Close packing in curved space by simulated annealing

1987 J. Phys. A: Math. Gen. 20 L1211

(http://iopscience.iop.org/0305-4470/20/17/014)

J. Phys. A: Math. Gen. **20** (1987) L1211-L1218. Printed in the UK

LETTER TO THE EDITOR

Close packing in curved space by simulated annealing

L T Wille†

SERC Daresbury Laboratory, Daresbury, Warrington WA4 4AD, UK

Received 24 August 1987

Abstract. The problem of packing spheres of a maximum radius on the surface of a four-dimensional hypersphere is considered. It is shown how near-optimal solutions can be obtained by packing soft spheres, modelled as classical particles interacting under an inverse power potential, followed by a subsequent hardening of the interaction. In order to avoid trapping in high-lying local minima, the simulated annealing method is used to optimise the soft-sphere packing. Several improvements over other work (based on local optimisation of random initial configurations of hard spheres) have been found. The freezing behaviour of this system is discussed as a function of particle number, softness of the potential and cooling rate. Apart from their geometric interest, these results are useful in the study of topological frustration, metallic glasses and quasicrystals.

Densely packed arrangements of hard spheres in the three-dimensional Euclidean space E^3 have been investigated for a long time, since they provide the necessary building blocks for a description of solids and liquids. However, the subject is also of independent mathematical interest and (in higher dimensions) has important applications in the design of efficient codes for data transfer (Sloane 1984). The same problem in the spherical space S^2 consists in finding the closest packing of N circles on the surface of a sphere and is relevant to coordination problems in stereochemistry. Only for certain specific cases, usually corresponding to small N values, are rigorous geometric proofs available to show that particular configurations are optimal. In general, one has to proceed numerically and there will be no guarantee that the resulting structures are the best possible. Lists of solutions have been tabulated by Melnyk *et al* (1977) and more recently by Clare and Kepert (1986). Going one dimension higher, tesselations of the four-dimensional hypersphere S^3 have been proposed by Sadoc and co-workers as generic models for amorphous structures (see the review by Venkataraman and Sahoo (1985)). Computer results for packings in S^3 have been given by Mackay (1980). The central structure in S^3 is the regular polytope $\{3, 3, 5\}$ which has 120 vertices, all in an ideal icosahedral environment. Unlike the icosahedron in E^3 this system has no topological frustration and therefore S^3 can be tesselated perfectly with tetrahedra. This property has been used to study the effects of frustration on the freezing temperature for soft spheres (Straley 1984, 1986). The present letter uses computer simulations of freezing to obtain efficient packings in S^3.

The determination of the ground-state configuration of a system of N interacting particles is a difficult problem, mainly because of the presence of a large number of local minima. It has been argued (Wille and Vennik 1985a) that this problem is

† Present address: Department of Materials Science and Mineral Engineering, University of California, Berkeley, CA 94720, USA.

0305-4470/87/171211+08$02.50

NP-hard, a property which it shares with such difficult optimisation problems as graph bipartitioning (Fu and Anderson 1986), spin glasses (Bachas 1984), etc. The same is true for the packing of hard objects in two dimensions which remains NP-hard even under very restrictive conditions (Fowler *et al* 1981). An additional difficulty of packing hard spheres is that, in general, the optimal configuration will be non-rigid and might contain holes that are large enough to allow spheres to 'rattle'. This means that it may be difficult to decide if a particular structure cannot be further improved by constraining these free spheres. The standard method for obtaining packings starts from a random initial configuration in which points are moved away from each other until the minimum angular distance cannot be improved. The corrections to a given configuration are obtained by solving a set of non-linear equations for the coordinates, based on an estimate for the optimal distance (Mackay (1980), Clare and Kepert (1986): the latter authors use a preliminary soft-sphere optimisation to generate a starting structure). This procedure requires a good deal of manual intervention in adjusting the target distance and moreover can easily get trapped in local minima that are bad approximations to the global one.

