ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois 60439

$\mathcal{E}$-OPTIMAL SOLUTIONS TO DISTANCE GEOMETRY PROBLEMS VIA GLOBAL CONTINUATION

Jorge J. Moré and Zhijun Wu

Mathematics and Computer Science Division

Preprint MCS-P520-0595

May 1995

This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Computational and Technology Research, U.S. Department of Energy, under Contract W-31-109-Eng-38 and by the Argonne Director's Individual Investigator Program.

ABSTRACT

We show that a continuation approach to global optimization with global smoothing techniques can be used to obtain ε-optimal solutions to distance geometry problems. We show that determining an ε-optimal solution is still an NP-hard problem when ε is small. A discrete form of the Gaussian transform is proposed based on the Hermite form of Gaussian quadrature. We show that the modified transform can be used whenever the transformed functions cannot be computed analytically. Our numerical results show that the discrete Gauss transform can be used to obtain ε-optimal solutions for general distance geometry problems, and in particular, to determine the three-dimensional structure of protein fragments.

$\mathcal{E}$-OPTIMAL SOLUTIONS TO DISTANCE GEOMETRY PROBLEMS VIA GLOBAL CONTINUATION

Jorge J. Moré and Zhijun Wu

1 Introduction

Distance geometry problems arise in the interpretation of NMR data and in the determination of protein structures. For a general review of the distance geometry problem and its relationship to macromolecular conformation, see Crippen and Havel [4], Havel [10], Kuntz, Thomason, and Oshiro [17], and Brünger and Nilges [1].

A distance geometry problem is specified by a subset $\mathcal{S}$ of all atom pairs and by the distances $\delta_{i,j}$ between atoms i and j for $(i,j) \in \mathcal{S}$. A solution to the distance geometry problem is a set of positions $x_1, \ldots, x_m$ in $\mathbb{R}^3$ such that

$$\|x_i - x_j\| = \delta_{i,j}, \qquad (i,j) \in \mathcal{S}. \tag{1.1}$$

Usually, $\mathcal{S}$ is sparse; in other words, only a small subset of distances is known.

In practice, lower and upper bounds on the distances are specified instead of their exact values. The distance geometry problem with lower and upper bounds is to find a set of positions $x_1, \ldots, x_m$ such that

$$l_{i,j} \le \|x_i - x_j\| \le u_{i,j}, \qquad (i,j) \in \mathcal{S}, \tag{1.2}$$

where $l_{i,j}$ and $u_{i,j}$ are lower and upper bounds on the distance constraints, respectively.

The distance geometry problem (1.1) is computationally intractable in general because the restriction of the distance geometry problem to a one-dimensional space is equivalent to the set partition problem, which is known to be NP-complete [6]. Even stronger results have been obtained showing that k-dimensional distance geometry problems are strongly NP-hard for all $k \ge 1$. For detailed proofs of these results, see Saxe [22] and Crippen and Havel [4]. One of the purposes of this paper is to study the complexity of solving the bounded distance geometry problem (1.2).

The distance geometry problem, exact or bounded, can be formulated as a global optimization problem. The objective function for the distance geometry problem is defined so that the constraints are satisfied at a global minimizer of the problem. Special optimization techniques for this class of problems have been developed by Crippen and Havel [4], Havel [10], Hendrickson [11, 12], Glunt, Hayden, and Raydan [8, 9], and Moré and Wu

This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Computational and Technology Research, U.S. Department of Energy, under Contract W-31-109-Eng-38 and by the Argonne Director's Individual Investigator Program.

[19]. In particular, Moré and Wu [19] used global smoothing techniques and a continuation approach to solve distance geometry problems.

We extend the work of Moré and Wu [19], by showing that the global continuation approach can also be used to obtain ε-optimal solutions to the distance geometry problem, that is, a set of positions $x_1, \ldots, x_m$ such that

$$\left| \|x_i - x_j\| - \delta_{i,j} \right| \leq \varepsilon, \qquad (i,j) \in \mathcal{S}, \tag{1.3}$$

where ε is a positive number. An ε-optimal solution to the distance geometry problem is all that can be expected in practice, with the choice of $\varepsilon > 0$ dictated by the accuracy in the data. An ε-optimal solution can also be useful when the exact solution does not exist because of small errors in the data. The latter case can happen, for example, if the triangle inequality

$$\delta_{i,j} \leq \delta_{i,k} + \delta_{k,j}$$

is violated for atoms $\{i, j, k\}$ because of possible inconsistency in the experimental data.

We study problem (1.3) as a special case of the bounded distance geometry problem (1.2). Our approach is to formulate problem (1.2) in terms of finding the global minimum of the function

$$f(x) = \sum_{i,j \in \mathcal{S}} h_{i,j}(x_i - x_j), \tag{1.4}$$

where

$$h_{i,j}(x) = \min^2 \left\{ \frac{\|x\|^2 - l_{i,j}^2}{l_{i,j}^2}, 0 \right\} + \max^2 \left\{ \frac{\|x\|^2 - u_{i,j}^2}{u_{i,j}^2}, 0 \right\}. \tag{1.5}$$

Clearly, $x = \{x_1, \ldots, x_m\}$ solves problem (1.2) if and only if x is a global minimizer of f and $f(x) = 0$ (see Crippen and Havel [4]).

