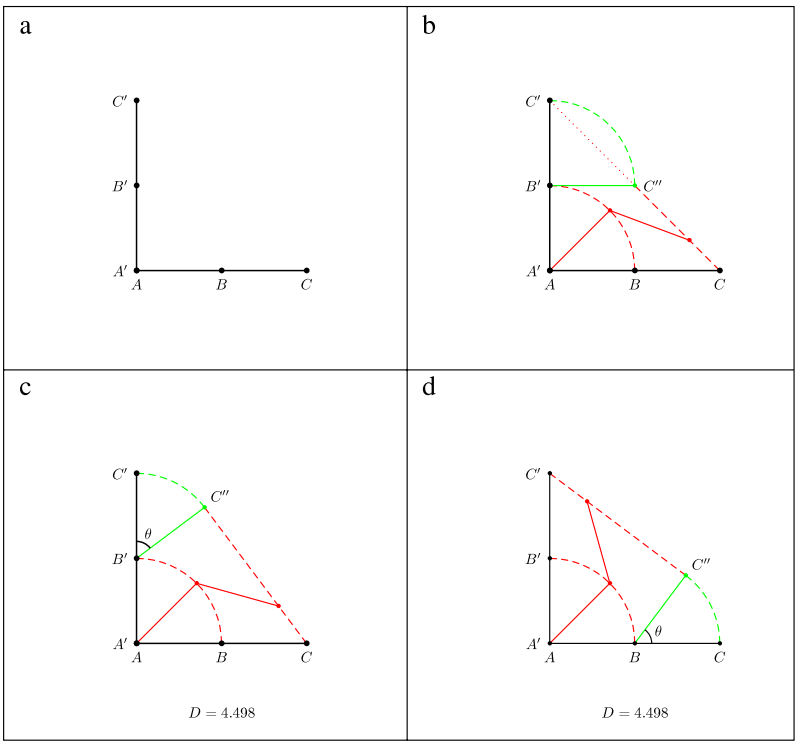


Figure 8. (a) Initial and final states for a chain of two links. The transformation in (b) is non-extremal because it violates a corner condition at C″. (c) and (d) are degenerate minima—rotations occurring about B′ or B both have the same length. Intermediate states shown in red have opposite convexities in (c) and (d).

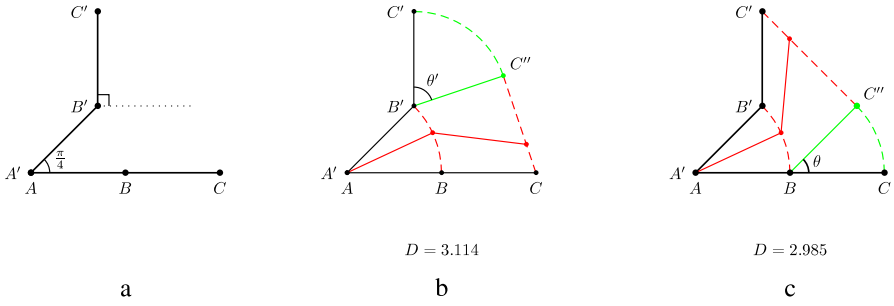


Figure 9. (a) Initial and final states for a polymer of two links. The angle between $\overline{AB}$ and $\overline{A'B'}$ is $\pi/4$. The minimal transformations in (b) and (c) are now no longer degenerate. (c) The global minimum.

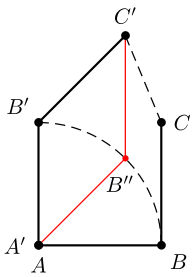


Figure 10. A transformation between two states of opposite convexity: ABC has convexity down and right, while A′B′C′ has convexity up and left. There is no extremal transformation in the plane that can connect them, without some apparent violation of corner conditions.

distance is different depending on where the rotation occurs, as in figures 5(a) and (b). In this case, the transformation in figure 9(c) has the minimal distance, and that in figure 9(b) is subminimal. Extensions of the transformation in figure 9 to large numbers of links were explored in [1].

4.1. Transformations involving a change in convexity

Transformations between configurations with opposite convexity involve motion out of the plane, even if the initial and final states lie in the plane. If the transformation is constrained to lie in plane, the trajectories of some points will be non-monotonic—these points must move farther away from their final positions before approaching them. We illustrate these ideas with some examples below.

Consider the initial and final states in figure 10. We again imagine B rotating to B′. If C were to translate to C′ one would have the intermediate configuration A′B″C′. Now C′ and A′ must remain at rest to satisfy corner conditions. Then the only way to finish the transformation is for B″ to rotate about the

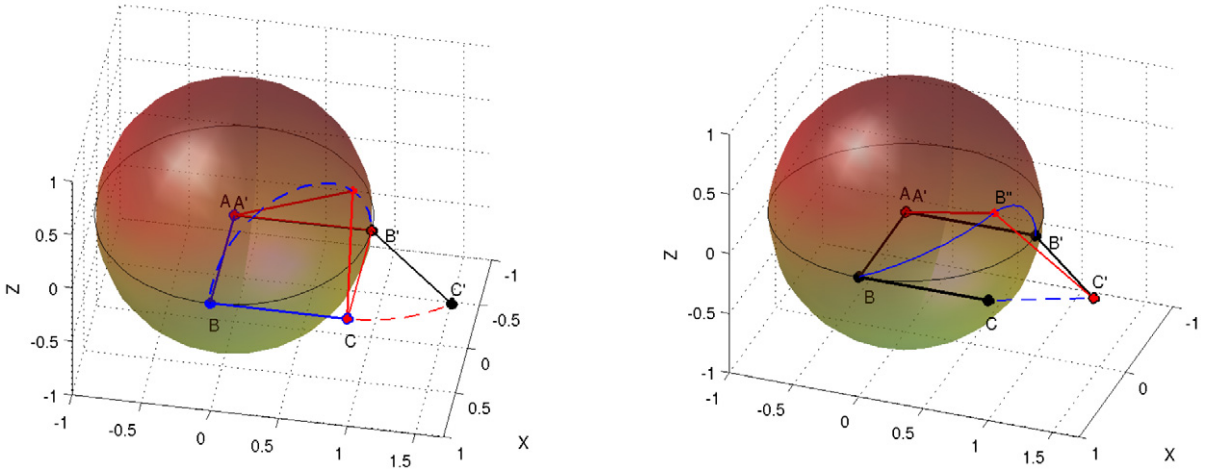


Figure 11. Subminimal (a) and minimal (b) transformations for the boundary conditions in figure 10. The distances for each transformation are approximately $3.007L^2$ and $2.576L^2$ respectively. Transformation (a) proceeds from ABC by first rotating B to B′ about axis $\overline{\mathrm{AC}}$, then rotating C about point B′. Transformation (b) proceeds from ABC by simultaneously translating C to C′ while rotating B about A on a great circle to point B″. Finally point B rotates from B″ to B′ about axis A′C′.

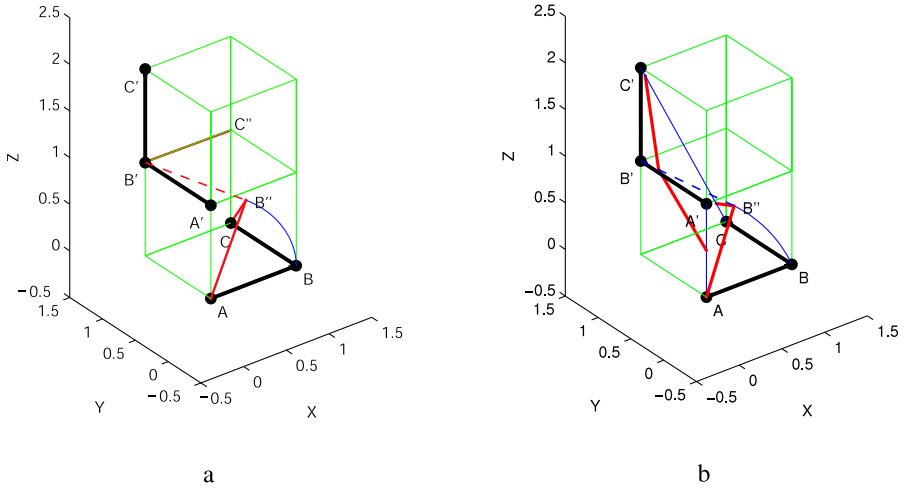


Figure 12. (a) Subminimal transformation and (b) minimal transformations between ABC and A′B′C′ (see text).

axis $\overline{\mathrm{A'C'}}$; however, then the trajectory of B violates corner conditions and so is not extremal. In appendix C we take up the issue of minimal transformations for this case when the links are *constrained* to lie in a plane.