It has become clear in recent years that a very efficient way of obtaining near-optimal solutions for difficult optimisation problems is provided by the simulated annealing method (Kirkpatrick *et al* 1983). This technique is based on the observation that, when a liquid is slowly cooled, it ends up in a configuration that has very few (energetically unfavourable) defects. Hence its energy is close to that of the ideal perfectly crystalline solid. The reason for this is that fluctuations at finite temperature allow the atoms to surmount potential energy barriers and thus to escape from local minima. If the cooling were to proceed infinitely slowly the system would end up in the true ground state, but in practice one finds an energy distribution whose position and width depend on the cooling rate. This behaviour can be studied by the techniques of statistical mechanics and also in computer simulations. In the Metropolis prescription of the Monte Carlo method, a small random displacement $\Delta \boldsymbol{r}$ is attempted for each atom in turn and the change in energy ΔE is calculated. The atom is put in its new position if $\exp(-\Delta E/T)$ is larger than a randomly selected number from the interval $[0, 1]$, otherwise the system is left unchanged. In either case the procedure is repeated for the next atom. It is known that for a sufficiently large number of moves this gives a proper description of the system in equilibrium at temperature T. By slowly reducing the temperature and re-equilibrating one would obtain a chain of configurations whose energies converge to the true ground state. In reality relaxation times become very long at low temperatures, equilibrium cannot be maintained during the simulation and therefore only an approximate solution will be obtained. The important contribution of Kirkpatrick *et al* (1983) was that they realised that this procedure can readily be applied to discrete optimisation problems by replacing the energy with an appropriate cost function and by considering the temperature as a control parameter. Algorithms based on the simulated annealing technique have been very successful in providing good approximate solutions to a large variety of discrete and continuous optimisation problems (for an overview see Wille (1986a)—an informal publication obtainable from W Smith at Daresbury).

The hard-sphere packing problem involves the maximisation of the minimum distance between points and has a remarkable potential energy surface (or, more accurately, 'cost function surface'). Moving a sphere that is sitting in a hole, i.e. one that has no neighbours on the minimal distance D, involves no change in energy. On the other hand, for truly hard spheres, moves that decrease D would cost an infinite

amount of energy. Clearly it would be useful to have soft spheres during the simulation. This can be done by considering a system of N particles interacting under a repulsive power law potential

$$V(r_{ij}) = (r_{ij})^{-n}$$

where r_{ij} is the Euclidean distance. The hard-sphere limit is obtained when $n \to \infty$ (with an appropriate distance scaling factor). However, in general the optimal packing of soft spheres will not correspond to that for hard spheres, nor is there any guarantee that a 'hardening' of the soft-sphere optimum will lead to the hard-sphere optimum. Some investigations of this point in the three-dimensional case were done by Melnyk *et al* (1977). These authors found that (for $2 \leq N \leq 16$) the configuration with $n = 100$ agreed with that for $n \to \infty$, except in the case $N = 15$ when a crossover seems to occur when $n \approx 2000$. Likewise, Clare and Kepert (1986) used a soft-sphere interaction (but no annealing procedure) with $n = 5000$–$10\,000$ to generate input configurations for the hardening routine. Introducing a power law potential is a form of 'energy surface sculpting', i.e. the hard-sphere potential energy surface is transformed into a smoother surface with an 'undulating' energy landscape. This has the advantage that low-lying minima become more easily accessible, but the disadvantage of introducing additional minima and displacing others. The proper choice of the power law poses a dilemma: if n is too small the deformation of the hard-sphere surface is too extensive and the soft-sphere minima do not lead to good hard-sphere minima. On the other hand, if n is too large, relaxation times become very long and the annealing algorithm converges very slowly. The best choice can only be found by trial and error and will also depend on the number of particles. This is confirmed by the results of actual simulations as will now be discussed.

The calculations followed the pattern set out in previous communications (Wille and Vennik 1985b, Wille 1986b, 1987). The simulation was started at a high enough initial temperature so that the system could be considered as molten. For each particle in turn a random displacement was attempted according to the Metropolis algorithm. This procedure was repeated a number of times at a fixed temperature (typically 1000 iterations were performed, unless stated otherwise), after which the temperature was reduced by a factor $\chi_T = 0.9$. This cooling schedule was repeated until the system was frozen in a (local) minimum; subsequently, the local optimisation was accomplished by means of a steepest descent routine. This then provided low-lying local minima of the soft-sphere potential. Next a 'hardening' of the potential was realised by using a steepest descent optimiser to maximise the minimum distance between particles. This optimiser used function values only and its convergence was, of course, slower than the other routine based on analytical gradient information. Nevertheless, this method gave adequate results if the desired accuracy was not too high and it may have some advantages over the matrix equation approach used in previous work (Mackay 1980, Clare and Kepert 1986).

A list of minima obtained in this way is presented in table 1. For comparison, Mackay's results (Mackay 1980) are also given and the n value for which the minimum was attained is tabulated. The results for $N = 3$–8 are known to be optimal; that for $N = 24$ is believed to be so. In general, the values found by the present approach are better than those of Mackay (1980). There are however a number of cases in which the latter's packings are superior. The small discrepancy for $N = 12$ is probably due to a loss of accuracy in the hardening routine and leads to an estimate of -0.005 for the error in the present results. In the other cases $N = 14$, 16, 18 the convergence of

Table 1. Packing of points on a four-dimensional hypersphere. Closed packing of N hard spheres of maximum radius on a four-dimensional hypersphere found in the present letter. D is the angular distance (in degrees), n is the power used in the soft-sphere simulation, D' is the result of Mackay (1980), 'opt' means that a result is known to be optimal.