In the continuation approach to global optimization an integral transformation is used to smooth the original function into a function with fewer minimizers. An optimization algorithm is then used to trace the minimizers of the transformed functions back to the original function. The smoothing transformation was first proposed and studied in the diffusion equation method for protein conformation by Scheraga and coworkers [21, 14, 15, 16, 23]. Similar transformations have been used in the packet annealing algorithm of Shalloway [25, 24], and in the algorithms used by Coleman, Shalloway, and Wu [2, 3] for molecular conformation problems. Recent development of this approach from a mathematical and computational point of view can be found in Wu [28] and Moré and Wu [19]. A general review on this approach can also be found in Pardalos, Shalloway, and Xue [20].

A major issue in applying the continuation method to problem (1.2) is that the potential function (1.4) cannot be transformed analytically. While the function can certainly be transformed by using standard techniques for numerical integration, the cost of evaluating the transformed functions would be prohibitive. In this paper we introduce a Gauss-Hermite

discrete transformation that can be evaluated at a reasonable cost and is applicable to a wide class of functions, including (1.4).

Our main concern is with the determination of ε-optimal solutions to the exact distance geometry problem, that is, vectors $x \in \mathbb{R}^n$ that satisfy (1.3). Complexity issues are examined in Section 2, where we show that determining ε-optimal solutions to distance geometry problems is an NP-hard problem when the atoms are restricted to $\mathbb{R}$.

In Section 3 we define the Gaussian transform and show how the Gaussian transform of (1.4) can be computed in terms of one-dimensional integrals. We then turn to computational issues. In particular, Section 4 introduces the discrete Gauss-Hermite transform as an approximation to the Gaussian transform, while Section 5 contains numerical results for a continuation method based on this transformation. We consider a set of model distance geometry problems, as well as the problem of determining the three-dimensional structure of a 63-atom protein fragment. Although preliminary, these results show that continuation algorithms based on smoothing techniques can solve these global optimization problems with a high degree of reliability.

2 Approximate Solutions

The solution of the exact distance geometry problem (1.1) may not exist because, for example, the constraints may not be consistent. Even if the constraints are consistent, a small perturbation may render the constraints inconsistent. Since experimental data is not 100% accurate, it is quite reasonable to search for approximate solutions to the exact distance geometry problem.

In this section we analyze the complexity of determining the approximate solution to the exact distance geometry problem (1.1). Any $x \in \mathbb{R}^n$ that satisfies (1.3) is an ε-optimal solution to th exact distance geometry problem. We wish to show that if ε is small, problem (1.3) is just as difficult as the original exact distance geometry problem. We prove this result by restricting all the atoms to lie in $\mathbb{R}$, and all distances to be positive integers. The result can be extended to higher dimensions, but we do not pursue these extensions.

Given n positive integers $a_i \in \mathbf{Z}^+$, the standard set partition problem is to find a partition S_1, S_2 of $\{1, \ldots, n\}$ such that

$$\sum_{i \in S_1} a_i = \sum_{i \in S_2} a_i.$$

This problem is one of the basic NP-hard problems (see Garey and Johnson [6]). We now generalize this problem by allowing approximate solutions. Given n positive integers $a_i \in \mathbf{Z}^+$ and $\varepsilon > 0$, the ε-set partition problem is to find a partition $S_1(\varepsilon)$, $S_2(\varepsilon)$ of

$\{1, \ldots, n\}$, and numbers $a_i(\varepsilon) \in \mathbb{R}$ such that

$$|a_i(\varepsilon) - a_i| \leq \varepsilon, \qquad \left| \sum_{i \in S_1(\varepsilon)} a_i(\varepsilon) - \sum_{i \in S_2(\varepsilon)} a_i(\varepsilon) \right| \leq \tfrac{1}{2}. \tag{2.1}$$

Clearly, any solution to the standard set partition problem is a solution of the ε-set partition problem for any $\varepsilon > 0$. This seems to suggest that the ε-set partition problem may be easier than the standard set partition problem. However, we now show that obtaining a solution to the ε-set partition problem is still NP-hard when ε is small.

Theorem 2.1 *If $a_i(\varepsilon) \in \mathbb{R}$ solves the ε-set partition problem and $n\varepsilon < \frac{1}{2}$, then*

$$\sum_{i \in S_1(\varepsilon)} a_i = \sum_{i \in S_2(\varepsilon)} a_i$$

Proof. The triangle inequality and the second inequality in (2.1) show that

$$\left| \sum_{i \in S_1(\varepsilon)} a_i - \sum_{i \in S_2(\varepsilon)} a_i \right| \leq \sum_{i \in S_1(\varepsilon)} |a_i(\varepsilon) - a_i| + \sum_{i \in S_2(\varepsilon)} |a_i(\varepsilon) - a_i| + \tfrac{1}{2}.$$

Now use the first inequality in (2.1), and recall that S_1, S_2 is a partition of $\{1, \ldots, n\}$ to obtain that

$$\left| \sum_{i \in S_1(\varepsilon)} a_i - \sum_{i \in S_2(\varepsilon)} a_i \right| \leq n\varepsilon + \tfrac{1}{2} < 1.$$

The sums in the above inequality add integers, and thus the sums must be integers. Since the difference between these two sums is less than one, they must be equal. This is the desired result. ∎

An immediate consequence of Theorem 2.1 is that the ε-set partition problem is NP-hard. We now show that determining an ε-optimal solution to the distance geometry problem in $\mathbb{R}$ is NP-hard. The proof shows that any ε-set partition problem can be directly reduced to a distance geometry problem with $n + 1$ atoms in $\mathbb{R}$.

