Parameter continuation, semi-classical path integrals, and looping probability of DNA

John Maddocks, EPFL

BIRS November 2012

Introduction Part 1: semi-classical looping probabilities for elastic polymers

Explain how the computation of the looping probability for an elastic polymer, such as DNA, can be cast in terms of a ratio of path integrals to give a probability density function (or pdf) on the special Euclidean group $SE(3)$.

This quotient can then be approximated with a leading order term that involves an isoperimetric variational principle from the $1D$ calculus of variations, along with the first, or semi-classical, correction which involves the determinant of certain solutions to the associated Jacobi equations.

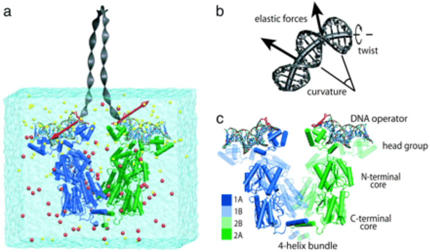
This part is joint work with L. Cotta-Ramusino and R.S. Manning, both published C-R & jhm, Phys. Rev E, 2010, and ongoing. Supported by Swiss National Science Foundation.

Introduction Part 2: Jacobi systems and exchange of stability in parameter-dependent variational principles

As time permits I will then describe connexions between Part 1, and results concerning exchange of stability and distinguished bifurcation diagrams in parameter dependent variational principles, including isoperimetric variational principles where the parameters are Lagrange multipliers.

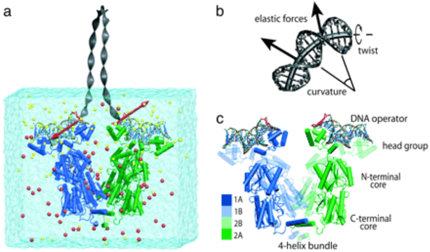
DNA loop formation is important in Biology

One biologically important class of examples involve DNA looping. For example the LAC-repressor loop of length 75 bp (image from Villa, Balaeff and Schulten, PNAS 2005)



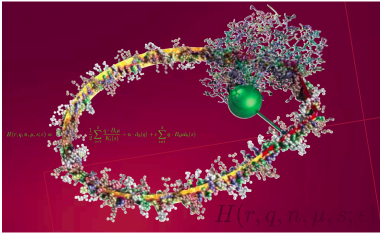
DNA loop formation is important in Biology

One biologically important class of examples involve DNA looping. For example the LAC-repressor loop of length 75 bp (image from Villa, Balaeff and Schulten, PNAS 2005)



The probability of the first end of the loop binding depends only on concentration, but then for the second end to bind the DNA must deform, and that free energy will have a large effect.

The probability of minicircle formation is a particular and experimentally important case of looping



DNA minicircle cyclization rates or J-factors, i.e. probability of the formation of closed loops of DNA with or without a bound protein, are exquisitely dependent upon sequence, particularly intrinsic curvatures, at scales of 50 – 250 bp. Such in vitro cyclization experiments measuring J-factors à la Shore-Baldwin, Crothers group, Kahn, Widom, Vologodskii, Orozco, etc, are a longstanding and still currently active experimental area.

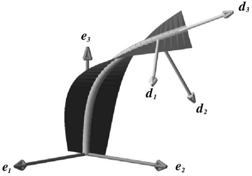
The J-factor as an experimental probe of the physical properties of DNA

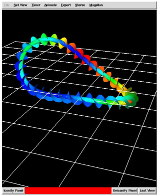
"DNA cyclization is potentially the most powerful approach for systematic quantitation of sequence-dependent DNA bending and flexibility."
First sentence of abstract in Zhang & Crothers, Biophysical J. **84** (2003)

More recently the importance of methylation patterns and epigenetics has become apparent. Does methylation affect the physical properties of DNA in a significant way?

Need more experimental data, and better models to interpret the experimental data. Note however that there is no single "correct" level of coarse-graining DNA.

A Comparatively Simple Coarse Grain Model of DNA





Today talk about the analysis of an extensible, shearable (or Cosserat) rod model, and its inextensible, unshearable (Kirchhoff) limit. In these models the configuration is determined by the strains $\mathbf{w}(s) \in se(3)$. The geometry of the reconstruction is nonlinear.

Depending on how $SE(3)$ is treated the strains $\mathbf{w}$ will involve both the parametrization $q(s)$ of $SE(3)$ and its space derivative $q'(s)$.

