FI MU

Faculty of Informatics
Masaryk University

Smoothing Integral Transforms for Molecular Mechanics Force Fields

by

Petr Mejzlík

FI MU Report Series FIMU-RS-95-02

 April 1995

Smoothing Integral Transforms for Molecular Mechanics Force Fields

Petr Mejzlík[†]
mejzlik@muni.cz

April 30, 1995

Abstract

A new class of methods, based on a special type of smoothing integral transforms, has recently been developed to solve problems concerning conformational optimization in computational chemistry. These methods do not apply an optimization procedure directly to the original potential function, but trace low minima through a sequence of transformed potential functions with decreasing level of smoothing. This work studies the integral smoothing transforms in general and applies the theory to a class of potential functions which are typically used by molecular mechanics and related methods for computations with large molecules. It also addresses the problems of computational complexity of the transforms and their approximations.

AMS subject classification: 44A20, 49M37, 92E10

[†]Department of Information Technology, Faculty of Informatics, Masaryk University, Botanická 68a, 602 00 Brno, Czech Republic.

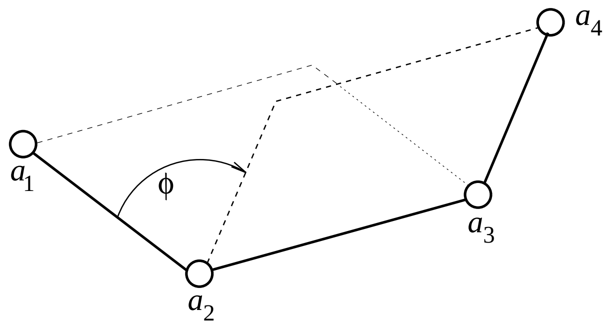
1 Introduction

Molecular mechanics and related methods are the main tools for computations with large molecules. During the past 15 years, hundreds of studies have been done using these methods to elucidate behavior of biologically important molecules, mostly proteins, DNA and RNA.

Molecular mechanics is based on repeated calculation of a potential function, called *force field*, which depends on coordinates of the atoms of the molecule. A typical force field can be expressed by the following equation:

$$\begin{aligned} H = & \sum_{\text{bonds}} S_k (b_k - b_{k,0})^2 + \sum_{\substack{\text{bond}\\ \text{angles}}} B_\ell (\theta_\ell - \theta_{\ell,0})^2 \\ & + \sum_{\text{torsions}} T_m [1 + \cos(n_m \phi_m - \delta_m)] + \sum_{\substack{\text{nonbonded}\\ ij \text{ pairs}}} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{D r_{ij}} \right]. \end{aligned} \tag{1}$$

Here b_k is the length of the k-th bond, θ_ℓ is the angle of the ℓ-th pair of consecutive bonds, ϕ_m is the value of the m-th torsion, and r_{ij} is the distance between the i-th and the j-th non-bonded atoms. A *torsion* is defined for every set of four atoms $\{a_1, \ldots, a_4\}$ such that a_k is bonded to a_{k+1} for $k = 1, 2, 3$. The value of the torsion is equal to the angle between the two planes defined by coordinates of $\{a_1, a_2, a_3\}$ and $\{a_2, a_3, a_4\}$:



The symbols $b_{k,0}$ (equilibrium length of the k-th bond), $\theta_{\ell,0}$ (equilibrium value of the ℓ-th bond angle), n_m, δ_m, A_{ij}, and C_{ij} in (1) are constants, which are specific to the types (orbital configurations) of atoms taking part in each particular interaction (types of atoms do not change during computation). q_i is the value of electrostatic charge on the i-th atom, which also does not change during the computation. D represents electrostatic properties of the environment (solvent). Most frequently, $D = 1$, when solvent is explicitly included, or $D = cr_{ij}$, where $c \geq 1$, when solvent is not present.

The function H is a function of Cartesian coordinates of atoms in the molecule, i.e. $H : \mathbb{R}^{3N} \longrightarrow \mathbb{R}$, where N is the number of atoms. Thus, in this sense b_k, θ_ℓ, ϕ_m, and r_{ij} in (1) are also functions of a $3N$-dimensional vector. We will write them in the form of $b_k(x)$, $\theta_\ell(x)$, $\phi_m(x)$, and $r_{ij}(x)$ whenever we need to make this fact explicit.

Computations with (1) spend usually a vast majority of computer time by minimization of the potential function. The minimization procedure is either a simple local minimization (molecular mechanics), or a local minimization combined with random perturbations to the conformation of the molecule (molecular dynamics). Both methods have a common fault: they are *local* in the sense that they do not move the molecule into a completely different conformation. Local minimization cannot do it in principle and molecular dynamics can simulate only a time period that is too short even for relatively small molecules, because of limitations of current computers.