Theorem 2.2 *Determining an ε-optimal solution to the distance geometry problem in $\mathbb{R}$ is NP-hard.*

Proof. Given an instance of the ε-set partition problem, consider the distance geometry problem with $n + 1$ atoms in $\mathbb{R}$ with

$$\delta_{i,i+1} = a_i, \quad 1 \leq i \leq n, \qquad \delta_{1,n+1} = \tfrac{1}{2}.$$

If $x_1, \ldots, x_{n+1}$ is an ε-optimal solution to this distance geometry problem, then

$$\left| |x_i - x_{i+1}| - \delta_{i,i+1} \right| \leq \varepsilon, \qquad |x_1 - x_{n+1}| \leq \tfrac{1}{2}.$$

In particular, if we define $a_i(\varepsilon) = |x_i - x_{i+1}|$, then $|a_i(\varepsilon) - a_i| \leq \varepsilon$. Now note that

$$\left| \sum_{i=1}^{n} (x_i - x_{i+1}) \right| = |x_1 - x_{n+1}| \leq \tfrac{1}{2} \tag{2.2}$$

and that, if $S_1(\varepsilon) = \{i : x_i > x_{i+1}\}$ and $S_2(\varepsilon) = \{i : x_i \leq x_{i+1}\}$, then (2.2) implies that

$$\left| \sum_{i \in S_1(\varepsilon)} |x_i - x_{i+1}| - \sum_{i \in S_2(\varepsilon)} |x_i - x_{i+1}| \right| \leq \tfrac{1}{2}.$$

Since $a_i(\varepsilon) = |x_i - x_{i+1}|$, we have shown that the $a_i(\varepsilon)$ solve the ε-set partition problem. ∎

3 Global Smoothing

In the continuation approach to global optimization, the objective function is gradually transformed into a smoother function with fewer local minimizers. An optimization algorithm is then applied to the transformed function, tracing the minimizers back to the original function. In this section we define the Gaussian transform and describe some of the interesting properties of this transformation. The emphasis is on computational considerations; motivation and additional details can be found in Wu [28] and Moré and Wu [19].

The Gaussian transform depends on a parameter λ that controls the degree of smoothing. The original function is obtained if $\lambda = 0$, while smoother functions are obtained as λ increases.

Definition 3.1 *The Gaussian transform $\langle f \rangle_\lambda$ of a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is*

$$\langle f \rangle_\lambda(x) = \frac{1}{\pi^{n/2} \lambda^n} \int_{\mathbb{R}^n} f(y) \exp\left(-\frac{\|y - x\|^2}{\lambda^2} \right) dy. \tag{3.1}$$

The value $\langle f \rangle_\lambda(x)$ is an average of f in a neighborhood of x, with the relative size of this neighborhood controlled by the parameter λ. The size of the neighborhood decreases as λ decreases so that when $\lambda = 0$, the neighborhood is the center x. The Gaussian transform $\langle f \rangle_\lambda$ can also be viewed as the expectation value of f with respect to the Gaussian density function

$$\rho_\lambda(y) = \frac{1}{\pi^{n/2} \lambda^n} \exp\left(-\frac{\|y\|^2}{\lambda^2} \right).$$

We could have used other density functions in the definition of the transform, but Gaussian density function has stronger smoothing properties.

In principle the computation of the Gaussian transform requires the evaluation of n-dimensional integrals, but for many functions that arise in practice, it is possible to compute

the Gaussian transform explicitly, or in terms of one-dimensional integrals. In particular, if the function is *decomposable*, that is, if the function $f : \mathbb{R}^n \mapsto \mathbb{R}$ can be written in the form

$$f(x) = \sum_{k=1}^{m} f_k(x), \qquad f_k(x) = \prod_{j=1}^{n} f_{k,j}(x_j),$$

for some set of functions $\{f_{k,j}\}$, where $f_{k,j} : \mathbb{R} \mapsto \mathbb{R}$, then a computation shows that

$$\langle f \rangle_\lambda(x) = \sum_{k=1}^{m} \left(\prod_{j=1}^{n} \langle f_{k,j} \rangle_\lambda(x_j) \right).$$

Thus, computing $\langle f \rangle_\lambda$ for a decomposable function requires the computation of only the one-dimensional integrals for each $\langle f_{k,j} \rangle_\lambda$. As we shall see in the next section, we avoid computing one-dimensional integrals by using special-purpose quadratures.

For the distance geometry problem, as well as for other problems in macromolecular conformation, we are interested in transforming a class of functions defined in terms of the distances between pairs of atoms. Given functions $h_{i,j} : \mathbb{R}^p \mapsto \mathbb{R}$ of the distance between atoms i and j, we consider the potential function

$$f(x) = \sum_{i,j \in \mathcal{S}} h_{i,j}(x_i - x_j), \tag{3.2}$$

where $\mathcal{S}$ is some subset of all pairs of atoms, and $x_i \in \mathbb{R}^p$ is the position of the i-th atom. In general we are concerned with three-dimensional problems where $p = 3$.

The following result of Moré and Wu [19] shows that computing the Gaussian transform of (3.2) requires only the Gaussian transform of $h_{i,j}$.