The Assumed Functional Form of Energy

The (free or potential) energy (or Hamiltonian in stat mech language) is assumed to be a shifted quadratic function of the strain variables

$$U\left(\mathbf{w} \approx (q, q')\right) = \frac{1}{2} \int_0^L (\mathbf{w} - \hat{\mathbf{w}})^T \mathbf{P} (\mathbf{w} - \hat{\mathbf{w}}) \; ds$$

The input parameters are the coefficients $\mathbf{P}(s) > 0$, a 6×6 symmetric matrix of stiffnesses, and the $\hat{\mathbf{w}}(s)$ which determine the intrinsic shape of an absolute minimum energy configuration.

Estimation of the Model Input Parameters

Both $\mathbf{P}(s) > 0$ and $\hat{\mathbf{w}}(s)$ encode the sequence-dependence of a particular DNA fragment or oligomer. They are both possibly rapidly varying functions of s, which opens the way to effective constant coefficient models determined by averaging, e.g. standard worm like chain models of polymer physics, which are typically constant coefficient and inextensible and unshearable.

One way to estimate these input coefficients as a function of the DNA sequence is from a training set of finer-grain molecular dynamics simulations, but that is another long story: for example jhm et al Phys. Chem. Chem. Phys. **11** (2009), and Gonzalez, Petkevicuite, jhm, J. Chem Phys. submitted (2012)

Todays Point of Departure

We make the assumption that the DNA interacts with a solvent heat bath in such a way that the configurations of a given oligomer are governed by an equilibrium, or stationary, pdf of the form:

$$\frac{1}{Z}\exp[-\beta U]$$

for a specific choice of configuration variable q parametrizing $SE(3)$, and associated known sequence-dependent potential $U(q)$ of the functional form already introduced, partition function Z, and heat bath temperature scale β.

In the current continuum context with $U(q)$ being a functional on paths in $SE(3)$ the precise sense of Z is delicate.

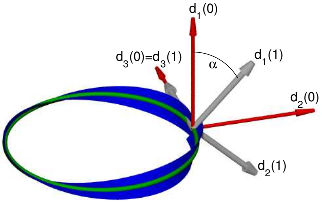
Looping Probability Density as a Marginal of the Configuration Probability Density

A special feature of polymers such as DNA is that they are one-dimensional objects in which the probability of interest is frequently not on the entire configuration $q(s)$ itself, but is instead a marginal probability on $q(L)$ or q-final or q_F at one end of the chain conditioned on the value $q(0)$ or q-initial or q_I at the other end of the chain.

For example one can have $q_F = (r_F, R_F) \in SE(3)$ an origin and orientation of a rigid base pair, or frame, and you want to know the pdf $p(q_F|q_I)$ on $SE(3)$ for the final frame being at a given place and orientation, given the origin and orientation q_I at the other end of the chain (which information typically just factors out a Euclidean symmetry of the potential).

The J-factor for cyclization

For example to model the 'looping' problem of DNA minicircle cyclization rates, you want to approximate the probability density function $p(q_F|q_I)$ in the particular case $q_F = R(\alpha)q_I$, i.e. the two end points and end tangents of the DNA fragment coincide and the terminal frames are rotated through an angle α.



(Technicality: The additional parameter α allows the ribbon to differ from the two DNA backbones.)

The Path Integral Expression for looping J-factors.

In the context of such looping problems the desired marginal probability density function (on the six dimensional group $SE(3)$) involves a 'summation' over all intervening paths satisfying both initial and final end conditions:

$$p(q_F | q_I) = \frac{1}{Z} \int_{q_I}^{q_F} \exp[-\beta U(q)] \quad D[q]$$

while the partition function is the sum over all paths satisfying only the initial boundary condition:

$$Z = \int_{q_I} \exp[-\beta U(q)] \quad D[q]$$

In our continuum context the integration measure $D[q]$ is over all paths satisfying the prescribed boundary condition(s): at both ends in numerator, at only one end in denominator.

“semi-classical” approximation of real path integrals

If:

- there is a unique energy minimizing configuration $\bar{q}$ with initial and final configurations q_I and q_F
- and β is large in an appropriate sense,

then can write $q = \bar{q} + h$ and make the ‘semi-classical’ approximation in terms of the second variation $\delta^2 U(\bar{q}; h)$ of the potential U at the minimum $\bar{q}$

$$p(q_F|q_I) \approx \exp[-\beta U(\bar{q})] \quad \frac{\int_{h_I=0}^{h_F=0} \exp[-\beta\delta^2 U(\bar{q}; h)] \quad D[h]}{Z}$$

The second term is now the ratio of two quadratic path integrals: numerator in h with two boundary conditions, denominator in strains w with no boundary condition. This ratio is a well-behaved object.