Global optimization methods have also been applied to chemical force fields. Among them, the largest number of studies concern variants of the simulated annealing method [1]. In principle, the simulated annealing algorithm guarantees to find the global minimizer of the potential function, if it has enough time to run. Nevertheless, in the search for global minima of potential functions like (1) it spends a lot of time "inefficiently" by searching high energy regions of the state space, which are usually highly populated with local minima [2]. While simulated annealing suffers from its non-specificity to computational chemistry (which implies its relative inefficiency), more specific methods based on decomposition of the molecule, like "rotamers" (see [3]) or the build-up approach (see [4]) require deep understanding of the analyzed molecule. Therefore their use is limited by the necessary careful considerations preceding the calculation itself. This situation opens a tantalizing challenge not only to theoretically-oriented chemists, but also to mathematicians and computer scientists.

2 Spatial Averaging Approach

During the past few years, a new class of conformational analysis methods has been developed [5–7]. They share the idea of minimizing a *smoothed* force field instead of the original one. All these methods can be considered as special cases of a *spatial averaging* template [8]. The value of a spatially averaged function in a given point x is, if fact, obtained as a weighted sum of values of the original function in points sampled in a neighborhood of x.

For the sake of clarity, we need to introduce some definitions so that we can discuss the spatial averaging transforms formally.

Notation. For every real matrix $A = (a_{ij})$ we define $\|A\|$ to be the Euclidean matrix norm, i.e.

$$\|A\| = \sqrt{\sum_{i,j} a_{ij}^2}.$$

The symbol *0* will denote a null vector or matrix.

Definition 1. We say that a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an *isometry*, or an *isometric transformation*, if

$$\|f(x) - f(y)\| = \|x - y\| \qquad \text{for all } x, y \in \mathbb{R}^n.$$

The general form of a *spatially averaged energy function* is based on a convolution with a sampling function p_Λ, where the parameter Λ characterizes the extent of averaging,

$$\langle H \rangle_{\Lambda,p,g}(x) = g^{-1}\left\{ \int_\Omega g[H(u)]p_\Lambda(x-u)\,du \right\}, \tag{2}$$

where
Λ is a $3N \times 3N$ real nonsingular matrix, [1]
g is a continuous real function, increasing on $(-\infty, \infty)$,
p_Λ is a continuous real function of $3N$ variables, for every Λ, and
Ω is the state space, i.e. $\Omega = \mathbb{R}^{3N}$.

The functions p and g can be chosen to satisfy the following property for all $x \in \mathbb{R}^{3N}$:

$$\lim_{\|\Lambda\| \to 0} \langle H \rangle_{\Lambda,p,g}(x) = H(x).$$

Then we can construct a sequence of transformed hypersurfaces

$$\langle H \rangle_{\Lambda_0}, \langle H \rangle_{\Lambda_1}, \ldots, \langle H \rangle_{\Lambda_n}$$

(with the same p and g), which approach the shape of the original hypersurface H. For instance, we can build the sequence by setting $\Lambda_i = \lambda_i \Lambda_0$ $(i = 0, 1, \ldots, n)$,

[1]Pardalos, Shalloway and Xue [8] used a more general formulation, where Λ is a tensor. However, the present definition covers all special cases which have been practically used.

where $1 = \lambda_0 > \lambda_1 > \cdots > \lambda_n = 0$ and Λ_0 is a given regular matrix (for the sake of completeness, we have to say that $\langle H \rangle_{0,p,g}(x) \stackrel{\text{def}}{=} H(x)$, as $\langle H \rangle_{0,p,g}$ is not defined by (2)). Under certain conditions, the minima of H can be traced through the sequence of the transformations.

An important class of transforms of the type (2) are those, which are invariant to any isometric transformation of the coordinates of points (atoms) in $\mathbb{R}^3$. Then we can consider them to be potential functions of a molecule (we assume that a potential function should be independent to transitions and rotations).

Definition 2. A real nonsingular matrix A is said to be *orthogonal*, if

$$A^{-1} = A^{\mathsf{T}}.$$

We will use the following lemmas to represent isometric transformations of structures in $\mathbb{R}^3$.

Lemma 1. *A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an isometry if and only if $f(x) = Ax + b$ for all $x \in \mathbb{R}^n$, where A is an orthogonal matrix and $b \in \mathbb{R}^n$.*

Proof. See [9]. □

Lemma 2. *If $x_1, \ldots, x_k, y_1, \ldots, y_k \in \mathbb{R}^3$ satisfy the condition*

$$\|x_i - x_j\| = \|y_i - y_j\| \qquad (1 \le i, j \le k),$$

then there exist orthogonal matrix A and vector $b \in \mathbb{R}^3$ such that

$$y_i = Ax_i + b \qquad (i = 1, \ldots, k).$$

Proof. $(x_1, \ldots, x_k)$ can be transformed into $(y_1, \ldots, y_k)$ by a sequence of transpositions and rotations around the axes. As each of these elementary transformations can be expressed in the form of $x_i' = Ax_i + b$, where A is an orthogonal matrix, the same is true also for their composition. □

Definition 3. Let $f : \mathbb{R}^{3k} \longrightarrow \mathbb{R}^{3k}$ be a linear function,

$$f(x) = Ax + b \qquad (x \in \mathbb{R}^{3k}).$$

Then f is said to be a *3-dimensional isometry*, if

$$A = \begin{pmatrix} D & & 0 \\ & \ddots & \\ 0 & & D \end{pmatrix}, \tag{3a}$$

$$b = (\underbrace{d, d, \ldots, d}_{k \text{ times}})^{\mathsf{T}}, \tag{3b}$$

where D is a 3×3 orthogonal matrix and $d \in \mathbb{R}^3$.