We thus seek a point B″ and resulting trajectory $\overrightarrow{\mathrm{BB''B'}}$ such that arc $\widehat{\mathrm{BB''}}$ satisfies corner conditions with arc $\widehat{\mathrm{B''B'}}$.

One solution is to effectively place B″ at position B′ by considering the boundary condition with C at rest (and A at rest). Then B rotates to B′ about axis $\overline{\mathrm{AC}}$, and the trajectory of B lies on a circle defined by the intercept of two unit spheres centred at A and C. The sphere about A is drawn in figure 11 as a visual aid. Along arc $\widehat{\mathrm{BB'}}$ both $\lambda_{\mathrm{AB}} \neq 0$ and $\lambda_{\mathrm{BC}} \neq 0$. Once in configuration A′B′C, C can then undergo rotation about B′ to C′, with A′ and B′ stationary.

The transformation in 11(a) is a local minimum in distance; however, it is not the global minimum. A shorter distance transformation can be seen by considering the reverse transformation. Imagine A′ and C′ stationary while B′ rotates about axis $\overline{\mathrm{A'C'}}$ in figure 11(b). This rotation of B′ follows a circular trajectory defined by the intercept of two unit spheres centred at A′ and C′. The rotation occurs until point B″, which is the point where the above circle is tangent to a great circle on the unit sphere about A and passing through B. The arc $\widehat{\mathrm{BB''}}$ is a great circle because this is a geodesic for point B given A is fixed, which follows from the Euler equations (39*b*) and (39*c*) when $\lambda_{\mathrm{BC}} = 0$. The great circle is defined by the plane containing the points A, B, and B″.

The angle between the (variable) vector $\overrightarrow{\mathrm{BC}}$ of link BC and the tangent to the arc $\widehat{\mathrm{B'B''}}$ is always $\pi/2$, so once the corner

condition is met, point C on link BC can move in straight line motion from C′ to C while B moves on the great circle from B″ to B. That is, the quadrilateral criterion of section 3.1 is met for □BB″C′C.

To find point B″, let its position be $\mathbf{r}_{B''} = (x_o, y(x_o), z(x_o))$. The great circle is defined by the plane passing through the points A, B, and B″. This plane has normal $\mathbf{n} \equiv \overrightarrow{AB} \times \overrightarrow{AB''} = (1,0,0) \times (x_o, y(x_o), z(x_o)) = (0, -z(x_o), y(x_o))$. At the point B″ the normal is orthogonal to the tangent vector of the circle defined by rotation about the AC′ axis. This tangent vector is $\hat{\mathbf{t}} = \partial\mathbf{r}/\partial s = x_s(1, y_x, z_x)$ by the chain rule. At B″, $\hat{\mathbf{t}} \cdot \mathbf{n} = 0$, or

$$-z(x_o)y_x(x_o) + y(x_o)z_x(x_o) = 0. \tag{40}$$

The functions $y(x)$ and $z(x)$ are defined by the intercept of two unit spheres centred at $(0,0,0)$ and $(1/\sqrt{2}, 1+1/\sqrt{2}, 0)$, giving

$$\begin{aligned} y(x) &= 1 - \frac{\sqrt{2}}{2+\sqrt{2}}x \\ z(x) &= \sqrt{1 - x^2 - y(x)^2}\,. \end{aligned} \tag{41}$$

Together (40) and (41) give

$$\mathbf{r}_{B''} = \begin{pmatrix} \sqrt{2}-1 \\ 2(\sqrt{2}-1) \\ \sqrt{2(5\sqrt{2}-7)} \end{pmatrix}.$$

The distance travelled along arc $\overset{\frown}{BB''}$ is $\theta_{BB''}$, where $\cos\theta_{BB''} = x_o = \sqrt{2}-1$. The distance travelled along arc $\overset{\frown}{B''B}$ can similarly be shown to be $r\theta_{B''B'} = \sin(\pi/8)\cos^{-1}(2\sqrt{2}-3)$. Adding the distance $\overline{CC'}$, the total (minimal) distance is thus $\mathcal{D} = 2.576$. There is of course a degenerate solution to the above with $z \to -z$.

4.2. *Transformations with initial and final states in 3D*

We now give a representative example where the initial and final configurations do not lie in the same plane, as shown in figure 12. Because $\overline{AB} \perp \overline{AA'}$ and $\overline{BC} \perp \overline{CC'}$, neither A nor C will rotate about B as part of the transformation. Nor can ABC simultaneously translate directly to A′B′C′, because for example quadrilateral □AA′B′B does not satisfy the rule of opposite angles $\geqslant \pi/2$, so link AB cannot slide (translate) to A′B′.

This leaves three options for the initial stages of the transformation.

(1) A translates, B rotates, C remains fixed. B then rotates about C in the CBB′ plane. The initial direction of motion of B is then $\hat{\mathbf{v}}_B = (-\hat{\mathbf{i}} + \hat{\mathbf{k}})/\sqrt{2}$; however, then $\hat{\mathbf{v}}_A$ can only move backward to preserve link length ($\hat{\mathbf{v}}_A = -\hat{\mathbf{k}}$), similar to figure A.1. This rules out case (1).

(2) A remains fixed, B rotates, C remains fixed. B then rotates towards B′ about axis $\overline{AC}$ until it reaches a critical angle where line $\overline{B''B'}$ is tangent to its circular trajectory (see figure 12(a)). At this point the quadrilateral □B″CC′B′ does not have opposite obtuse angles, so a straight line transformation to A′B′C′ is not possible. It is possible to transform to a configuration A′B′C″, where C″ is at position (1, 1, 1) and angle ∠B′C″C = π/2, so that $\hat{\mathbf{v}}_C = \hat{\mathbf{k}}$. Then the transformation is completed by a π/2 rotation of C″ about B′. This transformation is subminimal.

(3) A remains fixed, B rotates, C translates. In this case, B rotates toward B′ in the BAB′ plane, while C translates to C′, until the state AB″C″ is reached (see figure 12(b)). State AB″C″ can be found as follows. Because the rotation of B is about the axis $(0, -1/\sqrt{2}, 1/\sqrt{2})$, the position $\overrightarrow{AB''}$ of B″ after rotation of the (critical) angle θ is $(\cos\theta, \sin\theta/\sqrt{2}, \sin\theta/\sqrt{2})$. This angle is then determined by the condition $\overrightarrow{AB''} \cdot \overrightarrow{B''B'} = 0$, where $\overrightarrow{B''B'} = \overrightarrow{AB'} - \overrightarrow{AB''}$. The solution to this condition is simply $\theta = \pi/4$. The location of C″ is then determined from the condition that the link length from B″ to C″ is one: $|\overrightarrow{B''C''}| = 1$, where $\overrightarrow{B''C''} = \overrightarrow{AB''} + t\overrightarrow{CC'}$. Solving this condition for t gives the position of C″ as $(\frac{3+\sqrt{2}}{5}, 1, \frac{2(2-\sqrt{2})}{5})$. At this point the quadrilateral □B′B″C″C′ has opposite obtuse angles, and quadrilateral □AB″B′A′ has opposite angles = π/2, so it is in a bow-tie configuration as in the end point configurations in figure 4. Therefore all points AB″C″ can translate from this intermediate state to their final positions A′B′C′. The total distance travelled is $\theta + |AA'| + |CC'| + |B''B'|$ or $\mathcal{D} = 2 + \pi/4 + \sqrt{5} \approx 5.022$. The reverse of this transformation is also possible, where point B′ rotates about A′ in the plane B′AB, while C′ translates along $\overrightarrow{C'C}$. Inspection reveals that the distance covered is the same as the forward transformation.