Theorem 3.2 *If $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $h : \mathbb{R}^p \mapsto \mathbb{R}$ are related by*

$$f(x) = h(P^T x),$$

for some matrix $P \in \mathbb{R}^{n \times p}$ such that $P^T P = \sigma^2 I$, then

$$\langle f \rangle_\lambda(x) = \langle h \rangle_{\sigma\lambda}(P^T x).$$

Theorem 3.2 reduces the computation of the Gaussian transform of the mapping f, which is defined on $\mathbb{R}^n$, to the computation of the Gaussian transform of h, which is defined on $\mathbb{R}^p$. As an application of this result, note that

$$\langle f \rangle_\lambda(x) = \sum_{i,j \in \mathcal{S}} \langle h_{i,j} \rangle_{\sqrt{2}\lambda}(x_i - x_j)$$

is the Gaussian transform of the potential function defined by (3.2). In this case f is defined on $\mathbb{R}^{pn}$, but $h_{i,j}$ is defined on $\mathbb{R}^p$.

In some application we can compute the Gaussian transform $\langle h_{i,j}\rangle_\lambda$ explicitly. For example, in the distance geometry problem (1.1), the function $h_{i,j} : \mathbb{R}^p \mapsto \mathbb{R}$ is of the general form

$$h(x) = \left(\|x\|^2 - \delta^2\right)^2.$$

This function is decomposable. Moreover, the Gaussian transform is explicitly given by

$$\langle h\rangle_\lambda(x) = h(x) + [3 + (p-1)]\lambda^2\|x\|^2 + \tfrac{1}{4}p(p+2)\lambda^4 - p\delta^2\lambda^2.$$

For details of this computation see Section 4 of Moré and Wu [19].

The Gaussian transform of the function defined by (1.4) requires a different approach since the element functions $h_{i,j}$ in (1.4) are not decomposable. The key observation is that the element functions $h_{i,j}$ in (1.4) are of the general form

$$h(x) = \mathcal{H}(\|x\|), \tag{3.3}$$

where $\mathcal{H} : \mathbb{R} \mapsto \mathbb{R}$ is given by

$$\mathcal{H}(s) = \min^2\left\{\frac{s^2 - l^2}{l^2}, 0\right\} + \max^2\left\{\frac{s^2 - u^2}{u^2}, 0\right\}. \tag{3.4}$$

We now show that the Gaussian transform of $h : \mathbb{R}^3 \mapsto \mathbb{R}$ can be reduced to the calculation of one-dimensional integrals.

Theorem 3.3 *Given $\mathcal{H} : \mathbb{R} \mapsto \mathbb{R}$ define $h : \mathbb{R}^3 \mapsto \mathbb{R}$ by (3.3). If $r = \|x\|$ then*

$$\langle h\rangle_\lambda(x) = \frac{1}{\lambda\sqrt{\pi}r}\int_0^\infty s\mathcal{H}(s)\left[\exp\left(-\frac{(r-s)^2}{\lambda^2}\right) - \exp\left(-\frac{(r+s)^2}{\lambda^2}\right)\right] ds. \tag{3.5}$$

Proof. First note that Theorem 3.2 shows, in particular, that $\langle h\rangle_\lambda(y) = \langle h\rangle_\lambda(x)$ for any vector y such that $\|y\| = \|x\|$. Thus, without loss of generality, we assume that $x = (0, 0, r)$ since this simplifies the proof. Now note that the definition of $\langle h\rangle_\lambda$ implies that

$$\langle h\rangle_\lambda(x) = \frac{1}{\lambda^3\sqrt{\pi}}\int_{\mathbb{R}^3}\mathcal{H}(\|y\|)\exp\left(-\frac{\|y-x\|^2}{\lambda^2}\right) dy.$$

Making a change to spherical coordinates

$$y = \begin{pmatrix} s\cos\theta \\ s\sin\theta \\ s\cos\varphi \end{pmatrix},$$

and noting that $y^T x = rs\cos\varphi$ when $x = (0, 0, r)$, we obtain that

$$\langle h\rangle_\lambda(x) = \frac{2}{\lambda^3\sqrt{\pi}}\int_0^\infty\int_0^\pi s^2\mathcal{H}(s)\exp\left(-\frac{r^2+s^2}{\lambda^2}\right)\exp\left(-\frac{rs\cos\varphi}{\lambda^2}\right)\sin\varphi\, d\varphi\, ds.$$

A computation now shows that

$$\int_0^{\pi} \exp\left(-\frac{rs\cos\varphi}{\lambda^2}\right)\sin\varphi\, d\varphi = \frac{\lambda^2}{2rs}\left[\exp\left(\frac{2rs}{\lambda^2}\right) - \exp\left(-\frac{2rs}{\lambda^2}\right)\right],$$

and this yields the desired result. ∎

Theorem 3.3 can be used to determine the Gaussian transform of h, but the computation can be poorly conditioned when r is small. The following result gives an alternative expression that can be used for small r.

Corollary 3.4 *If $\mathcal{H} : \mathbb{R} \mapsto \mathbb{R}$ and $h : \mathbb{R}^3 \mapsto \mathbb{R}$ are as in Theorem 3.3, then*

$$\lim_{x\to 0} \langle h\rangle_\lambda(x) = \frac{4}{\lambda^3\sqrt{\pi}} \int_0^{\infty} s^2\mathcal{H}(s)\exp\left(-\frac{s^2}{\lambda^2}\right)\, ds.$$

Proof. Since $r = \|x\|$, the result follows by taking the limit as $r \to 0$ in (3.5). ∎

For the general distance geometry problem the function $\mathcal{H}$ is defined by (3.4), and thus $\mathcal{H}$ is an even function. We now show that Theorem 3.5 can be simplified in this case.