From Lemma 2 it can be seen that every movement of a rigid structure of k atoms in 3-dimensional space can be expressed by $x \longmapsto f(x)$, where f is a 3-dimensional isometry in $\mathbb{R}^{3k}$ (here x is a sequence of their coordinates, i.e. $x = (x_{1x}, x_{1y}, x_{1z}, \ldots, x_{kx}, x_{ky}, x_{kz})$). The following theorem specifies a sufficient condition for function

$$\langle h \rangle_{\Lambda,p,g}(x) = g^{-1}\left\{\int_{\mathbb{R}^n} g[h(u)]p_\Lambda(x-u)\,du\right\} \tag{4}$$

to be invariant to an isometric transformation. Especially, it describes a certain class of functions $\langle h \rangle_{\Lambda,p,g}$ which are invariant to 3-dimensional isometries.

Theorem 1. *Let A be an orthogonal matrix of the type $n \times n$, $b \in \mathbb{R}^n$ and let h and p_Λ be measurable real functions of n variables such that*

$$\begin{aligned} h(Ax+b) &= h(x) \\ p_\Lambda(Ax) &= p_\Lambda(x) \quad \text{for all } x \in \mathbb{R}^n. \end{aligned}$$

Then

$$\langle h \rangle_{\Lambda,p,g}(Ax+b) = \langle h \rangle_{\Lambda,p,g}(x) \qquad \text{for all } x \in \mathbb{R}^n,$$

for any increasing continuous real function g.

Proof. By (4) we have

$$\langle h \rangle_{\Lambda,p,g}(x) = g^{-1}(\mathbf{I}(x)),$$

where

$$\mathbf{I}(x) = \int_{\mathbb{R}^n} g[h(u)]p_\Lambda(x-u)\,du\,.$$

Then

$$\begin{aligned} \mathbf{I}(Ax+b) &= \int_{\mathbb{R}^n} g[h(u)]p_\Lambda(Ax+b-u)\,du \\ &= |\det A| \int_{\mathbb{R}^n} g[h(Av+b)]p_\Lambda(Ax+b-Av-b)\,dv\,, \end{aligned}$$

after the variable substitution $u = Av + b$. Because A is orthogonal, we have $|\det A| = 1$. Finally,

$$\begin{aligned} \mathbf{I}(Ax+b) &= \int_{\mathbb{R}^n} g[h(v)]p_\Lambda(x-v)\,dv \\ &= \mathbf{I}(x). \end{aligned}$$

□

3 Gaussian Transforms

In the following, we will consider a special case of the spatial averaging transform (4), which has been analyzed by Wu [7]. Here g in (4) is the identity function and p comes from the density of a multidimensional normal distribution.

Definition 4. Let h be a measurable real function of n variables and let Λ be a real nonsingular symmetric matrix of the type $n \times n$. Then the *Gaussian transform* $\langle h \rangle_\Lambda$ is defined by

$$\langle h \rangle_\Lambda (x) = \mathcal{C}_\Lambda \int_{\mathbb{R}^n} h(x-u) e^{-u^\mathsf{T} \Lambda^{-2} u} \, du \,, \tag{5}$$

or equivalently,

$$\langle h \rangle_\Lambda (x) = \mathcal{C}_\Lambda \int_{\mathbb{R}^n} h(u) e^{-(x-u)^\mathsf{T} \Lambda^{-2} (x-u)} \, du \,, \tag{6}$$

where $\Lambda^{-2} = (\Lambda\Lambda)^{-1}$ and $\mathcal{C}_\Lambda$ is a normalization constant such that

$$\mathcal{C}_\Lambda \int_{\mathbb{R}^n} e^{-u^\mathsf{T} \Lambda^{-2} u} \, du = 1. \tag{7}$$

We define also $\langle \cdot \rangle_0$ as the identity transform, i.e. $\langle h \rangle_0 (x) = h(x)$.

From the *Laplace–Gauss integral* $\int\limits_0^\infty e^{-a^2 x^2} \, dx = \frac{\sqrt{\pi}}{2a}$ (where $a > 0$) it is easy to prove that

$$\mathcal{C}_\Lambda = \frac{1}{\pi^{n/2} \, |\det \Lambda|}. \tag{8}$$

For the case when Λ is a diagonal matrix, Wu [7] has shown that the minimizers of $\langle H \rangle_\Lambda$ can be traced continuously through layers with decreasing values of $\|\Lambda\|$. This approach has been used to find potential minima for atom clusters [10] with satisfactory results. Recently, the method has also been adapted to solve distance geometry problems [11], which are important in determining molecular structure from experimental data.

Our aim is to find an algorithm for efficient approximation of (5) for the potential function (1). The case when H consists of the "non-bonded" terms only has already been analyzed in [7]. We will follow this direction to extend the theory to the general case of (1).