5. Limit of large link number

From the transformation discussed in section 4.1, we see that if both ∠ABC and ∠A′B′C′ were π/2 as in figure 13(a) then the transformations in figures 11(a) and (b) would become degenerate, having distance $\mathcal{D} = \pi/\sqrt{2}$. The transformation is completed by a single rotation about axis $\overline{13}$.

We can now examine the effect of increasing the link number. Let the number of links increase to four, and let us preserve the symmetry that is present about the horizontal axis in figure 13(a), so the initial and final states become an octagon (figure 13(b)). In the limit $N \to \infty$, the figure becomes a circle.

If we separated the links in figure 13(a) by some distance in the y direction (perpendicular to axis $\overline{13}$), then the minimal transformation involves the same rotation of 2 about axis $\overline{13}$ until a critical angle θ_c, after which all three points 123 can translate in straight lines to 1′2′3′. In the same fashion, the minimal transformation for the octagonal transformation in figure 13(b) involves a rotation of point 3 out of the plane about axis $\overline{24}$ to a critical angle θ_c at which the point is located at position 3″. Once this critical angle is reached, point 3 translates in a straight line from 3″ to 3′.

Because points 1 and 5 are stationary to satisfy corner conditions, points 2 and 4 must move in great circles about points 1 and 5. However, points 2 and 4 cannot finish the transformation by moving on great circles. At the configuration 1′2″3′4″5′ in figure 13(b), point 3 has finished

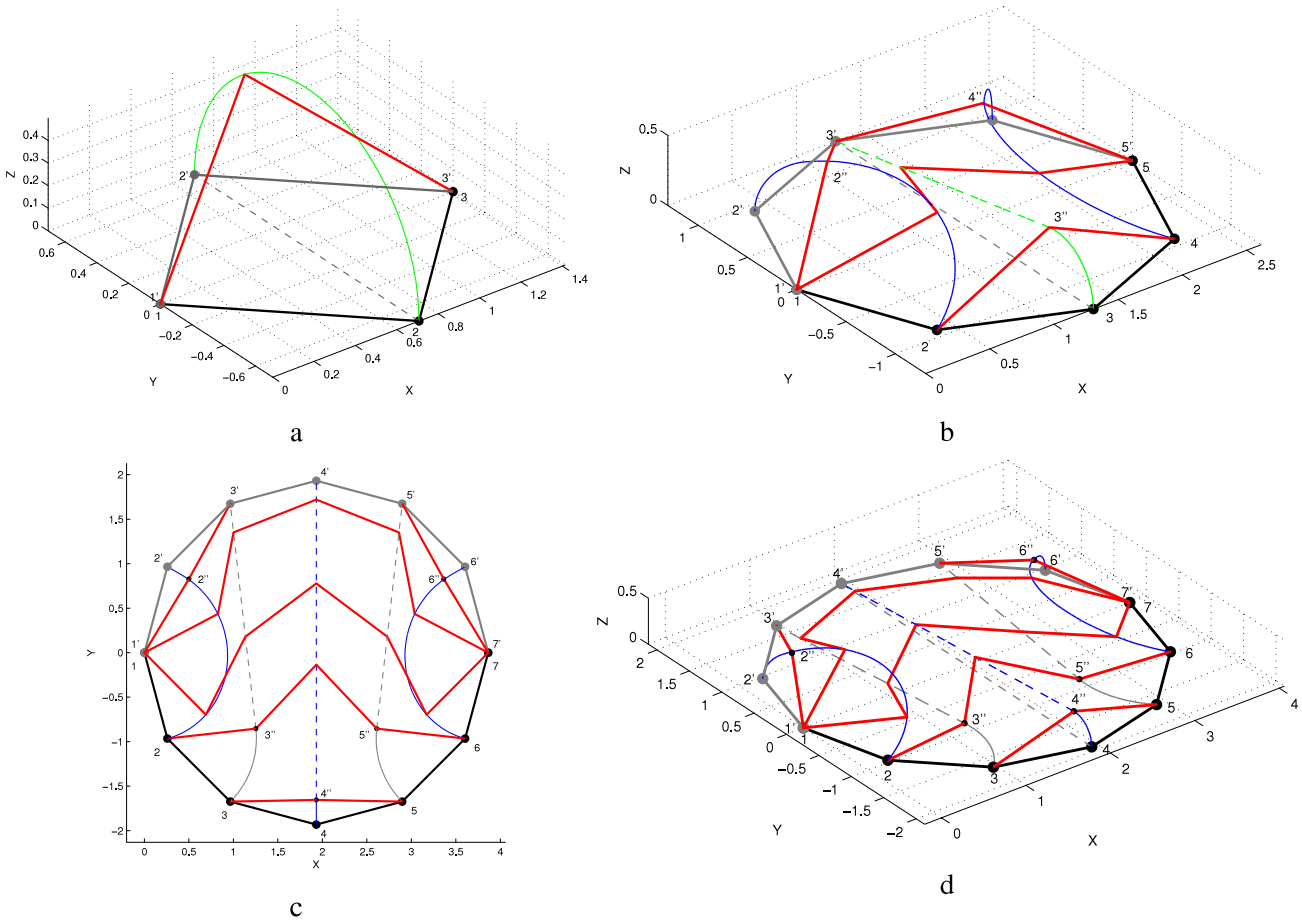


Figure 13. Examples of transformations between initial and final states of opposite convexity, for increasing numbers of links. (a) The transformation for $N = 2$ links. (b) $N = 4$ and initial and final state form an octagon. (c), (d) $N = 6$ and initial and final states form a dodecagon. (c) Top view. (d) View in perspective. Rotations are shown as solid colour lines (either green or blue). Translations are shown as dashed lines. The grey dashed lines underneath $\overline{3''3'}$ in (b) and $\overline{4''4'}$ in (d) are shown only to illustrate that these lines are above the plane.

the transformation, but points 2 and 4 have not. To satisfy corner conditions at the points $2''$ and $4''$, the great circles must be out of the plane as well. At points $2''$ and $4''$, the transformation finishes with rotations about axes $\overline{1'3'}$ and $\overline{3'5'}$. The total distance $\mathcal{D} \approx 7.93$.

Of course the time reverse of this transformation (equivalent to swapping primed and unprimed labels) is also a minimal transformation, as is the transformation obtained by reflection about the $z = 0$ plane.

Now consider increasing the chain to six links, so the combination of $\mathbf{r}_i(0)$ and $\mathbf{r}_i(T)$ becomes a dodecagon (12-sided polygon, see figures 13(c) and (d)). As before the midpoint vertex (here $\mathbf{r}_4$) must rotate out of the plane about axis $\overline{35}$ to a critical angle θ_c before translating in a straight line to $\mathbf{r}_{4'}$. This critical angle is where $\overrightarrow{34''} \cdot \overrightarrow{4''4'} = \overrightarrow{54''} \cdot \overrightarrow{4''4'} = 0$. The quadrilaterals $\square 22'3'3$ and $\square 655'6'$ are of the type in figure 5, so point 3 must rotate about $\mathbf{r}_2(0)$ to a critical angle where $\overrightarrow{23''} \cdot \overrightarrow{3''3'} = 0$, and likewise for point 5.

While point 3 rotates to its critical angle, point 4 translates along line $\overline{4''4'}$. Points $\mathbf{r}_1(0)$ and $\mathbf{r}_7(0)$ overlap with $\mathbf{r}_1(T)$ and $\mathbf{r}_7(T)$ and so remain fixed to satisfy corner conditions. After point 3 has reached its critical angle, it can translate along $\overline{3''3'}$ as point 2 rotates about $\mathbf{r}_1$. However, to satisfy corner conditions at point $2''$, the rotation cannot remain in the x–y plane. Point $\mathbf{r}_{2''}$ is determined as the point where $\hat{\mathbf{t}} \cdot \mathbf{n}_{\text{plane}} = 0$, where $\hat{\mathbf{t}}$ is the tangent to the arc $\widehat{22''}$ defined by rotation about axis $\overline{13'}$, and $\mathbf{n}_{\text{plane}}$ is the normal to the plane $122''$, i.e. $\mathbf{r}_{2/1} \times \mathbf{r}_{2''/1}$. The same process holds for point 6. These critical points and some intermediate states for the transformation are shown in figure 13(d). The total distance covered by the transformation is $\mathcal{D} \approx 16.3$.

