Computational complexity of the ground-state determination of atomic clusters

LETTER TO THE EDITOR

Computational complexity of the ground-state determination of atomic clusters

L T Wille and J Vennik

Laboratorium voor Kristallografie en Studie van de Vaste Stof, Rijksuniversiteit Gent, Krijgslaan 281, B9000 Gent, Belgium

Abstract. We prove that determining the ground state of a cluster of identical atoms, interacting under two-body central forces, belongs to the class of NP-hard problems. This means that as yet no polynomial time algorithm solving this problem is known and, moreover, that it is very unlikely that such an algorithm exists. It also suggests the need for good heuristics.

The study of the morphology of small aggregates of atoms is important in a number of fields, e.g. nucleation theory, astrophysics, electronic structure calculations for atomic clusters, etc (Hoare 1979). An important improvement over hard-sphere models is obtained by assuming that the atoms are interacting via two-body central forces, derived, for example, from the Lennard-Jones or the Morse potentials (Hoare and Pal 1971). Using pair potentials $v(r)$ between the atoms, the potential energy function of an N-atom cluster is defined as

$$V(r^N) = \tfrac{1}{2} \sum_{\substack{i,j \\ i \neq j}}^{N} v(r_i - r_j). \tag{1}$$

This function determines the potential energy surface in the $3N$-dimensional configuration space of all possible vectors $r^N = (r_1, r_2, \ldots, r_N)$, where r_i is the position vector of the ith atom in real space. In practice one uses a reduced configuration space of $3N-6$ dimensions by excluding translations and rotations of the cluster as a whole. The potential energy function $V(r^N)$ represents the total energy of the N atoms held motionless in the positions r_i, by assumption of zero temperature. The potential energy surface is a complicated geometrical object, which possesses a number of local minima and saddle points. The most efficient algorithms for the determination of the local minima appear to be those based on the steepest descent technique (Hoare 1979). The number of local minima·has been studied by Hoare and McInnes (1976, 1983) for Lennard-Jones and Morse potentials. This quantity seems to increase exponentially with the number of atoms in the cluster and this puts a limit on the cluster sizes that can be studied. It would be very interesting to investigate whether this great multiplicity also occurs in real clusters.

The binding energy of the most stable configuration is an important quantity about which very little is known. The results of Hoare and McInnes (1983) indicate that the energy gap between the most stable and next stable isomers is surprisingly large. This is in agreement (except for $N = 6$) with self-consistent quantum mechanical calculations

for Na clusters by Martins *et al* (1984). Another property of the absolute minimum configurations is the occurrence of non-crystallographic structures, e.g. those with fivefold symmetry. It would be very interesting to extend pair potential calculations to larger N values, in order to confirm these trends, and in this letter we investigate whether fast algorithms for the ground-state determination are possible. For this purpose we have to study the computational complexity of the cluster problem.

An important direction in the field of computational complexity concerns the study of non-deterministic polynomial-time complete (NP-complete) problems (Garey and Johnson 1979). Many classical problems in mathematics and computer science belong to this class. These problems are all equivalent in the sense that either all of them or none of them may be solved using polynomial-time algorithms. In spite of concentrated efforts in the last decades no polynomial-time algorithms for the problems in the class have been found and they are therefore considered to be intractable. It is generally believed that no such algorithms exist, although this conjecture has not been proved. There is some evidence that it might not be provable within the framework of formal set theory (Hartmanis and Hopcroft 1976). The term NP-complete is used for decision problems, whereas the corresponding optimisation problems are called NP-hard, since a polynomial-time algorithm for these problems would also solve the NP-complete problem. Recently it has been proved that two formulations of the frustration model of a three-dimensional spin–glass are NP-hard (Barahona 1982, Bachas 1984, see also Kirkpatrick 1979). In physics, optimisation problems arise quite naturally and it seems safe to speculate that many notoriously difficult physical problems are NP-hard. In this letter we will prove that a certain plausible formulation of the cluster ground-state problem indeed belongs to this class.