Corollary 3.5 *If the mapping $\mathcal{H}$ in Theorem 3.3 is an even function, then*

$$\langle h\rangle_\lambda(x) = \frac{1}{\lambda\sqrt{\pi}r} \int_{-\infty}^{+\infty} s\mathcal{H}(s)\exp\left(-\frac{(r-s)^2}{\lambda^2}\right)\, ds. \tag{3.6}$$

Proof. By direct computation,

$$\begin{aligned}
\langle h\rangle_\lambda(x) &= \frac{1}{\lambda\sqrt{\pi}r}\left[\int_0^{\infty} s\mathcal{H}(s)\exp\left(-\frac{(r-s)^2}{\lambda^2}\right)\, ds - \int_0^{\infty} s\mathcal{H}(s)\exp\left(-\frac{(r+s)^2}{\lambda^2}\right)\, ds\right] \\
&= \frac{1}{\lambda\sqrt{\pi}r}\left[\int_0^{\infty} s\mathcal{H}(s)\exp\left(-\frac{(r-s)^2}{\lambda^2}\right)\, ds + \int_{-\infty}^{0} s\mathcal{H}(s)\exp\left(-\frac{(r-s)^2}{\lambda^2}\right)\, ds\right] \\
&= \frac{1}{\lambda\sqrt{\pi}r}\int_{-\infty}^{+\infty} s\mathcal{H}(s)\exp\left(-\frac{(r-s)^2}{\lambda^2}\right)\, ds.
\end{aligned}$$

∎

Interestingly enough, Corollary 3.5 shows that the computation of the Gaussian transform of h when $\mathcal{H}$ is an even function reduces to the computation of the transform of the mapping $s \mapsto s\mathcal{H}(s)$.

4 Gauss-Hermite Approximation

We consider techniques for the approximation of the Gaussian transform of a function $f : \mathbb{R} \mapsto \mathbb{R}$. These techniques are useful for the computation of the Gaussian transform of a decomposable function defined on $\mathbb{R}^n$, and for the computation of the Gaussian transform

of the $3n$-dimensional potential function (3.2) when the element functions $h_{i,j}$ are of the form (3.3) with $\mathcal{H}$ an even function.

The change of variables $y = x + \lambda s$ in (3.1) shows that the Gaussian transform of a mapping $f : \mathbb{R} \mapsto \mathbb{R}$ can be written in the form

$$\langle f \rangle_\lambda(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} f(x + \lambda s) \exp(-s^2)\, ds. \tag{4.1}$$

We can certainly compute $\langle f \rangle_\lambda(x)$ with standard techniques for numerical integration (for example, an adaptive quadrature), but these techniques usually require a large number of function evaluations.

Another approach used to evaluate $\langle f \rangle_\lambda(x)$ is to approximate f with a linear combination of functions ϕ_i whose Gaussian transforms can be obtained analytically. If

$$f(x) \approx \sum_{i=1}^{m} \phi_i(x),$$

then the Gaussian transform can be evaluated by

$$\langle f \rangle_\lambda(x) \approx \sum_{i=1}^{m} \langle \phi_i \rangle_\lambda(x).$$

This approach was used in [15] for the Lennard-Jones potential with $\{\phi_i\}$ a set of Gaussian functions. The validity of this approach, however, requires the approximation to be valid over a wide range. Moreover, the approximation must be recalculated if the potential function changes.

We propose an alternative approach to the approximation of the Gaussian transform (4.1) based on Gaussian quadrature. This approach mixes elements of the numerical integration and function approximation approaches, but does not require the calculation of a different approximation for each application.

Given a positive weight function $w : \mathbb{R} \mapsto \mathbb{R}$, the Gaussian quadrature formula

$$\int_{-\infty}^{+\infty} w(s) f(s) ds \approx \sum_{i=1}^{m} w_i f(s_i), \tag{4.2}$$

is exact for all polynomials of degree less than $2m$. The weights w_i are the zeros of the m-th degree orthogonal polynomial p_m with respect to the weight function w; the nodes s_i are also determined by p_m. The w-orthogonal polynomials can be generated by a three term recurrence relationship. For the weight function $w(s) = \exp(-s^2)$ the w-orthogonal polynomials are the Hermite polynomials, which are generated by the recurrence

$$p_{k+1}(x) = x p_k(x) - (\tfrac{1}{2}k) p_{k-1}(x), \qquad k \geq 0,$$

with $p_{-1} \equiv 0$, and $p_0 \equiv 1$. Details about w-orthogonal polynomials for various weight functions are discussed in standard numerical analysis textbooks, for example, Stoer and

Bulirsch [26] and Kincaid and Cheney [13]. For an advanced treatment of Gaussian quadrature, see Stroud and Secrest [27] and Davis and Rabinowitz [5].

If we use Gaussian quadrature on the integral in (4.1), we obtain an approximation

$$\langle f \rangle_{G,\lambda}(x) = \frac{1}{\sqrt{\pi}} \sum_{i=1}^{m} w_i f(x + \lambda s_i) \tag{4.3}$$

to the Gaussian transform of f, which we call the Gauss-Hermite transform of f. This terminology is appropriate because (4.3) defines a discrete transformation that shares many of the properties of the standard Gaussian transform. In particular, note that the Gauss-Hermite transform coincides with the Gaussian transform if f is a polynomial of degree less than $2m$.

The weights w_i and nodes s_i in the Gauss-Hermite transform are independent of f and can be found in the literature (for example, Stroud and Secrest [27]) or can be computed with the `gauss` subroutine in ORTHOPOL (Gautschi [7]). The computation of the Gauss-Hermite transform requires m function evaluations.