It is sensible to consider the total length of chain as fixed to say $L = 1$, and to let the link length ds_N for the chain of N links be determined by $N ds_N = L$. Because distances scale as ds_N^2, the $N = 2, 4, 6$ cases have $\mathcal{D}_2 \approx 0.555L^2$, $\mathcal{D}_4 \approx 0.496L^2$, $\mathcal{D}_6 \approx 0.445L^2$. Note that this distance decreases with increasing number of links: the constraints on the motion of the various beads during the transformation are relaxed as the number of links is increased.

We can then imagine resting a piece of string on a table in the shape of a semicircular arc, and then asking how one can move this string to a facing semicircle of opposite convexity. So long as the string has some nonzero persistence length ℓ_P, the transformation of minimal distance must involve lifting the string off of the table to change its local convexity. The vertical height the string must be lifted (see figure 13(d)) is of order

$\sim \sin(\pi \ell_P/L) \sim \ell_P/L$, which goes to zero for an infinitely long chain.

As the number of links $N \to \infty$, some simplifications emerge. In particular the contribution to the total distance due to rotations becomes negligible, and the translational component dominates. To see this note that the distance due to straight line motion scales as

$$\mathcal{D}(\text{st. line}) \sim \mathrm{d}s\, NL \sim L^2$$

while the distance travelled during rotations scales as

$$\mathcal{D}(\text{rot.}) \sim \mathrm{d}s\, N(\overline{\theta_c}\mathrm{d}s) \sim L^2/N$$

where we assume the worst case scenario where an extensive number of links must rotate before translating. Because translation dominates the distance as $N \to \infty$, the distance travelled converges to L times the mean root square distance (MRSD), i.e.

$$\begin{aligned}\mathcal{D}_\infty &\to \mathrm{d}s \sum_{i=1}^{N+1} |\mathbf{r}_i(T) - \mathbf{r}_i(0)| \\ &= L\frac{1}{N}\sum_i \sqrt{(\mathbf{r}_{Bi} - \mathbf{r}_{Ai})^2} \\ &= L\,(\text{MRSD})\,. \end{aligned} \tag{42}$$

The MRSDs for the examples in figures 13(b) and (d) are $0.394L$ and $0.400L$ respectively, which are both less than the actual distances travelled (in units of L). In the limit $N \to \infty$, where the polygon becomes a circle, the distance converges to $\mathcal{D}_\infty = 4L^2/\pi^2 \approx 0.4053L^2$. For large N systems then, it is a good first approximation to use the MRSD for the distance.

The MRSD is always less than the root mean square distance (RMSD), except in special cases when they are equal. To see this, we can apply Hölder's inequality

$$\sum_{k=1}^{N} (g_k)^\alpha\, (h_k)^\beta \leqslant \left(\sum_{k=1}^{N} g_k\right)^\alpha \left(\sum_{k=1}^{N} h_k\right)^\beta$$

where $g_k, h_k \geqslant 0$, $\alpha, \beta \geqslant 0$, and $\alpha + \beta = 1$. With the specific identifications $g_k = (\mathbf{r}_B k - \mathbf{r}_A k)^2 \equiv \Delta\mathbf{r}_k^2$, $h_k = 1$, and $\alpha = \beta = 1/2$, we have directly

$$\frac{1}{N}\sum_k \sqrt{\Delta\mathbf{r}_k^2} \leqslant \sqrt{\frac{1}{N}\sum_k \Delta\mathbf{r}_k^2}.$$

For example, the RMSD for the circle configuration discussed above is $\sqrt{2}L/\pi \approx 0.4502L$, which is greater than the MRSD.

The fact that the distance converges for large N to MRSD rather than RMSD suggests that RMSD may not be the best metric for determining similarity between molecular structures, although it is ubiquitously used. This fact warrants future investigation—it has implications in research areas from structural alignment based pharmacophore identification [3–5] to protein structure and function prediction [6, 7].

It was shown in [1] that chains with persistence length characterized by some radius of curvature R have extensive corrections to the MRSD-derived minimal distance, which do not vanish as $N \to \infty$, but remain so long as R/L is nonzero. Likewise, chains that cannot cross themselves have non-local EL equations and extensive corrections to the minimal distance. Nevertheless, it is worthwhile to investigate some more complex polymers with MRSD as an approximate distance metric. We pursue this in the next section.

5.1. MRSD as a metric for protein folding

Here we examine the use of MRSD as a metric or order parameter for protein folding. To this end we adopt an unfrustrated C_α model of segment 84–140 of *src tyrosine-protein kinase* (src-SH3), by applying a Gō-like Hamiltonian [8–10] to an off-lattice coarse-grained representation of the *src-SH3* native structure (pdb 1fmk). Amino acids are represented as single beads centred at their C_α positions. The Gō-like energy of a protein configuration α is given by the following Hamiltonian, which we will explain term by term:

$$\begin{aligned}\mathcal{H}(\alpha|N) &= k_r \sum_{\text{bonds}} (r_\alpha - r_N)^2 + k_\theta \sum_{\text{triples}} (\theta_\alpha - \theta_N)^2 \\ &+ \sum_{n=1,3} k_\phi^{(n)} \sum_{\text{quads}} [1 - \cos(n \times (\phi_\alpha - \phi_N))] \\ &+ \epsilon_N \sum_{j \geqslant i+3} \left[6\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{10} - 5\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}\right] + \epsilon_{NN} \sum_{j \geqslant i+3} \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}. \end{aligned} \tag{43}$$

Adjacent beads are strung together into a polymer through harmonic bond interactions that preserve native bond distances between consecutive C_α residues. Here r_α and r_N represent the distances between two subsequent residues in configurations α and the native state N. As with other parameters in the Hamiltonian, the distances r_N are based on the pdb structure and may vary pair to pair. The angles θ_N represent the angles formed by three subsequent C_α residues in the pdb structure, and the angles ϕ_N represent the dihedral angles defined by four subsequent residues. The dihedral potential consists of a sum of two terms, one with period 2π and another with $2\pi/3$, which give *cis* and *trans* conformations for angles between successive planes of three amino acids, with a global dihedral potential minimum at $\phi_N \in [-\pi, \pi]$.

The parameters k_r, k_θ, and k_ϕ, are taken to accurately describe the energetics of the protein backbone: we used the values $k_r = 50$ kcal mol^{-1}, $k_\theta = 20$ kcal mol^{-1}, $k_\phi^{(1)} = 1$ kcal mol^{-1} and $k_\phi^{(3)} = 0.5$ kcal mol^{-1} for molecular dynamics (MD) simulations using the AMBER software package. For MD simulations using LAMMPS, we had used slightly different values: $k_r = 80$ kcal mol^{-1}, $k_\theta = 16$ kcal mol^{-1}, $k_\phi^{(1)} = 0.8$ kcal mol^{-1} and $k_\phi^{(3)} = 0.4$ kcal mol^{-1}.

The last line in equation (43) deals with non-local interactions, both native and non-native. If two amino acids are separated by three more along the chain ($|i - j| \geqslant 3$), and have one or more pairs of heavy atoms within a cut-off distance of $r_c = 4.8$ Å in the pdb structure, the amino acids are said to have a native contact. Then the respective coarse-grained C_α residues are given a Lennard-Jones-like 10–12 potential of depth $\epsilon_N = -0.6$ kcal mol^{-1} (-0.8 kcal mol^{-1} for LAMMPS

simulations) and a position of the potential minimum equal to the distance of the C_α atoms in the pdb structure. That is, σ_{ij} is taken equal to the native distance between C_α residues i and j if i–j have a native contact.

If two amino acids are not in contact, their respective C_α residues sterically repel each other ($\epsilon_{NN} = +0.6$ kcal mol^{-1}). Thus $\epsilon_{NN} = 0$ if i–j is a native residue pair, while $\epsilon_{\text{N}} = 0$ if i–j is a non-native pair. For non-native residue pairs, $\sigma_{ij} = 4$ Å.