It is easy to verify that for all measurable functions h_1, h_2,

$$\langle h_1 + h_2 \rangle_\Lambda = \langle h_1 \rangle_\Lambda + \langle h_2 \rangle_\Lambda .$$

Thus, if H_i denotes the potential function for the i-th interaction in (1) (i.e. H_i corresponds to a bond, bond angle, torsion, or a non-bonded pair), we have

$$\begin{aligned} H &= \sum_i H_i \\ \langle H \rangle_\Lambda &= \sum_i \langle H_i \rangle_\Lambda. \end{aligned}$$

From (1) we obtain

$$\begin{aligned} \langle H \rangle_\Lambda = {} & \sum_{\text{bonds}} S_k \left\langle (b_k(x) - b_{k,0})^2 \right\rangle_\Lambda + \sum_{\substack{\text{bond} \\ \text{angles}}} B_\ell \left\langle (\theta_\ell(x) - \theta_{\ell,0})^2 \right\rangle_\Lambda \\ & + \sum_{\text{torsions}} T_m [\langle 1 \rangle_\Lambda + \langle \cos(n_m \phi_m(x) - \delta_m) \rangle_\Lambda] \\ & + \sum_{\substack{\text{nonbonded} \\ ij \text{ pairs}}} A_{ij} \left\langle 1/r_{ij}^{12}(x) \right\rangle_\Lambda - C_{ij} \left\langle 1/r_{ij}^{6}(x) \right\rangle_\Lambda + q_i q_j \left\langle 1/[D(x) r_{ij}(x)] \right\rangle_\Lambda. \end{aligned}$$

where $D(x)$ is either constant or $D(x) = c r_{ij}(x)$, $r_{ij}(x) = \|x_i - x_j\|$, $b_k(x) = r_{i_k j_k}(x)$, where i_k and j_k are indices of the atoms connected by the k-th bond, and $c > 0$. After multiplication, we get

$$\begin{aligned} \langle H \rangle_\Lambda = {} & \sum_{\text{bonds}} S_k \left[\left\langle b_k^2(x) \right\rangle_\Lambda - 2 b_{k,0} \left\langle b_k(x) \right\rangle_\Lambda + b_{k,0}^2 \left\langle 1 \right\rangle_\Lambda \right] \\ & + \sum_{\substack{\text{bond} \\ \text{angles}}} B_\ell \left[\left\langle \theta_\ell^2(x) \right\rangle_\Lambda - 2 \theta_{\ell,0} \left\langle \theta_\ell(x) \right\rangle_\Lambda + \theta_{\ell,0}^2 \left\langle 1 \right\rangle_\Lambda \right] \\ & + \sum_{\text{torsions}} T_m [\langle 1 \rangle_\Lambda + \langle \cos(n_m \phi_m(x) - \delta_m) \rangle_\Lambda] \\ & + \sum_{\substack{\text{nonbonded} \\ ij \text{ pairs}}} A_{ij} \left\langle 1/r_{ij}^{12}(x) \right\rangle_\Lambda - C_{ij} \left\langle 1/r_{ij}^{6}(x) \right\rangle_\Lambda + q_i q_j \left\langle 1/[D(x) r_{ij}(x)] \right\rangle_\Lambda, \end{aligned} \tag{9}$$

where $\langle 1 \rangle_\Lambda = \mathcal{C}_\Lambda \displaystyle\int_{\mathbb{R}^{3N}} e^{-u^\mathsf{T} \Lambda^{-2} u} \, du = 1.$

Notation. We use $u.v$ to denote the scalar product of vectors u and v. We also write $u^2 = u.u = \|u\|^2$. The expression $a \times b$ means the vector product of vectors a and b.

From (9) it can be seen that the computation of $\langle H \rangle_\Lambda$ can be reduced to evaluation of the following basic functions:

$$\left\langle r_{ij}^p(x) \right\rangle_\Lambda \qquad (p = -12, -6, -2, -1, 1, 2) \tag{10a}$$

$$\left\langle \theta_{ijk}^p(x) \right\rangle_\Lambda \qquad (p = 1, 2) \tag{10b}$$

$$\langle \cos(n_m \phi_{ijkl}(x) + \delta_m) \rangle_\Lambda \qquad (m = 1, 2, \dots) \tag{10c}$$

where

$$\cos\theta_{ijk}(x) = \frac{(x_i - x_j).(x_k - x_j)}{r_{ij}\, r_{jk}},$$

$$\cos\phi_{ijkl}(x) = \frac{[(x_j - x_k) \times (x_i - x_j)].[(x_k - x_l) \times (x_j - x_k)]}{r_{ij}\, r_{jk}^2\, r_{kl}\, \sin\theta_{ijk}(x)\, \sin\theta_{jkl}(x)},$$

and $r_{ij} = r_{ij}(x) = \|x_i - x_j\|$.

The following theorem gives a sufficient condition for the Gaussian transform to "make sense" for molecular force fields, i.e. to be invariant to isometric transformations in $\mathbb{R}^3$.