We are interested in the use of Gaussian quadratures to approximate integrals of the form (3.6) where the function $\mathcal{H} : \mathbb{R} \mapsto \mathbb{R}$ is given by (3.4). Error bounds for this type of integrals can be obtained by using the Peano kernel theorem.

Theorem 4.1 *Let $\langle f \rangle_{G,\lambda}$ be the Gauss-Hermite transform of $f : \mathbb{R} \mapsto \mathbb{R}$. If $f^{(l)}$ is piecewise continuous on $\mathbb{R}$ for some $l \leq 2m$, there is a constant μ_l, independent of f, such that*

$$|\langle f \rangle_{\lambda}(x) - \langle f \rangle_{G,\lambda}(x)| \leq \mu_l \lambda^l \sigma(x),$$

where

$$\sigma(x) = \left\{ \int_{-\infty}^{+\infty} \exp(-s^2) \left| f^{(l)}(x + \lambda s) \right|^2 ds \right\}^{1/2}.$$

Proof. Since Gaussian quadratures are exact for polynomials of degree less than $2m$, the Peano kernel theorem (for example, Davis and Rabinowitz [5, Section 4.3]) shows that

$$|\langle f \rangle_{\lambda}(x) - \langle f \rangle_{G,\lambda}(x)| \leq \lambda^l \int_{-\infty}^{+\infty} \exp(-s^2) \left| f^{(l)}(x + \lambda s) \right| \left| K_l(s) \right| ds,$$

where K_l is the Peano kernel. Hence,

$$|\langle f \rangle_{\lambda}(x) - \langle f \rangle_{G,\lambda}(x)| \leq \mu_l \lambda^l \sigma(x),$$

where $\mu_l = \|K_l\|_2$ is the weighted L_2-norm of the kernel. ∎

Theorem 4.1 shows that the Gauss-Hermite transform is likely to be a good approximation to $\langle f \rangle_{\lambda}$ provided $\lambda < 1$, but that the accuracy is likely to deteriorate if $\lambda > 1$. This is not a serious difficulty because for large λ the Gauss-Hermite transform is used to guide an algorithm to a global minimizer, but for small λ we work with the original function f.

We can apply Theorem 4.1 to the mapping $f(x) = x\mathcal{H}(x)$, where $\mathcal{H} : \mathbb{R} \mapsto \mathbb{R}$ is given by (3.4). In this case, f'' is piecewise continuous, and thus Theorem 4.1 applies with $k = 2$ provided we use $m \geq 1$ nodes in the quadrature. Also note that in this case f is a piecewise polynomial of degree 5, and thus we can expect better results provided we use $m \geq 3$ nodes.

5 Numerical Results

We now present an algorithm for determining a solution to the bounded distance geometry problem (1.2). Our results are restricted to the problem of obtaining ε-optimal solutions to the distance geometry problem. In particular, we present preliminary numerical results on the determination of the three-dimensional structure of a 63-atom protein fragment with experimental data provided by Bruce Hendrickson. All computations were done on the IBM SP parallel system at Argonne National Laboratory with 128 nodes, each node an IBM RS/6000-370 with 128 Mbytes of memory and 125 Mflops of peak performance.

We attack the bounded distance geometry problem (1.2) by seeking a global minimizer of the function

$$f(x) = \sum_{i,j \in \mathcal{S}} h_{i,j}(x_i - x_j), \tag{5.1}$$

where

$$h_{i,j}(x) = \min^2 \left\{ \frac{\|x\|^2 - l_{i,j}^2}{l_{i,j}^2}, 0 \right\} + \max^2 \left\{ \frac{\|x\|^2 - u_{i,j}^2}{u_{i,j}^2}, 0 \right\}. \tag{5.2}$$

This function is appropriate for the bounded distance geometry problem because $f(x) \geq 0$ for all $x \in \mathbb{R}^n$, and $f(x) = 0$ if and only if x solves (1.2).

The computation of the Gaussian transform of f follows from the results in Section 3. Theorem 3.2 shows that

$$\langle f \rangle_\lambda(x) = \sum_{i,j \in \mathcal{S}} \langle h_{i,j} \rangle_{\sqrt{2}\lambda}(x_i - x_j).$$

The functions $h_{i,j}$ defined by (5.2) are not decomposable, but in this case Corollary 3.5 is applicable because

$$h_{i,j}(x) = \mathcal{H}_{i,j}(\|x\|), \tag{5.3}$$

where

$$\mathcal{H}_{i,j}(r) = \min^2 \left\{ \frac{r^2 - l_{i,j}^2}{l_{i,j}^2}, 0 \right\} + \max^2 \left\{ \frac{r^2 - u_{i,j}^2}{u_{i,j}^2}, 0 \right\} \tag{5.4}$$

is an even function. Hence, Corollary 3.5 shows that

$$\begin{aligned} \langle h_{i,j} \rangle_{\sqrt{2}\lambda}(x) &= \frac{1}{\lambda\sqrt{2\pi}\, r} \int_{-\infty}^{+\infty} s\mathcal{H}_{i,j}(s) \exp\left(-\frac{(r-s)^2}{2\lambda^2} \right) ds \\ &= \frac{1}{\sqrt{\pi}\, r} \int_{-\infty}^{+\infty} (r + \sqrt{2}\lambda s)\mathcal{H}_{i,j}(r + \sqrt{2}\lambda s) \exp\left(-s^2 \right) ds, \end{aligned}$$

where $r = \|x\|$. In summary, these computations show that

$$\langle f \rangle_\lambda(x) = \sum_{i,j \in \mathcal{S}} \frac{1}{\sqrt{\pi}\, r_{i,j}} \int_{-\infty}^{+\infty} (r_{i,j} + \sqrt{2}\lambda s) \mathcal{H}_{i,j}(r_{i,j} + \sqrt{2}\lambda s) \exp\left(-s^2\right)\, ds,$$

where $r_{i,j} = \|x_i - x_j\|$, is the Gaussian transform of the function defined by (5.1), (5.3), and (5.4). This expression for the Gaussian transform is valid for any function defined by (5.1) and (5.3), for any even function $\mathcal{H}_{i,j}$.