In an *arbitrary* configuration α, two C_α residues i and j are considered to have formed a native contact if they have a distance $r_{ij} \leqslant 1.2\sigma_{ij}$. The results do not strongly depend on the specific value of this cut-off. The fraction of native contacts present in the particular configuration α is then defined as Q (or Q_α).

The MRSD of configuration α is found by aligning this configuration to the native structure, by minimizing MRSD over three translational and three rotational degrees of freedom.

Constant temperature molecular dynamics simulations were run for this system using both AMBER and LAMMPS simulation packages. The probability for the system to have given values of Q and MRSD within $(Q, Q + \Delta Q)$ and (MRSD, MRSD + ΔMRSD) is proportional to the exponential of the free energy $F(Q, \text{MRSD})$. Thus the free energy can be directly obtained by sampling, binning, and taking the logarithm:

$$F(Q_1, \text{MRSD}_1) - F(Q_2, \text{MRSD}_2) = -k_{\text{B}}T \log\left(\frac{p(Q_1, \text{MRSD}_1)}{p(Q_2, \text{MRSD}_2)}\right) \tag{44}$$

with $F(1, 0) = E_N$, the energy of the native structure.

Figure 14 shows the free energy surfaces obtained using the above recipe, for the AMBER (figure 14(a)) and LAMMPS (figure 14(b)) molecular dynamics routines. The temperature is taken to be the transition or folding temperature T_{F}, where the unfolded and folded free energies are equal.

Notice that $F(Q)$ is comparable for both as it should be; moreover, F(MRSD) is as well. However, the free energy surface plotted as a function of both Q and MRSD shows a marked difference. In addition to a native minimum, the LAMMPS routine has an additional minimum at $Q \approx 0.95$ and MRSD ≈ 8.4. The conformational states in this bin are closely related, with an average MRSD between them of 1.8 Å. We can take the most representative state in this bin as that which has a minimum MRSD from all the others in the bin (at $Q \approx .95$, MRSD ≈ 8.4): $\min_i(\sum'_{j\neq i} \text{MRSD}_{ij} / \sum'_{j\neq i}) \approx 1.6$ Å. Inspection reveals that this state is a mirror image of the pdb structure (see figure 14(b)): if we reflect this structure about one plane, and subsequently align this reflected structure to the pdb one, the MRSD is only 1.1 Å.

The discrepancy in free energy surfaces corresponding to the presence of a low energy mirror-image structure arises because the COMPASS class 2 dihedral potentials in the LAMMPS algorithm do not ascribe a sign to the angle ϕ, so the full range $[-\pi, \pi]$ is projected onto $[0, \pi]$. This gives the set of actual dihedral angles $\{\phi_i + \pi\}$ the same energy as the set $\{\phi_i\}$, so that the dihedral potentials have two minima rather than one, and thus a protein chain of the opposite chirality (a mirror image) is allowed and has the same energy as the pdb structure. We found that the CHARMM and harmonic dihedral styles do not have this problem; however, they have less versatile function forms, so that we favoured modifying the COMPASS dihedrals to define ϕ over its full range.

6. Conclusions

Analogously to the distance between two points, the distance between two finite length space curves is a variational problem, and may be calculated by minimizing a functional of two independent variables s and t, where s is the arc-length along the chain, and t is the 'elapsed time' during the transformation.

We derived the Euler–Lagrange (EL) equation giving the solution to this problem, which is a vector partial differential equation, with extremal solution $\mathbf{r}^*(s, t)$. We also derived the sufficient conditions for the extremal solution to be a minimum, through the Jacobi equation. Once the minimal transformation $\mathbf{r}^*(s, t)$ is known, the distance $\mathcal{D}^* \equiv \mathcal{D}[\mathbf{r}^*]$ follows.

We provided a general recipe for the solution to the EL equation using the method of lines. The resulting $N + 1$ EL equations for the discretized chain are ODEs that can be interpreted geometrically and solved for minimal solutions. Solutions consist generally of rotations and translations pieced together so the direction of velocity of any link end point does not suddenly change (the Weierstrass–Erdmann corner conditions).

We explored the minimal transformations for the simplest polymers, consisting of one or two links, in depth. For transformations between two links, convexity becomes an issue (the analogue to the direction of the radius of curvature for a continuous string). For example, even if the initial and final states lie in the same plane, if the convexities of these states are of opposite sign the transformation must pass through intermediate states that are out of the plane. Similarly, given a semicircular piece of string lying on a table, to move it to a semicircle of opposite convexity using the minimal amount of motion, the string must be lifted off the table.

The study of minimal transformations between small numbers of links has applications to the inverse kinematic problem in robotics and movement control. In the inverse kinematic problem, one is given the initial and final positions of the end-effector (the hand of the robot), and asked for the functional form of the joint variables for all intermediate states. Generally there is no unique solution until some optimization functional is introduced, such as minimizing the time rate of change of acceleration (the jerk), torque, or muscle tension (see the review [11] and references therein). The minimal distance transformation would be relevant if one sought the fastest transformation between initial and final states, without explicit regard to mechanical limitations. The intermediate points can be handled variationally as a free boundary value problem.

In the limit of a large number of links, some simplifications emerge. For chains without curvature or non-crossing constraints, the distance converges to L times the mean root square distance (MRSD) of the initial and final conformations. So for example the distance between two