Explicit form of the second variation

$$\delta^2 U(\bar{q};\mathbf{h}) = \frac{1}{2}\int_0^L \left[\mathbf{h}'^T\mathbf{P}(s)\mathbf{h}' + \mathbf{h}^T\mathbf{Q}(s)\mathbf{h} + 2\mathbf{h}'^T\mathbf{C}(s)\mathbf{h}\right]\, ds$$

where $\mathbf{P}(s)$ is the already given stiffness matrix, while the anti-symmetric matrix $\mathbf{C}(\bar{q})$ and symmetric matrix $\mathbf{Q}(\bar{q})$ in general depend upon the minimum energy configuration being considered due to the nonlinear geometry of rods.

The anti-symmetric coefficient matrix $\mathbf{C}$ of the cross-terms does not vanish in any case of rods that I know of. In some sense $\mathbf{C}$ represents 'gyroscopic' terms.

And $\mathbf{C} \neq \mathbf{0}$ causes conniptions in all the standard path integral solution formula that I know.

Explicit form of the associated Jacobi system

The associated Jacobi equations have the linear Hamiltonian form

$$\begin{pmatrix} \mathbf{V} \\ \mathbf{W} \end{pmatrix}' = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{bmatrix} \mathbf{S} \begin{pmatrix} \mathbf{V} \\ \mathbf{W} \end{pmatrix}, \qquad \begin{bmatrix} \mathbf{V} \\ \mathbf{W} \end{bmatrix} (0) = \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}$$

where

$$\mathbf{S} = \begin{bmatrix} \mathbf{C}^T\mathbf{P}^{-1}\mathbf{C} - \mathbf{Q} & -\mathbf{C}^T\mathbf{P}^{-1} \\ -\mathbf{P}^{-1}\mathbf{C} & \mathbf{P}^{-1} \end{bmatrix}$$

and the 6×6 solution blocks $\mathbf{V}(s)$ and $\mathbf{W}(s)$ of the Hamiltonian Jacobi system for the particular initial conditions shown will be important and will reappear.

It is also true that $\mathbf{V}(s)$ and $\mathbf{W}(s)$ behave smoothly in the inextensible, unshearable limit in which certain entries of the stiffness coefficient matrix $\mathbf{P}$ become infinite, and $\mathbf{P}^{-1}$ smoothly approaches a non-invertible matrix.

Evaluation of the semi-classical approximation

'Time slice' (in s) in the ratio of quadratic path integrals take limit and arrive at

$$p(q_F|q_I) \approx \frac{\exp[-\beta U(\bar{q})]}{\sqrt{\det \frac{2\pi}{\beta}\mathbf{V}(L)}}$$

where the matrix $\mathbf{V}(s)$ is the solution block of the Hamiltonian form of the Jacobi system at the local minimizer $\bar{q}(s)$.

Computing the minimum energy shape $\bar{q}(s)$ is a nonlinear problem, but unavoidable, and efficient ways of doing that already known for quite some time, for example Manning, Maddocks, Kahn, J. Chem Phys, (1996).
Evaluating the semi-classical Jacobi field correction is then a single IVP solve for a linear system of matrix ODE, usually nonconstant coefficient.

First question for the expert(s) in the audience

The functional form of our formula, namely a pre-factor involving energy of a minimal path and a correction in terms of a determinant evaluated on a Jacobi field, is quite familiar in the large literature on various ways of evaluating quadratic path integrals, mostly couched in terms of the case of imaginary integrands that arises in the semi-classical treatment of quantum mechanics. Notable authors include Morette, and Gelfan'd-Yaglom. However for $\mathbf{C} \neq 0$ the only applicable evaluation formula we could find is due to G.J. Papadopoulos (Phys Rev D, 1975). It involves a nonlinear set of ODE related to, but different from, the linear Jacobi equations.

We then introduce a nonlinear change of variables, of Riccati type, to rewrite the Papadopoulos solution formula in terms of solutions to the linear Jacobi system. Any related citations known?

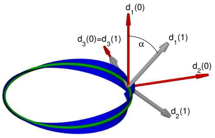
Second question for the expert(s) in the audience

The solution formula blows up when $\det \mathbf{V}(L) = 0$, i.e. at conjugate points. In the inextensible, unshearable case this conclusion requires the Hamiltonian form of the Bolza (1904) theory of the degenerate 'Jacobi' equations (actually DAE) appropriate for determining conjugate points and positive definiteness of the constrained second variation in the isoperimetric calculus of variations problems (see Manning, Rogers, and JHM, Proc Roy Soc A, 1998).