A discrete version of the cluster problem is the following. 'A number of points in real space is given and the interactions between any two points are known. In which way can N points be occupied so as to minimise the sum of their interactions?'. A few remarks are in order here. If the interactions have a repulsive character near the origin and tend to zero at infinity with a minimum in between, then it is obvious that the atoms will tend to be in a finite region of space and so a finite number of points will be sufficient. However, the number of points might as well be infinite, since this can always be reduced to the previous case by putting the interactions equal to $+\infty$ in the unwanted region. Furthermore, by taking the distance between neighbouring points arbitrarily small one can always approximate to an arbitrary degree of accuracy the minimal configuration corresponding to the continuum case. Using a discrete version of the problem is equivalent to the use of finite precision arithmetic in, for example, a steepest-descent method. Note that we do not specify that the points form a regular grid, nor do we force the interactions to be the same between points that lie the same distance apart. We return to this last point later.

Next we transform the discrete cluster problem to a graph theoretical version (Mehlhorn 1984). A graph $G = (V, E)$ consists of a set of vertices V and a set of edges, i.e. (unordered) pairs of vertices, E. We will deal with cliques, in which there is an edge between any two different vertices. If the edges are given weights $w(e)$, $\forall e \in E$, equal to the interaction between the vertices, then the discrete cluster problem is equivalent to selecting N different vertices $\{i; i = 1, \ldots, N\}$ so that

$$\tfrac{1}{2} \sum_{e=(i,j)} w(e) = \text{minimal.} \tag{2}$$

In order to show that the cluster problem is NP-hard we will use a proof by restriction

(Garey and Johnson 1979), i.e. we will show that it contains a known NP-hard problem as a special case. Thus a polynomial-time algorithm for the general problem would also solve the subproblem in polynomial time.

The problem that will be shown to be a special instance of the cluster optimisation is the well known travelling salesman problem (Garey and Johnson 1979, Held *et al* 1984). In its general form one asks for a minimal tour between N cities $\{c_i; i = 1, \ldots, n\}$ with known distances $d(c_i, c_j)$ (not necessarily obeying a triangle inequality). This problem remains NP-hard under severe restrictions, e.g. even if the distances are either 1 or 2. It is also NP-hard if the Euclidean distance is used (Papadimitriou 1977). We now transform the general travelling salesman problem as follows. Construct a graph with $N(N-1)$ vertices labelled by $(c_i, c_j) i, j = 1, \ldots, N; i \neq j$.

The weights of the edges,

$$e = \{(c_i, c_j), (c_k, c_l)\} \tag{3}$$

are defined as

$$w(e) \begin{cases} = 0 & \text{if } i \neq k, j \neq l \\ = +\infty & \begin{cases} \text{if } i = k, j \neq l \\ \text{if } j = l, i \neq k \\ \text{if } l = i, k = j \end{cases} \\ = d(c_k, c_l) & \text{if } l = i, k \neq j \\ = d(c_i, c_j) & \text{if } k = j, l \neq i \end{cases} \tag{4}$$

(see figure 1). Selecting N vertices, $\{(c_{\pi(i)}, c_{\pi(i+1)}); i = 1, \ldots, N, c_{\pi(N+1)} = c_{\pi(1)}\}$, so that the sum of the weights is minimal, evidently yields a minimal tour $\langle c_{\pi(1)}, c_{\pi(2)}, \ldots, c_{\pi(N)} \rangle$ solving the travelling salesman problem. Thus it is clear that the travelling salesman problem is a special instance of the cluster optimisation problem. Since the subproblem is NP-hard, so must be the general problem. This completes the proof of the NP-hardness of the cluster problem.

We conclude with a few remarks. In our formulation of the cluster problem we have not assumed that the interactions between points that lie.the same distance apart
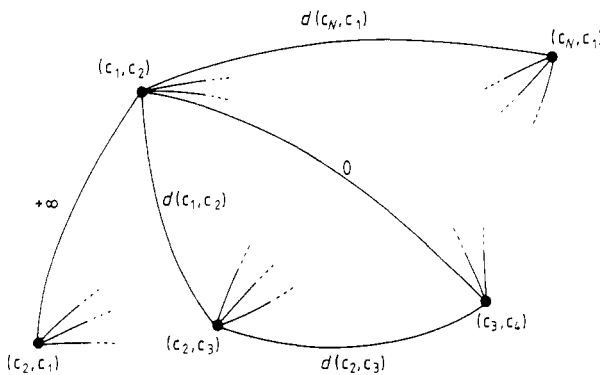