We approximate the Gaussian transform with the Gauss-Hermite transform as described in Section 4. For this problem the Gauss-Hermite transform is

$$\langle f \rangle_{G,\lambda}(x) = \sum_{i,j \in \mathcal{S}} \frac{1}{\sqrt{\pi}\, r_{i,j}} \sum_{k=1}^{n_q} w_k (r_{i,j} + \sqrt{2}\lambda s_k) \mathcal{H}_{i,j}(r_{i,j} + \sqrt{2}\lambda s_k), \tag{5.5}$$

where w_k and s_k are the weights and nodes for the Gaussian quadrature, and $r_{i,j} = \|x_i - x_j\|$.

We use a simple version of the global continuation algorithm that uses a prespecified set $\lambda_0 > \lambda_1 > \ldots > \lambda_p = 0$ of smoothing parameters.

Algorithm GMIN

Choose a random vector $x_0 \in \mathbb{R}^{m \times 3}$ and set $f_g = +\infty$.
for $k = 0, 1, \ldots, p$
  Determine x^* = **locmin** $(\langle f \rangle_{G,\lambda_k},\ x_0)$.
  Set $x_0 = x^*$.
  If $f(x^*) < f_g$ then $x_g = x^*$ and $f_g = f(x^*)$.
end do

We did not attempt to optimize the choice of smoothing parameters or the starting point for the local minimization procedure **locmin** in **GMIN**. We chose

$$\lambda_k = \lambda_0 \left(1 - \frac{k}{p}\right), \qquad 0 \le k \le p,$$

and the starting point of **locmin** as the local minimizer found at the previous iteration. We will show that these choices work remarkably well.

We tested **GMIN** with $\lambda_0 > 0$, but we also tested **GMIN** with a choice of $\lambda_0 = 0$, since for this choice **GMIN** reduces to the use of **locmin** on the original function f from a random starting point. A standard multistart method is obtained if **GMIN** is used from a set of randomly generated starting points with $\lambda_0 = 0$.

The choice of local minimization algorithm has to be done with some care because $\langle f \rangle_{G,\lambda_k}$ is not twice continuously differentiable. The Hessian matrix is discontinuous at points where the argument of $\mathcal{H}_{i,j}$ coincides with either $l_{i,j}$ or $u_{i,j}$. We cannot expect to avoid these discontinuities, in particular, if $l_{i,j}$ or $u_{i,j}$ are close. We used the variable-metric

Table 5.1: Results for the model problem, $\varepsilon = 0.1$

λ_0	$m = 64,\ r = 16$	$m = 125,\ r = 25$	$m = 216,\ r = 36$	$m = 343,\ r = 49$
1	70%	30%	60%	40%
0	0%	0%	0%	0%

limited-memory code VMLM in MINPACK-2, which is an implementation of the Liu and Nocedal [18] algorithm.

In our numerical results we consider **GMIN** to be successful from a starting point x_0 if x^* satisfies

$$\left| \|x_i - x_j\|^2 - \delta_{i,j}^2 \right| \leq \varepsilon\, \delta_{i,j}^2, \qquad (i,j) \in \mathcal{S}, \tag{5.6}$$

for a given $\varepsilon > 0$. This termination test may not be appropriate in practice because smaller distances tend to be determined more accurately. We used **GMIN** with

$$l_{i,j}^2 = \delta_{i,j}^2(1-\varepsilon), \qquad u_{i,j}^2 = \delta_{i,j}^2(1+\varepsilon).$$

In general, $\varepsilon = 0.1$ is a reasonable value, but we also tested **GMIN** with smaller ε.

We first tested the algorithm with the model distance geometry problem constructed by Moré and Wu [19]. In this problem the molecule has $m = s^3$ atoms located in the three-dimensional lattice

$$\{(i_1, i_2, i_3) :\ 0 \leq i_1 < s,\ 0 \leq i_2 < s,\ 0 \leq i_3 < s\}$$

for some integer $s \geq 1$. The problem is to determine the structure of this molecule if we are given $\delta_{i,j} \in \mathcal{S}$, where

$$\mathcal{S} = \{(i,j) :\ |i-j| \leq r\},$$

and r is an integer between 1 and m.

Numerical results for this problem are presented in Table 5.1. We used molecules with $(m, r) = (s^3, s^2)$, where $4 \leq s \leq 7$. For each problem, we ran **GMIN** from 10 randomly chosen starting points with $\lambda_0 = 1$ and $p = 20$ for continuation. The results show that the continuation method found the ε-optimal solutions to the problems successfully. On the other hand, the multistart method failed to find a solution in all cases.

We also tested **GMIN** with a 63-atom protein fragment. For this problem, 236 pairwise distances are specified. Hendrickson [11] showed that the molecule with this set of distance data is nearly rigid and difficult to determine. Our numerical results confirm this finding.