Theorem 2. *Let Λ be a $3k \times 3k$ diagonal matrix,*

$$\Lambda = \operatorname{diag}(\lambda_1, \lambda_1, \lambda_1, \lambda_2, \lambda_2, \lambda_2, \dots, \lambda_k, \lambda_k, \lambda_k), \quad \text{where } \lambda_i > 0 \quad (i = 1, \dots, N). \tag{11}$$

If $f : \mathbb{R}^{3k} \longrightarrow \mathbb{R}^{3k}$ is a 3-dimensional isometry and $h : \mathbb{R}^{3k} \longrightarrow \mathbb{R}^{3k}$ is a measurable function, which is invariant to f, then the function $\langle h \rangle_\Lambda$ is also invariant to f.

Proof. It is sufficient to verify the requirements of Theorem 1 for $p_\Lambda(x) = e^{-x^\mathsf{T} \Lambda^{-2} x}$. Let $f(x) = Ax + b$, where A and b are given by (3). Then for any $x = (x_1, \dots, x_k) \in \mathbb{R}^{3k}$,

$$\begin{aligned}
p_\Lambda(Ax) &= e^{-(Ax)^\mathsf{T} \Lambda^{-2} (Ax)} = e^{-x^\mathsf{T} A^\mathsf{T} \Lambda^{-2} A x} \\
&= \exp\left\{ -\sum\nolimits_{i=1}^k x_i^\mathsf{T} D^\mathsf{T} \left[\operatorname{diag}(\lambda_i^2, \lambda_i^2, \lambda_i^2)\right]^{-1} D x_i \right\} && \text{by (3a)} \\
&= \exp\left\{ -\sum\nolimits_{i=1}^k \frac{1}{\lambda_i^2} x_i^\mathsf{T} D^\mathsf{T} D x_i \right\} \\
&= \exp\left\{ -\sum\nolimits_{i=1}^k \frac{1}{\lambda_i^2} x_i^\mathsf{T} x_i \right\} && \text{as } D \text{ is orthogonal} \\
&= e^{-x^\mathsf{T} \Lambda^{-2} x} = p_\Lambda(x).
\end{aligned}$$

$\square$

The simplest kind of the Gaussian transformation is the case when

$$\Lambda = \mathrm{diag}(\underbrace{\lambda, \lambda, \ldots, \lambda}_{n \text{ times}}), \quad \text{where } \lambda > 0,$$

$$\langle h \rangle_\Lambda (x) = \langle h \rangle_\lambda (x) = \mathcal{C}_\lambda \int_{\mathbb{R}^n} h(x-u) e^{-\|u\|^2/\lambda^2} \, du \,, \tag{12}$$

where

$$\mathcal{C}_\lambda \int_{\mathbb{R}^n} e^{-\|u\|^2/\lambda^2} \, du = 1. \tag{13}$$

This type of transformation is called *isotropic*, as it averages the values of the potential function equally in all directions.

Anisotropic transformations (i.e. those, which are not isotropic) may also be of use in practical computations. For instance, the λ_i values in (11) assigned to the atoms of solvent can be different from the values for the atoms of a target molecule. Similarly, we can emphasize the motion of the atoms in the reaction center of an enzyme by assigning lower λ_i values to them than to the rest of the molecule.

4 Computational Feasibility

It is not practical to compute the values of a spatially averaged potential function directly by (9). The advantage of working with a spatially averaged hypersurface would easily be outweighted by the multidimensional numerical integration at every point of computation. For instance, computation of a transformed torsional term (10c) results in a 12-dimensional integration. We need a fast approximation of the transformed functions (10) to be able to apply the method to real problems.

The simplest approach to the approximation is to tabulate the transformed function for a grid of points. However, using the partial separability (9) only, we have to consider tables up to 12 dimensions for a function of the type (1), as the value of every torsion ϕ_m depends on 3×4 coordinates. If the transformation is invariant to 3-dimensional isometries (especially if it is of the type considered in Theorem 2), then it is sufficient to tabulate its value for only those elementar structures, which are in some "standard" position, like the first atom in the origin of coordinates, the second on the x axis and the third in the xy plane.

Even with this simplification, we still need tables of 6 dimensions. We will discuss some ideas, which could potentially overcome this difficulty at the end of this work. Here we will concentrate only on computation of a single value of $\langle H \rangle_\Lambda (x)$, which makes a baseline for any kind of approximation.

In the derivations, we will need the following formula:

$$\int_{-\infty}^{\infty} e^{-a^2x^2+bx}\, dx = \frac{\sqrt{\pi}}{a} e^{b^2/(4a^2)} \qquad (a > 0), \tag{14}$$

which can easily be derived from the Laplace–Gauss integral.

