Two-Point Boundary Value Problems for Curves: The Case of Minimum Free Energy Paths *
corrected

Robert D. Skeel
Department of Computer Science, Purdue University
West Lafayette, Indiana 47907-2107, U.S.A.

Abstract

The calculation of a minimum free energy path can be considered as a two-point boundary value problem. It is, however, nonstandard in a number of respects, which are detailed here. In particular, (i) the cost of a function evaluation depends on the location of the previous evaluation, and (ii) the value of a function will contain random noise of controllable amplitude. Hence, this is an unlikely application for an existing differential equation solver, and it underscores the importance of disseminating advances in numerical techniques by means of understandable expository articles (and pseudocode) in addition to black box solvers. *The second paragraph of Section 1 is corrected.*

Because free energy is defined in terms of energy by an averaging process, we discuss minimum energy paths and the practicalities of computing averages before discussing minimum free energy paths.

1 Minimum energy paths

Assume Newtonian dynamics with a diagonal matrix of masses M and a potential energy function $U(x)$ depending on the configuration x. The kinetic energy is $\frac{1}{2}p^{\mathsf{T}}M^{-1}p$ where p are momenta. Assume random initial conditions drawn from a Boltzmann-Gibbs distribution, for which the probability density is proportional to $\exp(-\beta(\frac{1}{2}p^{\mathsf{T}}M^{-1}p + U(x)))$, where β^{-1} is the inverse temperature. This probability density is invariant for the dynamics. If the potential energy surface $U(x)$ has a basin A in configuration space with barrier heights of several times β^{-1}, the system will be trapped in A for long periods at a time. In such a case, we call A a *metastable state*.

Of interest are the transition paths from one metastable state A to another B. A minimum energy path (MEP) $x = X(s), 0 \le s \le 1$, where s is an arbitrary parameter, is a representation of a bundle of such transition paths. Its derivation is based on a number of assumptions (omitted here). It is defined so that at each point $x = X(s)$, the energy $U(x)$ is a local minimum with respect to other points on the plane passing through $X(s)$ and normal to $D^{-1}X_s(s)$, where $D = \beta^{-1}M^{-1}$ and $X_s = (\mathrm{d}/\mathrm{d}s)X$. A necessary condition is that $X_s(s) \parallel -D\nabla U(X(s))$. Parameterization of a curve is arbitrary, so we normalize by choosing s to be *relative* arc length, meaning that $X_s^{\mathsf{T}}X_s =$ constant, $X(0) \in A$, and $X(1) \in B$. As a two-point boundary value problem, this would seem to be overdetermined. However, the direction of the curve is undetermined at critical points x^*, where $\nabla U(x^*) = 0$. This permits the construction of an

*This material is based upon work supported by NIH Grant P41-RR05969.

MEP from a given $X(0)$ to a given $X(1)$ via a sequence of critical points of $U(x)$: from $X(0)$ to the local minimum of its basin, then a chain of heteroclinic orbits from one critical point to another [1, App. A] until reaching the local minimum in the basin of $X(1)$, and finally to $X(1)$. The complete system of equations is

$$X_s \parallel -D\nabla U(X), \quad X_s^\mathsf{T} X_{ss} = 0, \quad \nabla U(X(0) = \nabla U(X(1)) = 0.$$

Since the singularities present in the MEP are not the main theme of this article, we also present an alternative to the MEP, which Zhao et al. [2] claim as a better representation of a bundle of transition paths, and which they call a maximum flux transition path (MFTP). It satisfies the condition

$$X_s \parallel -D\nabla U(X) + \frac{1}{\beta}\frac{1}{X_s^\mathsf{T} D^{-1} X_s} X_{ss},$$

which includes a finite-temperature correction to the MEP that reduces curvature. This condition combines with the arc length normalization to yield a standard system of 2nd order ordinary differential equation:

$$X_{ss} = (I - \Pi)\beta(X_s^\mathsf{T} D^{-1} X_s) D\nabla U(X) \quad \text{where } \Pi = X_s X_s^\mathsf{T}/(X_s^\mathsf{T} X_s).$$

In the limit of infinite temperature $\beta^{-1} \to \infty$, the path becomes a straight line, which is the desired behavior.

Standard numerical techniques apply to the MFTP, though provision must be made in cases of low temperature β^{-1}, for which the advection term (in Π) dominates, e.g., upwinded differencing.

2 Practicalities of calculating averages

Given some "observable" $A(x)$, the problem is to estimate the expectation

$$\langle A \rangle = \langle A(\mathbf{x}) \rangle = \int A(x)\rho(x)\mathrm{d}x$$

with respect to some probability density $\rho(x)$, which is known—except for its normalizing constant. This is estimated by an average

$$\langle A \rangle \approx \bar{A} = \frac{1}{N}\sum_{n=0}^{N-1} A(\mathbf{x}_n)$$

where the $\mathbf{x}_n$ are random configurations with probability density $\rho(x)$.

In practice, the generation of random configurations is done using a Markov chain Monte Carlo method (MCMC):

$$x_0 \to \mathbf{x}_1 \to \mathbf{x}_2 \to \cdots \to \mathbf{x}_{N-1}$$

where only in the limit $n \to \infty$ does the density of $\mathbf{x}_n$ converge to $\rho(x)$.