N	D	n	D'	$D-D'$
3	120	40	120	opt
4	109.47	40	109.47	opt
5	104.48	40	104.48	opt
6	90	40	90	opt
7	90	40	90	opt
8	90	40	90	opt
9	80.662	40	80.64	0.022
10	80.406	40	80.406	0
11	76.678	60	76.669	0.009
12	75.519	100	75.522	−0.003
13	71.978	100	71.98	0
14	71.247	100	71.440	−0.193
15	69.436	100	69.227	0.209
16	67.076	120	67.191	−0.115
17	65.584	70	65.196	0.388
18	64.916	120	64.939	−0.023
19	64.258	60	64.117	0.141
20	64.254	60	64.070	0.184
24	60	40	60	0
25	57.254	80	55.58	1.674

the annealing scheme was slower than for the other particle numbers; this is partly reflected in the higher n values for $N = 16$ and 18. This indicates that for these N values the system has a relatively high degree of frustration, leading to glassy relaxation behaviour. It is not clear if the fact that these numbers are even contributes to this phenomenon (note that $N = 12$ also fits this pattern). The actual structures corresponding to these packings will be described elsewhere and the remainder of this letter will be devoted to a discussion of the freezing behaviour of these curved space systems.

The Monte Carlo algorithm generates a Markov chain of states, whose macroscopic parameters fluctuate around their average value according to a Boltzmann distribution at each temperature T. An important quantity is the average energy $\langle E(T)\rangle$, obtained by taking the average over all configurations generated at a given temperature and where the previous configuration is counted again if an attempted step is rejected. The derivative of $\langle E\rangle$ with respect to T is the specific heat $C(T)$, which is a measure of the state of order in the system. A large value of C indicates the onset of freezing, i.e. the system getting trapped in a minimum. It is not necessary to take a numerical derivative of the energy to obtain $C(T)$, in view of the fluctuation-dissipation theorem:

$$C(T) = T^{-2}(\langle E^2\rangle - \langle E\rangle^2).$$

As mentioned before, relaxation times becomes very long near the phase transition and equilibrium cannot be maintained in the simulation. This means that it is not certain that the global minimum will be found, but rather a distribution of minima will be obtained, whose mean value should converge towards the energy of the global minimum as the cooling rate is decreased. This can be accomplished in the simulation by increasing the number of attempted steps and/or the cooling factor χ_T (although

the latter was kept fixed in the present work). A further consequence of the critical slowing down near the transition is that the freezing temperature T_f is hard to locate exactly. In general the system will behave like an undercooled liquid and the freezing temperature will be underestimated (see Wille 1987). Bearing these considerations in mind the thermodynamic properties of a slowly annealed inverse power law system in curved space will now be discussed.

Figure 1 shows the freezing temperature, determined by the major peak in the specific heat, as a function of particle number N. These results were obtained for $n = 80$ and a fixed cooling scheme of 1000 attempted steps for each particle and $\chi_T = 0.9$. It should be noted that each data point corresponds to a single simulation and that T_f varies slightly in different runs. No averaging over several calculations was performed, since the main purpose was to investigate the general trends. The curve shows the expected increase in T_f with increasing particle number, in agreement with simulations of Lennard-Jones particles in three dimensions (Briant and Burton 1975, Kaelberer and Etters 1977). The curve also seems to saturate towards a 'bulk' freezing temperature, but this should not be taken too literally since the temperature grid in this region was rather coarse, and the 'bulk' freezing temperature should only be expected at much larger particle numbers. Finally it needs to be re-emphasised that the T_f are the freezing temperatures for an undercooled system. The actual freezing temperatures can be obtained by much slower cooling rates ($\approx 10^6$ attempted steps or more) and may be higher by 40% (an estimate based on the results of Kaelberer and Etters (1977)). Next, figure 2 shows the effect of different power laws on T_f. These simulations were for a nine-particle system, the other parameters being the same as before. These results are shown in a semi-logarithmic plot in view of the scaling property of the inverse power law potential (Hoover *et al* 1971): the detailed dynamic evolution with identical scaled initial conditions and with the same value of $T^{3/n}/\nu$, where ν is the available volume per particle, are identical. Thus one expects $(T_f)^{3/n}$ to be a constant in a Monte Carlo simulation if the same random numbers and scaled initial conditions and step lengths are used. If, on the other hand, the parameters in the simulation are fixed and n is increased, the probability of overcoming energy barriers decreases and relaxation times