Is there a literature on semi-classical formula close to conjugate points for real quadratic path integrals?

How good is the semi-classical approximation?

Can gain some insight from a toy example where all computations explicit: $\alpha = 0$, untwisted loop closure for intrinsically straight untwisted strip, with diagonal, constant stiffness matrix $\mathbf{P} = diag(K_1, K_2, K_3, A_1, A_2, A_3)$.



Equilibrium configuration is an untwisted circle (a simple case of a helical, or relative, equilibrium for which the Jacobi equations can be made constant coefficient).

Explicit formula (don't sweat the detail)

Our semi-classical approximation to the probability density function for cyclization can be calculated analytically (here for the case $K_1 < K_2$ and $K_1 < K_3$):

$$p(L) = e^{-2\frac{\beta\pi^2 K_1}{L}} \frac{4\pi\,\beta^3}{L^6\,A} \left[\sqrt{\frac{K_1^5(K_3-K_1)}{B\,(1-\cos\lambda)}} \right]$$

$$= e^{-2\frac{\beta\pi^2 K_1}{L}} \left[2\left(\frac{\beta}{2\pi}\right)^{3/2} \sqrt{\frac{(2\pi)^4 K_1^3}{L^7 A^2}} \right]_{\mathrm{p}} \left[\left(\frac{\beta}{2\pi}\right)^{3/2} \sqrt{\frac{(2\pi)^4 K_1^2(K_3-K_1)}{L^5 B(1-\cos\lambda)}} \right]_{\mathrm{op}}$$

where

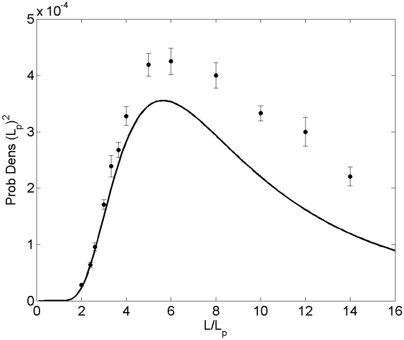
$$A = \left[1 + \left(\frac{1}{A_2} + \frac{1}{A_3}\right)\left(\frac{2\pi}{L}\right)^2 K_1 \right] \qquad B = \left[1 + \left(\frac{2\pi}{L}\right)^2 \frac{(K_3-K_1)}{A_1} \right]$$

and

$$\lambda = 2\pi\sqrt{\frac{(K_3-K_1)(K_2-K_1)}{K_2K_3}}.$$

In inextensible unshearable limit $A \to 1$, $B \to 1$.

Comparison with Monte Carlo for one particular set of stiffness parameters as a function of non-dimensional length of the loop



Experiments usually done at less than $2L_p$ where events are so rare that it is hard to get good Monte Carlo sampling.

Part II: Jacobi fields in parameter-dependent variational principles

Long realized that families of critical points, i.e. solutions of first-order necessary conditions, e.g. Euler-Lagrange equations, contain information about second-order necessary conditions at each extremal.

Describe various older results surrounding the idea of distinguished bifurcation diagrams, including the construction of a Lyapunov function which implies a stability result for the N-soliton of KdV (Maddocks and Sachs, 1993, 1995).

Then first attempt at a distinguished bifurcation diagram for semi-classical looping problems.

Classic exchange of stability result—Poincaré or earlier

If a branch of solutions to a set of equations

$$f(u, \lambda) = 0$$

with a scalar parameter λ has a (simple) fold, then one branch being 'stable' with respect to some spectral problem

$$f_u(u, \lambda)\xi = \mu\xi$$

implies that the other branch is 'unstable', e.g. B stable implies both A and C unstable

Stability and Folds 303

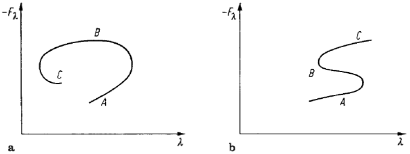


Fig. 2a and b. Two possible forms of two successive simple folds

Exchange of stability in distinguished diagram for variational problems (jhm ARMA **99**,1987)

A branch of solutions to a set of first-order necessary equations

$$F_u(u, \lambda) = 0$$

are not (local) minima if it is an upper branch in a fold opening to the left, or a lower branch in a fold opening to the right, in the distinguished bifurcation diagram $(\lambda, -F_\lambda)$. Thus B are not minima, and C are not minima in part a). No assumption on A and no assumption of simplicity.