The initial value x_0 is arbitrary but, in practice, not random. A significant shortcoming of an MCMC method is that the estimate *and its accuracy* depend on the initial value, so we write $\bar{A} = \bar{A}(x_0)$. Therefore, we first run the MCMC method until it produces a value x_m which is as good as random, a process called *equilibration*, and use instead $\bar{A}(x_m) = \frac{1}{N}\sum_{n=0}^{N-1} A(\mathbf{x}_{m+n})$. To be precise, a value x_m can be said to be *equilibrated* for a given MCMC process with respect to an observable $A(x)$ if $\langle(\bar{A}(x_m) - \langle A\rangle)^2\rangle \le \langle(\bar{A}(\mathbf{x}) - \langle A\rangle)^2\rangle$ for sufficiently large N, though this cannot be verified in practice. To reduce the cost of equilibration m, a plausible value is chosen for x_0.

The number of samples N is to be chosen so that $|\bar{A} - \langle A \rangle| \leq$ *tolerance* with (say) 95% confidence, though this is not possible in practice (without prior assumptions).

A good averaging routine might automatically choose values for m and N but would still require from the user a plausible x_0 and a *tolerance* for the statistical error.

3 Minimum free energy paths

Quite likely the transition paths will *not* cluster into bundles in configuration space, but they may do so if we describe these paths in a reduced set of variables called *collective variables*:

$$\zeta = \xi(x).$$

With assumptions similar to those for the MEP and MFTP, one can derive conditions for a representative transition path [1, 2].

Instead of the potential energy $U(x)$, the relevant quantity is the free energy $F(\zeta)$. The potential energy satisfies

$$U(x) = -\beta^{-1} \log \rho(x) + \text{constant}$$

where $\rho(x)$ is the Boltzmann-Gibbs probability density for configuration space. Analogously the free energy is defined by

$$F(\zeta) = -\beta^{-1} \log \rho_\xi(\zeta) + \text{arbitrary constant}$$

where $\rho_\xi(\zeta)$ is the probability density for $\zeta = \xi(\mathbf{x})$ It can be shown that the density has the formula

$$\rho_\xi(\zeta) = \int \delta(\xi(x) - \zeta)\rho(x)\mathrm{d}x,$$

where $\delta(\zeta)$ is the product of Dirac delta functions of the components of ζ. The integral defining ρ_ξ is, in fact, an expectation $\langle \delta(\xi(x) - \zeta) \rangle$. The most convenient way to calculate this is to use a blurred delta function

$$\delta(s) \approx (2\pi)^{-1/2} \varepsilon^{-1} \exp(\frac{1}{2}(s/\varepsilon)^2)$$

as an approximation.

The appropriate analog of D is the following conditional expectation:

$$D(\zeta) = \beta^{-1} \langle \xi_x M^{-1} \xi_x^\mathsf{T} \rangle_{\xi(x)=\zeta}.$$

With these definitions, the maximum flux transition path $\zeta = Z(s)$, $0 \leq s \leq 1$, satisfies

$$Z_{ss} = (I - \Pi) \left(\beta (Z_s^\mathsf{T} D^{-1} Z_s) D \nabla F(Z) + D_s D^{-1} Z_s \right) \quad \text{where } D = D(Z) \text{ and } \Pi = Z_s Z_s^\mathsf{T} / (Z_s^\mathsf{T} Z_s),$$

with boundary values given for $Z(0)$ and $Z(1)$. The *minimum free energy path* is the limiting zero temperature case $\beta^{-1} \to 0$ with boundary conditions $\nabla F(Z(0)) = \nabla F(Z(1)) = 0$.

Discretization reduces the problem to a system of nonlinear equations for nodes Z_j, $j = 0, 1, \ldots, J$, along the curve. To solve these equations, it is practical to compute D, ∇F, and the Hessian of F, at approximations to the nodes, as averages from a single (yet very long) MCMC simulation.

The solution of the discrete equations is obtained from an iteration, given some initial guess $\{Z_j^{(0)}\}$, which could be a straight line for a simple problem. Due to the need to do sampling, there should also be given values $\{x_j^{(0)}\}$ in configuration space such that $\xi(x_j^{(0)}) \approx Z_j^{(0)}$. An iterative solver generates values $\{Z_j^{(k+1)}\}$ from $\{Z_j^{(k)}\}$. To begin the sampling for an iterate

$Z_j^{(k+1)}$, use for a value $x_j^{(k+1)}$ one that is generated during the production phase of the sampling for $Z_j^{(k)}$, e.g., the last value. The length m of the equilibration phase will depend on how different $Z_j^{(k+1)}$ is from $Z_j^{(k)}$. Greater increments $Z_j^{(k+1)} - Z_j^{(k)}$ will result in longer equilibration times m. Therefore, *advanced iterative solvers* that take large steps *may not be helpful*. Currently, methods modeled after gradient descent are used.

References

[1] L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti. String method in collective variables: Minimum free energy paths and isocommittor surfaces. *J. Chem. Phys.*, 125:024106 (15 pages), 2006.

[2] R. Zhao, J. Shen, and R. D. Skeel. Maximum flux transition paths of conformational change, 2009. Submitted.