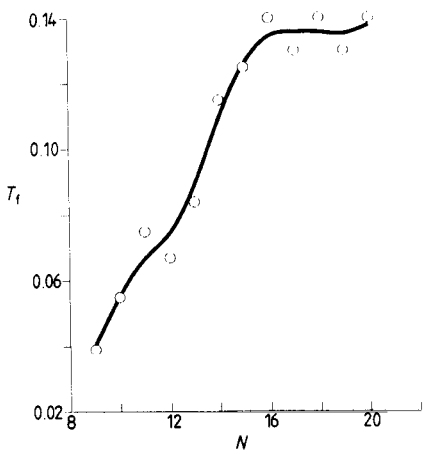

Figure 1. Freezing temperature (determined by the major peak in the specific heat) as a function of particle number. Interparticle potential of the form r^{-80}, 1000 attempted steps per particle at each temperature, cooling factor $\chi_T = 0.9$. The curve is drawn to guide the eye.

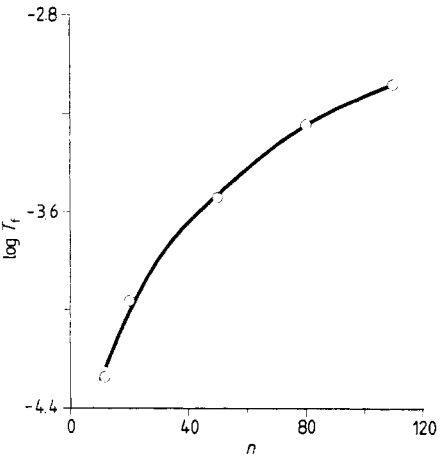

Figure 2. Freezing temperature for a nine-particle system as a function of the softness of the potential (r^{-n}); other parameters same as in figure 1. The curve is drawn to guide the eye.

increase accordingly. Therefore the T_f of the undercooled system should deviate more and more from a straight line $\log(T_f) \propto n$, as n is increased, in agreement with figure 2.

It is clear that the role of the cooling rate is central to the simulated annealing idea. If cooling is too fast local minima with a much higher energy than the global one will be found. The extreme limit of this case consists in the quenching of random initial configurations. Since the number of minima is believed to grow exponentially with the particle number this approach will fail for all but the smallest systems. On the other hand, a very slow cooling rate is very time consuming and gives no guarantee of actually finding the global minimum. This point has been investigated in Monte Carlo simulations by Grest *et al* (1986) for a (discrete) spin-glass model. These authors find that $\varepsilon(\tau)$, the difference between the ground-state energy and the average value, at a given inverse cooling rate τ, obeys an empirical (asymptotic) law:

$$\varepsilon(\tau) \propto (\log \tau)^{-1}.$$

On the other hand, Huse and Fisher (1986) find on very general grounds

$$\varepsilon(\tau) \propto (\log \tau)^{-\zeta}$$

where ζ is a system-dependent parameter. Both of these relations were obtained for the case of a linear cooling scheme, i.e. T was replaced by $T - \Delta T$, whereas the present letter uses a logarithmic cooling scheme ($T \rightarrow \chi_T T$). However, this distinction should be irrelevant in the asymptotic regime. Figure 3 shows the effect of changes in the number of iterations (τ) on the resulting energy for a nine-particle system ($n = 12$, $\chi_T = 0.9$). In figure 3(a) the results of five runs are shown, whereas figure 3(b) depicts the average value of these runs as a function of $1/\log(\tau)$. Note that the potential energy of the system is plotted, although the purpose of the present work was to generate close packings. For a small nine-particle system, however, all minimum distances were quite close, so that the energy gives a much better indication of the quality of the optimisation. Again it is not the intention to provide reliable statistics, but rather to analyse general trends. Figure 3(b) shows the expected decrease in average energy with increasing number of attempted steps. The uncertainty in the data does not rule out a $1/\log(\tau)$ behaviour, although a slightly faster convergence appears