Lemma 3. *Let $a > 0$ and $v = (v_1, \ldots, v_n)$ be an n-dimensional real vector. Then*

$$\int_{\mathbb{R}^n} e^{-a^2x^2+v.x}\, dx = \frac{\pi^{n/2}}{a^n} e^{v^2/(4a^2)}. \tag{15}$$

Proof.

$$\begin{aligned}
\int_{\mathbb{R}^n} e^{-a^2x^2+v.x}\, dx &= \int_{\mathbb{R}^n} \exp\left[\sum\nolimits_{i=1}^{n} -a^2x_i^2 + v_ix_i\right] dx \\
&= \prod_{i=1}^{n} \int_{-\infty}^{\infty} e^{-a^2t^2+v_it}\, dt \\
&= \prod_{i=1}^{n} \frac{\sqrt{\pi}}{a} e^{v_i^2/(4a^2)} \qquad \text{by (14)} \\
&= \frac{\pi^{n/2}}{a^n} \exp\left[\frac{1}{4a^2}\sum_{i=1}^{n} v_i^2\right] \\
&= \frac{\pi^{n/2}}{a^n} e^{v^2/(4a^2)}.
\end{aligned}$$

□

Let h be a real function, which depends on coordinates of k atoms:

$$x_i = (x_{ix}, x_{iy}, x_{iz}) \qquad (i = 1, \dots, k).$$

Then for any matrix Λ of the type (11) we have

$$\langle h \rangle_\Lambda (x_1, \dots, x_k) = \mathcal{C}_{\Lambda,k}\, \mathbf{J}(x),$$

where

$$\mathbf{J}(x) = \int_{\mathbb{R}^{3k}} h(x-u) e^{-\sum_{i=1}^{k} \|u_i\|^2/\lambda_i^2}\, du\,,$$

$$\mathcal{C}_{\Lambda,k} = \pi^{-\frac{3k}{2}} \prod_{i=1}^{k} \lambda_i^{-3} \qquad \text{by (8).}$$

We suppose that function h does not change its value when the atoms are shifted in $\mathbb{R}^3$ (i.e. it depends only on the *relative* position of the atoms), which is the case of all terms in (1). Then there exists a function $\bar{h} : \mathbb{R}^{3(k-1)} \longrightarrow \mathbb{R}$ such that for all $x = (x_1, \dots, x_k) \in \mathbb{R}^{3k}$,

$$h(x_1, x_2, \dots, x_k) = \bar{h}(x_2 - x_1, x_3 - x_1, \dots, x_k - x_1).$$

Under this assumption we can reduce the $3k$-dimensional integral $\mathbf{J}(x)$ to an integral of $3(k-1)$ dimensions. We have

$$\begin{aligned}
\mathbf{J}(x) &= \int_{\mathbb{R}^{3k}} h(x-u) e^{-\sum_{i=1}^{k} u_i^2/\lambda_i^2}\, du \\
&= \int_{\mathbb{R}^{3k}} \bar{h}\big((x_2-u_2)-(x_1-u_1), \dots, (x_k-u_k)-(x_1-u_1)\big) e^{-\sum_{i=1}^{k} u_i^2/\lambda_i^2}\, du \\
&= \int_{\mathbb{R}^{3k}} \bar{h}\big((x_2-x_1)-(u_2-u_1), \dots, (x_k-x_1)-(u_k-u_1)\big) e^{-\sum_{i=1}^{k} u_i^2/\lambda_i^2}\, du\,.
\end{aligned}$$

Let $v_k = u_1$, $v_i = u_{i+1} - u_1$ $(i = 1, \dots, k-1)$ (i.e. $u_i = v_{i-1} + v_k$ for $i = 2, \dots, k$) and $\bar{x} = (x_2 - x_1, x_3 - x_1, \dots, x_k - x_1)$. Then

$$\begin{aligned}
\mathbf{J}(x) &= \int_{\mathbb{R}^{3k}} \bar{h}(\bar{x}_1 - v_1, \dots, \bar{x}_{k-1} - v_{k-1}) \exp\Big[-\tfrac{1}{\lambda_1^2} v_k^2 - \sum\nolimits_{i=1}^{k-1} \tfrac{1}{\lambda_{i+1}^2}(v_i^2 + 2 v_i.v_k + v_k^2)\Big]\, dv \\
&= \int_{\mathbb{R}^{3(k-1)}} \bar{h}(\bar{x}_1 - v_1, \dots, \bar{x}_{k-1} - v_{k-1}) \exp\Big[-\sum\nolimits_{i=1}^{k-1} \tfrac{1}{\lambda_{i+1}^2} v_i^2\Big] \\
&\qquad \times \left\{ \int_{\mathbb{R}^3} \exp\Big[-\Big(\sum\nolimits_{i=1}^{k} \tfrac{1}{\lambda_i^2}\Big) v_k^2 - 2\Big(\sum\nolimits_{i=1}^{k-1} \tfrac{1}{\lambda_{i+1}^2} v_i\Big).v_k\Big]\, dv_k \right\} dv_1 \cdots dv_{k-1}\,.
\end{aligned}$$

From (15) we obtain

$$\begin{aligned}\mathbf{J}(x) = \left[\frac{\pi}{\sigma}\right]^{\frac{3}{2}} \int_{\mathbb{R}^{3(k-1)}} & \bar{h}(\bar{x}_1 - v_1, \ldots, \bar{x}_{k-1} - v_{k-1}) \\ & \times \exp\left[\frac{1}{\sigma}\left(\sum_{i=1}^{k-1} \tfrac{1}{\lambda_{i+1}^2} v_i\right)^2 - \sum_{i=1}^{k-1} \tfrac{1}{\lambda_{i+1}^2} v_i^2\right] dv_1 \cdots dv_{k-1}\,,\end{aligned} \tag{16}$$

where $\sigma = \sum_{i=1}^{k} \frac{1}{\lambda_i^2}$.