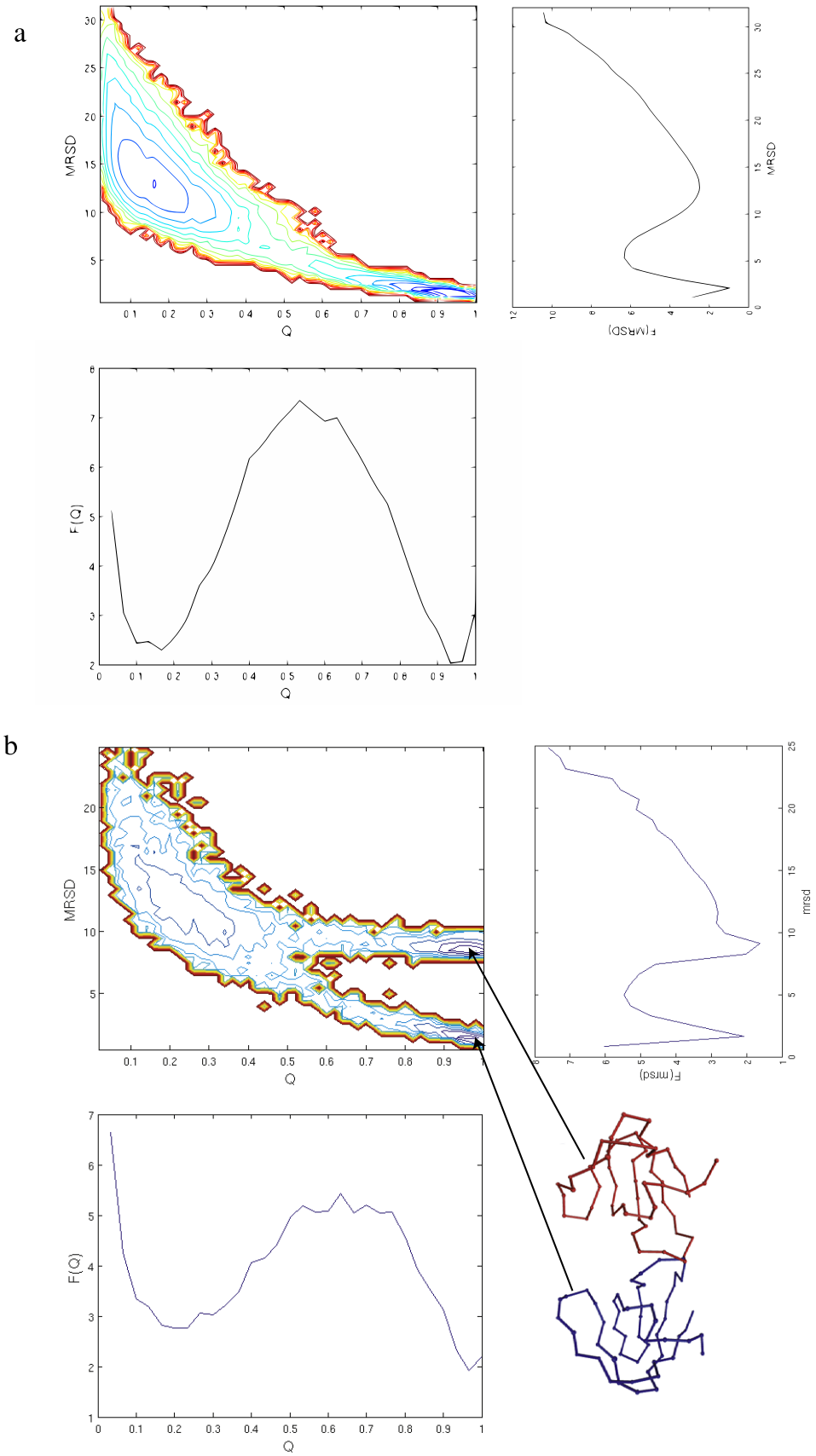


Figure 14. Free energy surfaces for the folding of Gō-model *src-SH3* using two molecular dynamics simulation packages, AMBER (a) and LAMMPS (b). The contour plots give $F(Q, \text{MRSD})$. The projections $F(Q)$ and $F(\text{MRSD})$ are also shown on each side. The COMPASS class 2 dihedral potential in LAMMPS allows for a mirror image of the folded structure (red colour structure in inset) that is not immediately evident from the $F(Q)$ or $F(\text{MRSD})$ surfaces. Future implementations of LAMMPS using COMPASS dihedrals for biomolecular simulations must then correct for dihedral angles defined on the interval $[-\pi, \pi]$.

strings of length L forming the top and bottom halves of a circle respectively is $4L^2/\pi^2$, the distance between horizontal and vertical straight lines of length L which touch at one end is $L^2/\sqrt{2}$, and the distance to fold a straight line upon itself (to form a hairpin) is $L^2/4$.

The fact that for large N the distance (over L) converges to MRSD rather than RMSD suggests that RMSD may not be the best metric for determining similarity between molecular structures, although it is ubiquitously used. Adopting MRSD may lead to improvements in structural alignment algorithms.

The MRSD was investigated as an approximate metric for protein folding. Free energy surfaces for folding were constructed for two simulation packages, AMBER and LAMMPS. It was found that including MRSD as an order parameter uncovered discrepancies between the two molecular dynamics algorithms. Because dihedral angles in LAMMPS (at least in COMPASS class 2 style) are only defined on $[0, \pi]$, the potential admits a mirror-image structure degenerate in energy with the native structure. This is easily remedied and should not be interpreted as a deficiency in the LAMMPS simulation package so long as one is aware of it. It should be mentioned that the mirror-image structure would also have been seen had RMSD been used as an additional order parameter.

It will be important for future studies to address the effects of persistence length and non-crossing on the distance between biopolymer conformations [1]. Also important is the role of entropy of paths or transformations in describing the accessibility of a particular biomolecular structure. Along these lines it will be interesting to investigate whether the distance can be a predictor of folding kinetics, or proximity to the native structure.

It is also an interesting question to ask whether the actual dynamics between polymer configurations resembles the minimal transformation, after a suitable averaging over trajectories. This question is linked with the role of the entropy of transformations described above. It is also related to the problem of finding the dominant pathway for a chemical reaction [12], which has recently been applied to the problem of protein folding [13]. We have focused here on the question of geometrical distance for complex systems, which can be separated from the calculation of quantities such as reaction paths that depend intrinsically on energetics, i.e. on the specific Hamiltonian of the system. Quantifying the relationship between geometrical distance and the dominant reaction path is an interesting future question worthy of investigation.

The notion of distance and corresponding optimal transformation for a system with many degrees of freedom is fundamental to a diverse array of research subjects. Hence we saw potential applications for this metric in areas ranging from drug design to robotics. It is not clear at present how useful the calculation of the true Euclidean distance between high-dimensional objects will be for practical applications, but we are optimistic.

Acknowledgments

We are thankful to Shirin Hadizadeh, Mike Prentiss, and Peter Wolynes for helpful discussions. SSP gratefully acknowledges support from the Natural Sciences and Engineering Research Council and the AP Sloan Foundation.

Appendix A. Necessary conditions for straight line transformations

It was shown in section 3.1 that to have straight line transformations between links it is sufficient to have facing obtuse angles on opposite sides of the quadrilateral defined by the transformation as shown in figure 3(a). We now show that it is a necessary condition as well, i.e. we show that a slide in the correct direction is not possible in the absence of obtuse angles.

Without loss of generality assume that the link is initially along the z axis. The paths travelled by the link ends are shown in the figure. Note that the end point trajectories of A and B are in 3D space so the paths travelled by A and B need not cross or lie in the same plane. Let the unit vector along A's path be $\hat{\mathbf{v}}_\mathrm{A}$ and the unit vector along B's path be $\hat{\mathbf{v}}_\mathrm{B}$. Because the angles that the path of A and the path of B make with the link are acute, the z-component of $\hat{\mathbf{v}}_\mathrm{B}$ ($\equiv z_\mathrm{B}$) is negative and the z-component of $\hat{\mathbf{v}}_\mathrm{A}$ (z_A) is positive. One can write $\hat{\mathbf{v}}_\mathrm{A}$ and $\hat{\mathbf{v}}_\mathrm{B}$ as

$$\hat{\mathbf{v}}_\mathrm{A} = \boldsymbol{\rho}_\mathrm{A} + z_\mathrm{A}\hat{\mathbf{z}}$$

$$\hat{\mathbf{v}}_\mathrm{B} = \boldsymbol{\rho}_\mathrm{B} + z_\mathrm{B}\hat{\mathbf{z}}$$

where $\boldsymbol{\rho}_\mathrm{A}$ and $\boldsymbol{\rho}_\mathrm{B}$ are vectors in the xy plane and $z_\mathrm{A} > 0$ and $z_\mathrm{B} < 0$.

Let $\mathbf{r}_\mathrm{A}(t)$ and $\mathbf{r}_\mathrm{B}(t)$ denote the positions of the A and B ends at time t:

$$\mathbf{r}_\mathrm{A} = t\hat{\mathbf{v}}_\mathrm{A}$$

$$\mathbf{r}_\mathrm{B} = g(t)\hat{\mathbf{v}}_\mathrm{B} + \hat{\mathbf{z}}$$

The rigid link constraint dictates that

$$(\mathbf{r}_\mathrm{A} - \mathbf{r}_\mathrm{B}) \cdot (\mathbf{r}_\mathrm{A} - \mathbf{r}_\mathrm{B}) = 1$$

which translates to

$$g^2 + 2g\,(z_\mathrm{B} - t\,(c + z_\mathrm{A}\,z_\mathrm{B})) - 2tz_\mathrm{A} + t^2 + 1 = 1$$

with $c = \boldsymbol{\rho}_\mathrm{A} \cdot \boldsymbol{\rho}_\mathrm{B}$. Solving for g as a function of t, keeping in mind that $g(0) = 0$,

$$g(t) = -(z_\mathrm{B} - t\ (c + z_\mathrm{A}\,z_\mathrm{B})) + \sqrt{(z_\mathrm{B} - t\ (c + z_\mathrm{A}\,z_\mathrm{B}))^2 - t^2 + 2tz_\mathrm{A}}.$$

Now if $g'(t) > 0$ it means that the B-end of the link is travelling in the assumed direction, and if $g'(t) < 0$ it means that B-end is travelling in the opposite direction (which means that the angle is not acute anymore). Writing $g'(0)$ we get

$$g'(0) = \frac{2\,z_\mathrm{B}\,c + 2\,z_\mathrm{A}\,z_\mathrm{B}^2 - 2\,z_\mathrm{A}}{2\,|z_\mathrm{B}|} + c + z_\mathrm{A}z_\mathrm{B} = \frac{-z_\mathrm{A}}{|z_\mathrm{B}|} < 0.$$

Thus point B can only travel in the opposite direction from what was assumed, which in turn means an all-acute slide

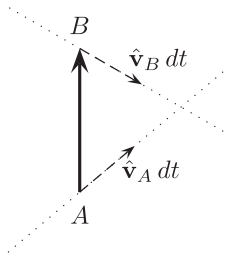


Figure A.1. Initial positions and unit velocities for the end points of the link.

is not possible. We conclude that the condition of 'facing obtuse angles' is necessary and sufficient for transformations consisting only of pure translations.