Figure 1. Graph representation of the travelling salesman problem. The $N(N-1)$ vertices are labelled by pairs of cities and the edges have certain weights associated with them. The problem is to select N vertices, so as to minimise the sum of the weights on the edges between selected vertices.

have the same value, i.e. the pair potential was not assumed to be spherically symmetric. However, even under this assumption the problem remains NP-hard, as can easily be seen by comparing it to the general and the Euclidean travelling salesman problem (Garey and Johnson 1979). Nevertheless, it is possible that extra assumptions on the nature of the interactions reduces the cluster problem to a tractable one. This is the case, for example, for the frustration model, which is NP-hard in three dimensions (Barahano 1982, Bachas 1984), but the two-dimensional version is solvable in poly-nomial time (Bieche *et al* 1980, Barahona *et al* 1982). Finally, we note that the problem of enumerating the number of local minima in a cluster (Hoare and McInnes 1976, 1983) must also be NP-hard. Indeed, a polynomial-time algorithm for this problem must necessarily give a number of local minima which is a polynomial function of the number of atoms in the cluster. It is then trivial to determine the absolute minimum configuration in polynomial time, in contradiction with the fact that this problem is NP-hard.

Proving that a problem is NP-hard is not a goal in itself, but rather an indication that a heuristic method of solution, giving a near-optimal solution in polynomial time is needed. Recently, a very interesting optimisation method for NP-complete problems was proposed by Kirkpatrick *et al* (1983). It is based on the similarity between combinatorial optimisation and statistical mechanisms. The method uses the Metropolis procedure (Metropolis *et al* 1953), but considers the temperature as a control parameter. By slowly decreasing this parameter ('cooling'), the system can be 'frozen' into the state of minimal energy. The applicability of this approach to optimisation problems was analysed by Vannimenus and Mézard (1984). The method was applied to the travelling salesman problem (Kirkpatrick *et al* 1983) and to the optimisation of computer design (Kirkpatrick *et al* 1983, Siarry and Dreyfus 1984). We are currently investigating the applicability of this 'simulated annealing' procedure to the cluster problem and will present our results in future publications.

References

Bachas C P 1984 *J. Phys. A: Math. Gen.* **17** L709
Barahona F 1982 *J. Phys. A: Math. Gen.* **15** 3241
Barahona F, Maynard R, Rammal R and Uhry J P 1982 *J. Phys. A: Math. Gen.* **15** 673
Bieche I, Maynard R, Rammal R and Uhry J P 1980 *J. Phys. A: Math. Gen.* **13** 2553
Garey M and Johnson D 1979 *Computers and Intractability* (New York: Freeman)
Hartmanis J and Hopcroft J E 1976 *SIGACT News* **8** 13
Held M, Hoffman A J, Johnson E L and Wolfe P 1984 *IBM J. Res. Dev.* **28** 476
Hoare M R 1979 *Adv. Chem. Phys.* **40** 49
Hoare M R and McInnes J 1976 *Faraday Discuss. Chem. Soc.* **61** 12
—— 1983 *Adv. Phys.* **32** 791
Hoare M R and Pal P 1971 *Adv. Phys.* **20** 161
Kirkpatrick S 1979 *Ill Condensed Matter* ed R Balian, R Maynard and G Toulouse (Amsterdam: North-Holland) p 329
Kirkpatrick S, Gellat C D and Vecchi M P 1983 *Science* **220** 671
Martins J L, Buttet J and Car R 1984 *Phys. Rev. Lett.* **53** 655
Mehlhorn K 1984 *Graph Algorithms and NP-completeness* (Berlin: Springer)
Metropolis N, Rosenbluth A, Rosenbluth M N, Teller A H and Teller E 1953 *J. Chem. Phys.* **21** 1087
Papadimitriou C H 1977 *Theor. Comput. Sci.* **4** 237
Siarry P and Dreyfus G 1984 *J. Physique Lett.* **45** L39
Vannimenus J and Mézard M 1984 *J. Physique Lett.* **45** L1145