The formula (16) can be directly employed to reduce the computational cost of transformed functions (10b) and (10c). In such a way we obtain a 6-dimesional integral instead of a 9-dimensional one for a bond angle term and a 9-dimensional integral instead of a 12-dimensional one for a torsional term.

Now we can apply (16) to an important special case when h is a pairwise distance function, i.e. $h : \mathbb{R}^6 \longrightarrow \mathbb{R}$, $h(x_1, x_2) = g(\|x_1 - x_2\|)$, where g is a one dimensional real function. In this case,

$$\bar{h}(y) = h(0, y) = g(\|y\|) \qquad (y \in \mathbb{R}^3).$$

By substitution $k = 2$ in (16) we obtain

$$\begin{aligned}\mathbf{J}(x_1, x_2) &= \left[\frac{\pi}{\sigma}\right]^{\frac{3}{2}} \int_{\mathbb{R}^3} \bar{h}(\bar{x} - v_1) \exp\left(\frac{1}{\sigma\lambda_2^4} v_1^2 - \frac{1}{\lambda_2^2} v_1^2\right) dv_1 \\ &= \left[\frac{\pi}{\sigma}\right]^{\frac{3}{2}} \int_{\mathbb{R}^3} \bar{h}(\bar{x} - v_1) \exp\left[v_1^2 \frac{1}{\sigma\lambda_2^2}\left(\frac{1}{\lambda_2^2} - \sigma\right)\right] dv_1\,,\end{aligned} \tag{17}$$

where $\sigma = \frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2}$, $\bar{x} = x_2 - x_1$.

Using the equivalence $\sigma = \frac{\lambda_1^2 + \lambda_2^2}{\lambda_1^2 \lambda_2^2}$, we can rearrange the argument of the exponential function in (17) as follows:

$$\begin{aligned}\frac{1}{\sigma\lambda_2^2}\left(\frac{1}{\lambda_2^2} - \sigma\right) &= \frac{1}{\sigma\lambda_2^2}\left(\frac{1}{\lambda_2^2} - \frac{\lambda_1^2 + \lambda_2^2}{\lambda_1^2\lambda_2^2}\right) = \frac{1}{\sigma\lambda_2^2}\,\frac{\lambda_1^2 - \lambda_1^2 - \lambda_2^2}{\lambda_1^2\lambda_2^2} = \\ &= -\frac{1}{\sigma}\,\frac{1}{\lambda_1^2\lambda_2^2} = -\frac{\lambda_1^2\lambda_2^2}{\lambda_1^2 + \lambda_2^2}\,\frac{1}{\lambda_1^2\lambda_2^2} = -\frac{1}{\lambda_1^2 + \lambda_2^2}\,.\end{aligned}$$

Thus, we have

$$\begin{aligned}
\mathbf{J}(x_1,x_2) &= \left[\frac{\pi}{\sigma}\right]^{\frac{3}{2}} \int_{\mathbb{R}^3} \bar{h}(\bar{x}-v_1) \exp\left(-\frac{1}{\lambda_1^2+\lambda_2^2}\, v_1^2\right) dv_1 \\
&= \pi^{\frac{3}{2}} \left[\frac{\lambda_1\lambda_2}{\sqrt{\lambda_1^2+\lambda_2^2}}\right]^3 \int_{\mathbb{R}^3} \bar{h}(\bar{x}-v_1) \exp\left(-\frac{1}{\lambda_1^2+\lambda_2^2}\, v_1^2\right) dv_1 \\
&= \pi^{\frac{3}{2}} \left[\frac{\lambda_1\lambda_2}{\sqrt{\lambda_1^2+\lambda_2^2}}\right]^3 \int_{\mathbb{R}^3} g(\|x_2-x_1-v_1\|) \exp\left(-\frac{1}{\lambda_1^2+\lambda_2^2}\, \|v_1\|^2\right) dv_1\,,
\end{aligned} \tag{18}$$

which is the expected symmetric formula.

5 Approximations

In the previous section we saw that the direct computation of the transformation $\langle H\rangle_\Lambda$ is impractical and its precomputation results in tables of several dimensions. The main difficulty consists in transformation of torsional terms (10c), which requires 6-dimensional tables. A suitable approximation of this part of computation would clearly make our approach more useful.