Appendix B. Critical angles

The concept of critical angle was first introduced in section 3.2. In order for a straight line slide of both ends to be possible, at some stage during the transformation the link needs to rotate about one of the ends, with the other end being stationary. In principle the rotation can be about either of the two ends and it can happen at the beginning or the end of the transformation. The conditions on the critical angle or orientation can be readily derived from the broken extremal conditions. It was seen from (18*a*) and (19), the nontrivial corner conditions read

$$\hat{\mathbf{v}}_i|_+ = \hat{\mathbf{v}}_i|_-. \tag{B.1}$$

We know that the path travelled by the moving bead during the rotation is circular and the path that is travelled during the slide part is a straight line. The broken extremal condition forces these two paths to be patched smoothly, which means that the straight line path should be tangent to the circle. In the 3D case, for the broken extremal condition to be satisfied, the straight line slide path and the circular rotation path should lie in the same plane. For example, in figure 5 where B is rotating about A initially to B_1 and then slides to B′, the rotation has to be in the plane formed by the three points ABB′.

Matching the directions of velocity as in (B.1) does not itself mean that a link can subsequently slide in a straight line; however, at the tangent point, the tangent line to the circle is perpendicular to the radius, hence one satisfies this second condition as well. Below we derive an analytical expression for the critical angle for a particular case of a single link problem, as an example and illustration of the discussed concepts. Furthermore, the particular example will be used later in appendix C to introduce minimal transformations in two dimensions.

Consider the single link action with the particular parametrization $s = s(\theta)$, as discussed in section 3.2:

$$\int (\sqrt{\dot{s}^2 + 1 + 2\dot{s}\cos\theta} + \sqrt{\dot{s}^2})\,\mathrm{d}\theta \tag{B.2}$$

where $s \equiv \overrightarrow{A(\theta)\hat{A}}$ is the (signed) distance of the A-end from its initial position, and θ is the angle between the link and the horizontal line (see figure B.1).

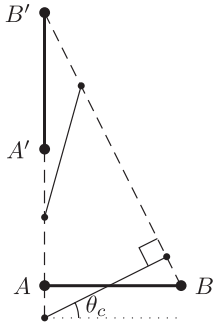


Figure B.1. Transformation in which both ends stay on a linear track.

The Euler–Lagrange equation of motion reads

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{\dot{s}}{\sqrt{\dot{s}^2}} + \frac{\dot{s} + \cos\theta}{\sqrt{\dot{s}^2 + 1 + 2\dot{s}\cos\theta}}\right) = 0 \tag{B.3}$$

We consider a transformation which is not (necessarily) a minimum:

$$s = a\cos\theta - \sin\theta + b \tag{B.4}$$

with a and b parameters to be determined.

Such a transformation in fact forces the two ends to travel on a straight line (right from the beginning), but the A side may in fact retreat and then move forward. We call such a transformation a 'hyperextended transformation'. A sample transformation of this kind is shown in figure B.1. The parameters a and b in (B.4) can be tuned to meet the boundary conditions (see below).

In fact, it is seen that point A on the link retreats backwards until it reaches some critical angle, which is when link $\overline{\mathrm{AB}}$ makes an angle $\frac{\pi}{2}$ with the straight line $\overline{\mathrm{BB}'}$ that point B travels on. Subsequently A then moves forward towards A′.

Assume that θ runs from θ_1 to θ_2, where $0 < \theta_2 < \pi/2$. For simplicity assume that both these angles are between 0 and $\frac{\pi}{2}$.

The boundary conditions dictate that

$$s(\theta_1) = 0 \tag{B.5}$$

$$s(\theta_2) = l \tag{B.6}$$

where l is the distance between A and A′.

a and b can be explicitly solved to give

$$a = \frac{-\sin\theta_2 + \sin\theta_1 - l}{\cos\theta_1 - \cos\theta_2} \tag{B.7}$$

$$b = -\frac{\cos\theta_1\ (-\sin\theta_2 - l) + \sin\theta_1\ \cos\theta_2}{\cos\theta_1 - \cos\theta_2}. \tag{B.8}$$

For our purposes we only need to note that the critical angle occurs when $\dot{s} \equiv \frac{\mathrm{d}s}{\mathrm{d}\theta}$ becomes zero, that is when A stops going backward and starts moving forward:

$$\dot{s} = -a\sin\theta - \cos\theta = 0 \tag{B.9}$$

where a is given in equation (B.7).

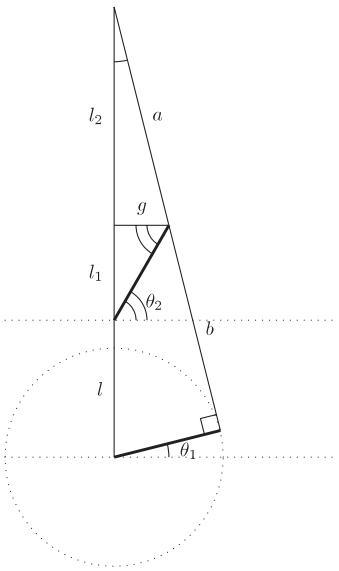


Figure B.2. Geometric proof for critical angle condition.

We can now ask what θ_1 should be so that there is no need for the link to go backward, i.e. it moves forward from the beginning and the transformation is monotonic. Equations (B.9) and (B.7) give

$$\cos\theta + \frac{-\sin\theta_2 + \sin\theta - l}{\cos\theta - \cos\theta_2}\sin\theta = 0. \tag{B.10}$$

For pedagogical reasons we prove condition (B.10) using analytic geometry as well. Looking at figure B.2 we have the following:

$$g^2 + l_1^2 = 1 \tag{B.11}$$

$$g^2 + l_2^2 = a^2 \tag{B.12}$$

$$\frac{g}{a} = \frac{1}{l + l_1 + l_2}. \tag{B.13}$$

We can solve $g = \sqrt{1 - l_1^2}$ and $a = \sqrt{1 - l_1^2 + l_2^2}$ from the first two equations and substitute in the third equation to give

$$l = \frac{\sqrt{1 - l_1^2 + l_2^2}}{\sqrt{1 - l_1^2}} - l_1 - l_2. \tag{B.14}$$

On the other hand based on our results for g and a we have

$$\sin\theta_1 = \frac{\sqrt{1 - l_1^2}}{\sqrt{1 - l_1^2 + l_2^2}} \tag{B.15}$$

$$\cos\theta_1 = \frac{l_2}{\sqrt{1 - l_1^2 + l_2^2}} \tag{B.16}$$

$$\sin\theta_2 = l_1 \tag{B.17}$$

$$\cos\theta_2 = \sqrt{1 - l_1^2}. \tag{B.18}$$

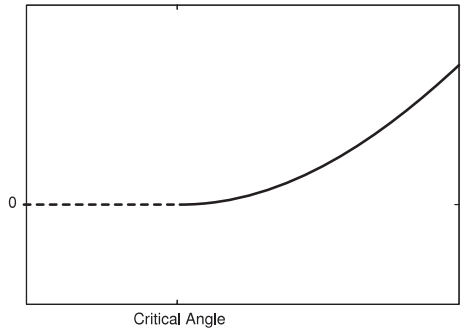


Figure B.3. A minimal transformation in $s(\theta)$ parametrization. The horizontal segment corresponds to pure rotation and the curved section corresponds to slide on straight paths. Here the corner conditions demand that the derivative $\dot{s}$ be continuous at the critical angle.

Substitution of equations (B.15)–(B.18) in equation (B.10) gives equation (B.14) after some simplification.