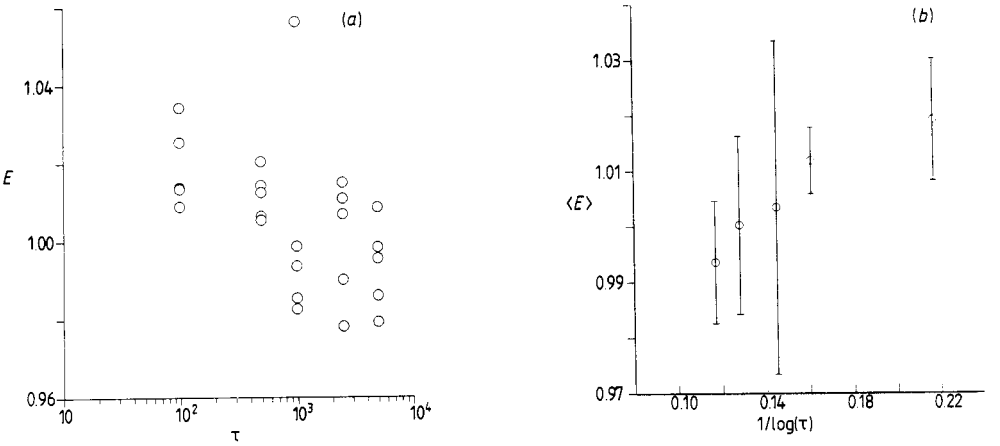

Figure 3. (*a*) Final energies for five runs of a nine-particle system as a function of τ (the number of attempted steps per particle); interparticle potential r^{-12}. (*b*) Average energy of the previous results is compatible with a dependence $\langle E\rangle \propto (\log \tau)^{-\zeta}$.

to be the case. Figure 3(*a*) reveals the distribution of the minima around the averages of figure 3(*b*). It is interesting to note that the highest minimum is obtained for 1000 iterations and the lowest one for 2500 iterations, although the average values are monotonically decreasing. This illustrates clearly the difficulty of deciding on an appropriate cooling rate. From these results one may conclude that the value of 1000 attempted steps used in the other calculations gives a proper balance between the energy distribution and optimal use of computer resources.

In summary, the simulated annealing algorithm has been used to optimise the potential energy of a system of soft spheres distributed on the surface of a four-dimensional hypersphere. The resulting configuration has been 'hardened' to obtain efficient close packings on this surface. This method seems to be superior to previous algorithms based on local optimisation only and its relative efficiency is expected to increase with increasing particle number. The freezing of the soft-sphere system, which is interesting in its own right, has been discussed and the influence of the cooling rate has been investigated. It is concluded that the simulated annealing algorithm gives low-lying minima, but that the determination of the actual ground state with a large probability would necessitate impractical cooling rates. Moreover, the present problem is additionally complicated by the need for soft spheres characterised by an inverse power law r^{-n} in the simulation. If the power n is low so that energy barriers can be overcome relatively easily, the resulting minima bear little relation to the hard-sphere minima; alternatively if n is larger the barrier heights increase and with them the relaxation times.

I should like to thank Dr Alan Mackay for bringing this problem to my attention and for many useful discussions.

References

Bachas C P 1984 *J. Phys. A: Math. Gen.* **17** L709-12
Briant C L and Burton J J 1975 *J. Chem. Phys.* **63** 2045-58

Clare B W and Kepert D L 1986 *Proc. R. Soc.* A **405** 329-44
Fowler R J, Paterson M S and Tanimoto S L 1981 *Inf. Proc. Lett.* **12** 133-7
Fu Y and Anderson P W 1986 *J. Phys. A: Math. Gen.* **19** 1605-20
Grest G S, Soukoulis C M and Levin K 1986 *Phys. Rev. Lett.* **56** 1148-51
Hoover W G, Gray S G and Johnson K W 1971 *J. Chem. Phys.* **55** 1128-36
Huse D A and Fisher D S 1986 *Phys. Rev. Lett.* **57** 2203-6
Kaelberer J B and Etters R D 1977 *J. Chem. Phys.* **66** 3233-9
Kirkpatrick S, Gelatt C D and Vecchi M P 1983 *Science* **220** 671-80
Mackay A L 1980 *J. Phys. A: Math. Gen.* **13** 3373-9
Melnyk T W, Knop O and Smith W R 1977 *Can. J. Chem.* **55** 1745-61
Sloane N J A 1984 *Sci. Am.* **250** 92-101
Straley J P 1984 *Phys. Rev.* B **30** 6592-5
—— 1986 *Phys. Rev.* B **34** 405-9
Venkataraman G and Sahoo D 1985 *Contemp. Phys.* **26** 579-615
Wille L T 1986a *Inf. Q. Comp. Sim. Cond. Phases* **20** 19-32
—— 1986b *Nature* **324** 46-8
—— 1987 *Chem. Phys. Lett.* **133** 405-10
Wille L T and Vennik J 1985a *J. Phys. A: Math. Gen.* **18** L419-22
—— 1985b *J. Phys. A: Math. Gen.* **18** L1113-7