A classical way how to simplify a chemical force field is to consider approximations, which hold only in a neighborhood of equilibrium state(s) of the molecule. This can easily be applied to computation of $\langle h\rangle_\Lambda(x_1,x_2,x_3,x_4)$, where h is a bounded function of the torsional angle x_1–x_2–x_3–x_4. If bond lengths ($\|x_i-x_{i+1}\|$, $i=1,2,3$) and bond angles ($\measuredangle\, x_i$–x_{i+1}–x_{i+2}, $i=1,2$) are close enough to their equilibrium values, we can approximate $\langle h\rangle_\Lambda$ by

$$\langle h\rangle_\Lambda(x_1,x_2,x_3,x_4) \approx \langle h\rangle_\Lambda(x_1',x_2',x_3',x_4'), \tag{19}$$

where $\measuredangle\, x_1'$–x_2'–x_3'–$x_4' = \measuredangle\, x_1$–$x_2$–$x_3$–$x_4$ and bond lengths $\|x_i'-x_{i+1}'\|$ ($i=1,2,3$), as well as bond angles $\measuredangle\, x_i'$–x_{i+1}'–x_{i+2}' ($i=1,2$) have the equilibrium values of the given quadruple of atom types.

If bond lengths and bond angles in x_1–x_2–x_3–x_4 are far from equilibrium, then almost any approximation is valid: in this case the torsional term forms only a small fraction of the contribution of x_1–x_2–x_3–x_4 to the force field, as the terms for bond lengths and/or bond angles strongly dominate. One of the simplest approximations possible is

$$\langle h \rangle_\Lambda (x) \approx h(x).$$

In order to approximate $\langle h \rangle_\Lambda (x)$ for intermediate states, we can use some combination of simple approximations, like

$$\langle h \rangle_\Lambda (x) \approx \alpha(x, x')h(x) + (1 - \alpha(x, x')) \langle h \rangle_\Lambda (x'), \tag{20}$$

where $x' = (x'_1, x'_2, x'_3, x'_4)$ has the same meaning as in (19) and $\alpha(x, x')$ is a simple metric such that $0 \leq \alpha(x, x') \leq 1$.

It can be seen that we need only a one-dimensional look-up table (for $\langle h \rangle_\Lambda (x')$) to compute the torsional term approximation (20) efficiently. Nevertheless, we generally need a separate table for *each* torsion in the molecule. The large ensemble of tables can be handled efficiently, as we can link torsions to the corresponding tables at the beginning of computation, provided that there is enough fast memory available.

Acknowledgements

The author is grateful to Drs. Jaroslav Koča, Luděk Matyska, and Jiří Sochor for fruitful discussions concerning this work. This research was supported, in part, by grant 201/94/1493 from the Grant Agency of the Czech Republic.

References

[1] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[2] H. Gotō and E. Ōsawa. An efficient algorithm for searching low-energy conformers of cyclic and acyclic molecules. *J. Chem. Soc. Perkin Trans.*, 2:187–198, 1993.

[3] P. Tufféry, C. Etchebest, S. Hazout, and R. Lavery. A new approach to the rapid determination of protein side chain conformations. *J. Biomol. Struct. Dyn.*, 8:1267–1289, 1991.

[4] M. Vásquez and H. A. Scheraga. Use of buildup and energy-minimization procedure to compute low-energy structures of the backbone of enkephalin. *Biopolymers*, 24:1437–1447, 1985.

[5] L. Piela, J. Kostrowicki, and H. A. Scheraga. The multiple-minima problem in the conformational analysis of molecules: Deformation of the protein energy hypersurface by the diffusion equation method. *J. Phys. Chem.*, 93:3339–3346, 1989.

[6] D. Shalloway. Application of the renormalization group to deterministic global minimization of molecular conformation energy function. *J. Global Optimization*, 2:281–311, 1992.

[7] Z. Wu. The effective energy transformation scheme as a special continuation approach to global optimization with application to molecular conformation. Technical Report MCS-P442-0694, Argonne National Laboratory, Mathematics & Computer Science Division, Argonne, IL, July 1994. Available at `ftp://ftp.mcs.anl.gov/pub/tech_reports/reports/P442.ps.Z`.

[8] P. M. Pardalos, D. Shalloway, and G. Xue. Optimization methods for computing global minima of nonconvex potential energy functions. *J. Global Optimization*, 4:117–133, 1994.

[9] S. Mac Lane and G. Birkhoff. *Algebra*. Macmillan, NY, 1967.

[10] T. F. Coleman and Z. Wu. Parallel continuation-based global optimization for molecular conformation and protein folding. Technical Report MCS-P443-0694, Argonne National Laboratory, Mathematics & Computer Science Division, Argonne, IL, July 1994. Available at `ftp://ftp.mcs.anl.gov/pub/tech_reports/reports/P443.ps.Z`.

[11] J. J. Moré and Z. Wu. Global continuation for distance geometry problems. Technical Report MCS-P505-0395, Argonne National Laboratory, Mathematics & Computer Science Division, Argonne, IL, March 1995. Available at `ftp://ftp.mcs.anl.gov/pub/tech_reports/reports/P505.ps.Z`.

Publications in the FI MU Report Series are in general accessible via WWW and anonymous FTP:

```
http://www.fi.muni.cz/informatics/reports/
ftp  ftp.fi.muni.cz (cd pub/reports)
```

Copies may be also obtained by contacting:

Faculty of Informatics
Masaryk University
Botanická 68a
602 00 Brno
Czech Republic