Stability and Folds 303

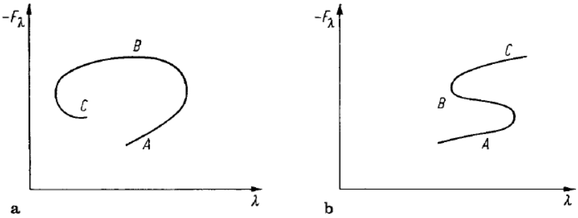


Fig. 2a and b. Two possible forms of two successive simple folds

Exchange of stability in distinguished diagram for isoperimetric variational problems

When λ enters the first-order necessary equations linearly it may or may not be a Lagrange multiplier:

$$G_u(u) - \lambda H_u(u) = 0$$

The distinguished diagram is now (λ, H), unconstrained index changes at vertical folds as before, and constrained index changes at horizontal folds.

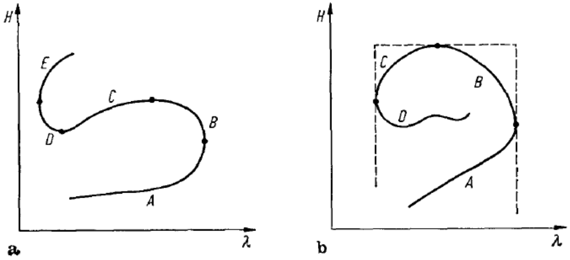
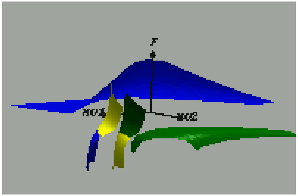
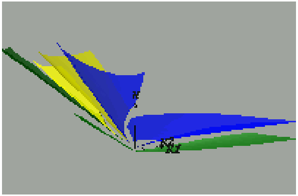


Fig. 9a and b. Possible bifurcation diagrams. The labelled segments have boundaries at the points with vertical and horizontal tangents

Exchange of stability in multiply constrained isoperimetric variational problems

To extend to multi-parameter problems, reinterpret distinguished diagram results in terms of plots of the Energy against the parameters. Then folds become cusps. In case of functionals with linear parameter dependence $F(u, \lambda) := G(u) - \lambda \cdot \mathbf{H}(u)$ have two natural and conjugate distinguished diagrams (λ, F) and $(\mathbf{H}, G)$





Bifurcation diagrams shown are for the Axes of Staude (the steady spins of an asymmetric, nonintegrable, heavy rigid body about a fixed point).

Signs of principal curvatures and connexions between constrained and unconstrained indices

For isoperimetric problems a single critical point has two indices, one as an unconstrained critical point where the λ are prescribed, and one as a constrained critical point where the $\mathbf{H}$ are prescribed. The difference between these two indices is given by the number of positive principal curvatures in (either of) the two distinguished diagrams.

For example in the Lax-Novikov-Kruskal variational characterization of the KdV N-soliton profiles as critical points of integrals of the motion, the unconstrained index can be shown to be $\mathrm{Int}[(N+1)/2]$,i.e. $1, 1, 2, 2, 3, 3, 4, 4, \ldots$, and the distinguished diagram can easily be shown to have the same number of positive curvatures. Thus the N-soliton profiles are all constrained minimizers, a Lyapunov function for the dynamics can then be constructed, and orbital stability concluded. (Benjamin, 1972, used the Bolza theory of conjugate points to prove the $N = 1$ case.)

Is there a useful distinguished bifurcation diagram for semi-classical looping problems?

As yet unclear. For the moment just remark that in the spirit of Hamilton-Jacobi theory one can write the action, which in the looping problem context is just the elastic energy U, as a function of the final state q_F. Then expand $q_F = \bar{q}_F + \delta q_F$ to obtain

$$U(\bar{q}_F + \delta q_F) = \bar{p}_F \cdot \delta q_F + \frac{1}{2} \delta q_F \cdot \mathbf{W}(L)\mathbf{V}^{-1}(L) \delta q_F + H.O.T.(P2C2E)$$

Here the matrices $\mathbf{W}(s)$ and $\mathbf{V}(s)$ are our old friends defined as solutions of the Hamiltonian version of the Jacobi equation which implies that $\mathbf{W}(L)\mathbf{V}^{-1}(L)$ is a symmetric matrix and $\mathbf{V}(s)$ is invertible up to the first conjugate point, and a distinguished diagram should encode $\det \mathbf{V}^{-1}(L)$.

Thank you for your attention.