For the particular case that we have discussed, the proposed transformation is in fact a minimal solution if θ_1 is greater than the critical angle, because in this case a simple slide would be possible. If θ_1 is less than the critical angle a locally minimum solution as we know is pure rotation to the critical angle and then straight line slide. Pure rotation has a nice geometric interpretation in our parametrization. It corresponds to the null solution $s = 0$. Since at the critical angle $\dot{s} = 0$ we see that $s = 0$ will be smoothly patched with $s = a\cos\theta - \sin\theta + b$, as mandated by the corner conditions in equation (18*a*). A plot of s as a function of q for such a transformation is given in figure B.3.

Appendix C. Minimal transformations in two dimensions

It was seen in section 4.1 that for the case of two links, when one is confined to moving in a plane, satisfying the constant link length constraints and corner conditions does not seem to lead to solutions which are extremal. However, given the additional constraint that the links must lie in a plane, there must be one or a set of minimal transformations. We need to look at other forms of transformations, namely compound straight line transformations. We will elaborate on the idea starting with single links.

The hyper-extended solution that was discussed previously in appendix B can be considered as a very special example of compound straight line transformation. These are transformations that are made strictly from straight line paths with no pure rotation. A more general transformation is shown in figure C.1 beside the old transformation.

Note that the corners do not technically violate the corner conditions because the speed of the 'A' bead is zero at the corner point in any parametrization that can simultaneously describe A motion and B motion: since at the corner point the link makes an angle of 90° with the path that B travels, the speed of B at the critical angle is infinitely larger than the speed of A. In fact, one sees that we have an instantaneous pure rotation about the A bead, when it is at the corner point. $\hat{v}_a$ is

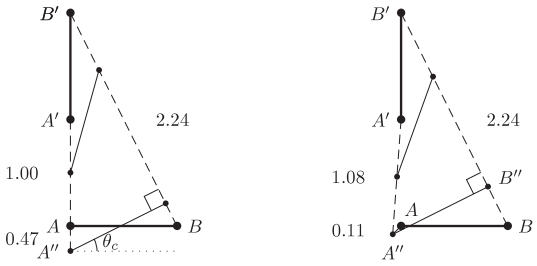


Figure C.1. The previous hyper-extended solution is shown along with a more general compound straight line transformation, where $\overrightarrow{AA''}$ travels in some general direction. The length of each line segment is written beside it. For the hyper-extended solution the value of AA″ is multiplied by two because the path is travelled twice.

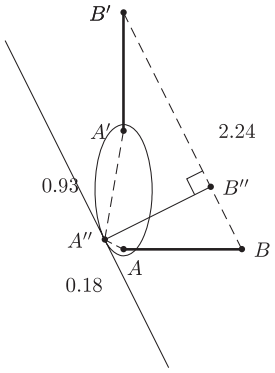


Figure C.2. Optimal compound straight line transformation.

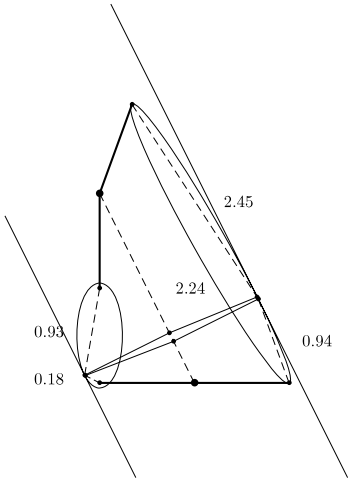


Figure C.3. An optimal compound straight line solution for two links. For this particular class of solutions, the problem is divided into two disjoint problems (one for each link) and solved separately.

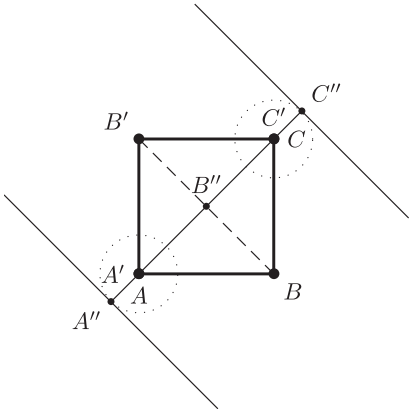


Figure C.4. Minimal transformation restricted to two dimensions, for two links of opposite convexity which form opposite sides of a square.

not clearly defined at the corners, and everywhere else (when the speed of the bead(s) is not zero) the two beads are travelling on a straight line. The two solutions depicted in the figure come from two different parametrizations of the most general form of the action and result in different distances. But each of them is a local minimum once the direction of $\overrightarrow{AA''}$ is picked, and these local minima have different values for the distance.

We can then ask about the best position to put the corner point, to minimize the distance travelled in the compound straight line transformation, with respect to other compound straight line transformations. We assume the corner occurs on one side and we take it to be the 'A' side.

Note that at the corner the link makes a 90° angle with the B-bead path $\overline{BB'}$, meaning that the distance from the corner point to the B path is always the length of the link, i.e. unity here. Also note that the total distance that the 'A' bead travels is the distance from the initial point A to the corner point A″, plus the distance from A″ to the final position A′.

The locus of points with equal sum of distances from two points A and A′ defines an ellipse with foci at A and A′. Moreover, the length of the major axis of the ellipse equals the sum of the distances from the foci. Thus the smaller the major axis of the ellipse with foci A and A′, the smaller the total distance travelled by the 'A' bead. Moreover A″ should sit on a line parallel to the B path at a distance of 1 from the B-path line $\overline{BB'}$. So in seeking the shortest distance travelled by the A end of the link, we seek the point A″ such that it lies on an ellipse with foci A and A′, the ellipse shares at least one point with a line parallel to $\overline{BB'}$ and distance 1 away from it, and lastly that the ellipse has the smallest possible major axis (see figure C.2). So the ellipse giving the minimal distance is tangent to the parallel line, and A″ is the tangent point. This is illustrated in figure C.2.

This solution can be straightforwardly extended to two links, as depicted in figure C.3. Consider then the example in figure 13(a), where the links are no longer allowed to move out of the plane (see figure C.4). Here $\mathbf{r}_A = \mathbf{r}_{A'}$ and $\mathbf{r}_C = \mathbf{r}_{C'}$ and the above ellipses turn into circles centred at A and C. The circles have radii $1 - 1/\sqrt{2}$, so that the perpendicular distance from line $\overline{BB'}$ to the farthest point on the circle is 1 and a fully extended intermediate state is allowed.

References

[1] Plotkin S S 2007 Generalization of distance to higher dimensional objects *Proc. Natl Acad. Sci. USA* **104** 14899–904

[2] Gelfand I M and Fomin S V 2000 *Calculus of Variations* (New York: Dover)
[3] Greene J, Kahn S, Savaj H, Sprague P and Teig S 1994 Chemical function queries for 3D database search *J. Chem. Inf. Comput. Sci.* **34** 1297–308
[4] Lemmen C and Lengauer T 2000 Computational methods for the structural alignment of molecules *J. Comput. Aided Mol. Des.* **14** 215–31
[5] Patel Y, Gillet V J, Bravi G and Leach A R 2002 A comparison of the pharmacophore identification programs: catalyst, DISCO and GASP *J. Comput. Aided Mol. Des.* **16** 653–81
[6] Gerstein M and Levitt M 1998 Comprehensive assessment of automatic structural alignment against a manual standard *Protein Sci.* **7** 445–56
[7] Baker D and Sali A 2001 Protein structure prediction and structural genomics *Science* **294** 93–6
[8] Ueda Y, Taketomi H and Gō N 1975 Studies on protein folding, unfolding and fluctuations by computer simulation. I. The effects of specific amino acid sequence represented by specific inter-unit interactions *Int. J. Peptide Res.* **7** 445–59
[9] Shea J and Brooks C III 2001 From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding *Annu. Rev. Phys. Chem.* **52** 499–535
[10] Clementi C and Plotkin S S 2004 The effects of nonnative interactions on protein folding rates: theory and simulation *Protein Sci.* **13** 1750–66
[11] Kawato M 1996 *Advances in Motor Learning and Control* ed H N Zelaznik (Champaign, IL: Human Kinetics) pp 225–59
[12] Onsager L and Machlup S 1953 Fluctuations and irreversible processes *Phys. Rev.* **91** 1505–12
[13] Sega M, Faccioli P, Pederiva F, Garberoglio G and Orland H 2007 Quantitative protein dynamics from dominant folding pathways *Phys. Rev. Lett.* **99** 118102