We ran **GMIN** from 100 starting points for $\varepsilon = 0.01, 0.02, 0.04, 0.06, 0.08, 0.1$. In all cases we set $\lambda_0 = 1$ and $p = 20$ for continuation. The results from these experiments,

Table 5.2: Results for the 63-atom protein fragment

λ_0	$\varepsilon = .01$	$\varepsilon = .02$	$\varepsilon = .04$	$\varepsilon = .06$	$\varepsilon = .08$	$\varepsilon = .10$
1	0%	4%	30%	69%	76%	96%
0	0%	0%	8%	23%	40%	66%

given in Table 5.2, clearly show that the problem becomes harder as ε is decreased. The continuation method failed for $\varepsilon = 0.01$, but was successful for all other larger values of ε. If compared with the multistart method, note that the continuation method has a significantly higher probability of success.

Acknowledgment

We thank Bruce Hendrickson for providing the data for the problems of determining the structures of protein fragments.

References

[1] A. T. Brünger and M. Nilges, *Computational challenges for macromolecular structure determination by X-ray crystallography and solution NMR-spectroscopy*, Q. Rev. Biophys., 26 (1993), pp. 49–125.

[2] T. F. Coleman, D. Shalloway, and Z. Wu, *Isotropic effective energy simulated annealing searches for low energy molecular cluster states*, Comp. Optim. Applications, 2 (1993), pp. 145–170.

[3] ——, *A parallel build-up algorithm for global energy minimizations of molecular clusters using effective energy simulated annealing*, J. Global Optim., 4 (1994), pp. 171–185.

[4] G. M. Crippen and T. F. Havel, *Distance Geometry and Molecular Conformation*, John Wiley & Sons, 1988.

[5] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Academic Press, second ed., 1984.

[6] M. R. Garey and D. S. Johnson, *Computers and Intractability*, W. H. Freeman, 1979.

[7] W. Gautschi, *Algorithm 726:* ORTHOPOL *– A package of routines for generating orthogonal polynomials and Gauss-type quadrature rules*, ACM Trans. Math. Software, 20 (1994), pp. 21–62.

[8] W. Glunt, T. L. Hayden, and M. Raydan, *Molecular conformation from distance matrices*, J. Comp. Chem., 14 (1993), pp. 114–120.

[9] ——, *Preconditioners for distance matrix algorithms*, J. Comp. Chem., 15 (1994), pp. 227–232.

[10] T. F. Havel, *An evaluation of computational strategies for use in the determination of protein structure from distance geometry constraints obtained by nuclear magnetic resonance*, Prog. Biophys. Mol. Biol., 56 (1991), pp. 43–78.

[11] B. A. Hendrickson, *The molecule problem: Determining conformation from pairwise distances*, Ph.D. thesis, Cornell University, Ithaca, New York, 1991.

[12] ——, *The molecule problem: Exploiting structure in global optimization*, SIAM J. Optimization, (1995). To appear.

[13] D. Kincaid and W. Cheney, *Numerical Analysis*, Brooks/Cole Publishing Company, 1991.

[14] J. Kostrowicki and L. Piela, *Diffusion equation method of global minimization: Performance for standard functions*, J. Optim. Theory Appl., 69 (1991), pp. 269–284.

[15] J. Kostrowicki, L. Piela, B. J. Cherayil, and H. A. Scheraga, *Performance of the diffusion equation method in searches for optimum structures of clusters of Lennard-Jones atoms*, J. Phys. Chem., 95 (1991), pp. 4113–4119.

[16] J. Kostrowicki and H. A. Scheraga, *Application of the diffusion equation method for global optimization to oligopeptides*, J. Phys. Chem., 96 (1992), pp. 7442–7449.

[17] I. D. Kuntz, J. F. Thomason, and C. M. Oshiro, *Distance geometry*, in Methods in Enzymology, N. J. Oppenheimer and T. L. James, eds., vol. 177, Academic Press, 1993, pp. 159–204.

[18] D. C. Liu and J. Nocedal, *On the limited memory BFGS method for large scale optimization*, Math. Programming, 45 (1989), pp. 503–528.

[19] J. J. Moré and Z. Wu, *Global continuation for distance geometry problems*, Preprint MCS-P505-0395, Argonne National Laboratory, Argonne, Illinois, 1995.

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

[21] 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 (1989), pp. 3339–3346.

[22] J. B. Saxe, *Embeddability of weighted graphs in k-space is strongly NP-hard*, in Proc. 17th Allerton Conference in Communications, Control and Computing, 1979, pp. 480–489.

[23] H. A. Scheraga, *Predicting three-dimensional structures of oligopeptides*, in Reviews in Computational Chemistry, K. B. Lipkowitz and D. B. Boyd, eds., vol. 3, VCH Publishers, 1992, pp. 73–142.

[24] D. Shalloway, *Application of the renormalization group to deterministic global minimization of molecular conformation energy functions*, J. Global Optim., 2 (1992), pp. 281–311.

[25] D. Shalloway, *Packet annealing: A deterministic method for global minimization, application to molecular conformation*, in Recent Advances in Global Optimization, C. Floudas and P. Pardalos, eds., Princeton University Press, 1992, pp. 433–477.

[26] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, 1980.

[27] A. H. Stroud and D. Secrest, *Gaussian Quadrature Formulas*, Prentice-Hall, Inc., 1966.

[28] Z. Wu, *The effective energy transformation scheme as a special continuation approach to global optimization with application to molecular conformation*, Preprint MCS-P442-0694, Argonne National Laboratory, Argonne, Illinois, 1994.