The packing of three-dimensional spheres on the surface of a four-dimensional hypersphere

1980 J. Phys. A: Math. Gen. 13 3373

(http://iopscience.iop.org/0305-4470/13/11/011)

J. Phys. A: Math. Gen. 13 (1980) 3373–3379. Printed in Great Britain

The packing of three-dimensional spheres on the surface of a four-dimensional hypersphere

A L Mackay

Department of Crystallography, Birkbeck College (University of London), Malet Street, London WC1E 7HX, UK

Received 30 November 1979, in final form 21 March 1980

Abstract. Results are given for computer searches for those configurations of N points on the surface of a four-dimensional hypersphere which have maximum values of the least angular distance between pairs of points. The generalised inverse was used to solve the equations giving the corrections to the hyperspherical coordinates in terms of the overlaps of the domains of neighbouring spheres. It is concluded that the use of this curved space offers an alternative to cyclic boundary values for simulations involving the packing of points in three dimensions.

1. Introduction

The Fejes problem (Fejes Toth 1952) consists in finding the distribution of N points on the surface of a sphere such that the least distance between any two points is a maximum. This is, of course, equivalent to packing equal circles or spheres of maximum radius on the surface of a sphere. The problem has been reviewed at intervals (Whyte 1952, Coxeter 1962) and the fullest table of putative solutions (running up to $N = 60$) is due to Székely (1974). In general, solutions are found only by trial and error and in only a few cases are proofs available that particular solutions are optima. Configurations are usually established by demonstration and remain until superseded by counterdemonstration.

In order to search for possible solutions a computer program (in BASIC) has been developed using time-sharing facilities. The algorithm is simply to take a random arrangement of points and to move them simultaneously away from each other until they are the maximum distance apart. Results for the three-dimensional Fejes problem will be published separately. It was shown that there are often a number of configurations with values of D (the value in degrees of the least distance between a pair of points) sometimes within 0·1° of the best packing. Such configurations with similar packing densities may be inaccessible one from the other without considerable rearrangement.

This program could easily be adapted for any desired measure of the distance between points and it has been used for the present study of the four-dimensional Fejes problem. Hyperspherical coordinates were used to define the position of a point on the surface of a hypersphere of unit radius:

$$x_1 = r \cos \theta$$

0305-4470/80/113373+07$01.50

$$x_2 = r \sin\theta \cos\phi$$

$$x_3 = r \sin\theta \sin\phi \cos\psi$$

$$x_4 = r \sin\theta \sin\phi \sin\psi.$$

With $r = 1$ this ensures that $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$.

2. Calculation

The angular distance d_{ij} between two points i and j can be calculated from an extension of the cosine rule, readily obtained from the scalar product of the two position vectors $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ as

$$\cos d_{ij} = \cos\theta_i \cos\theta_j + \sin\theta_i \sin\theta_j(\cos\phi_i \cos\phi_j + \sin\phi_i \sin\phi_j \cos(\psi_i - \psi_j)).$$

The program proceeds as follows.

(1) Set N, the number of points, and D, the target distance in degrees.

(2) Place the N points at random on the hypersphere. $\theta = \cos^{-1}(x_1)$ is chosen so that x_1 lies uniformly in the range $+1$ to -1 (θ is between 0° and 180°). The remaining angles are chosen to lie uniformly in the range 0° to 360°.

(3) Calculate the angular distances d_{ij} for all pairs. Select those for which d_{ij} is less than D, i.e. overlapping pairs.

(4) Make a table of all points concerned in overlaps and of the amounts by which they overlap. M points are involved in the P distances which are less than D. Since each point has three positional coordinates, we must calculate corrections to the $3M$ coordinates from the P values of discrepant distances. Usually $3M$ is greater than P, that is, there are more unknowns (the corrections to the coordinates) than there are data items (the values of the overlaps).

(5) The matrix M_{ij} of differential coefficients, describing how each of the P distances changes when each of the six coordinates on which it depends is changed, is calculated from formulae obtained by differentiating the cosine rule given above.

(6) We then have the matrix equation relating the $3M$ corrections $\Delta\theta_j$ to be applied to the coordinates to the P discrepancies Δh_i, $\Delta h_i = M_{ij}\Delta\theta_j$. This equation is solved by inverting M_{ij}, which is usually rectangular. It has no ordinary inverse, but a routine which calculates the generalised (Moore–Penrose) inverse is used. This gives a solution which minimises $\|\Delta h_i - M_{ij}\Delta\theta_j\|$.

(7) These corrections are applied, the new distances between points are calculated, and the whole procedure is repeated until the target distance is reached or approached. At each cycle the minimum distance between points is printed out.

(8) The target distance is reset at intervals until the points settle into a stable close-packed configuration. This is then printed out as a table of distances and of hyperspherical coordinates. From these a graphical plot or network of nearest neighbours can be obtained.

3. Calculation of the generalised inverse

The core of the program is the use of the generalised inverse. This is very convenient for handling molecular configurations where (in three dimensions) N atoms have $3N$

Cartesian coordinates but only $3N-6$ independent parameters describing their mutual positions and only $3N-6$ normal modes of vibration. In the present case of N points on a hypersphere, there are also $3N$ coordinates and only $3N-6$ are independent.

If $AX=H$ is a set of linear equations and A^+ is the generalised inverse of A, then the solution $X=A^+H$ gives the correct answer if A is non-singular. If there are more equations than unknowns (and these may be observational equations inconsistent with each other), then $X=A^+H$ gives the least-squares solution which minimises $\|AX-H\|$. If there are fewer equations than unknowns part of the solution remains arbitrary, the full solution being $X=A^+H+[I-A^+A]Z$ where Z is an arbitrary vector. If the arbitrary part is omitted, then $X=A^+H$ minimises the shifts applied.

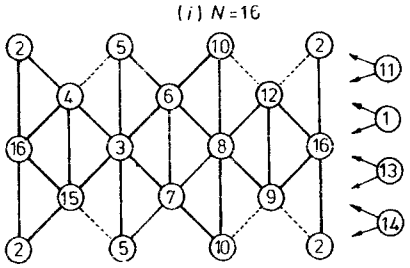
Routines for the generalised inverse are available in ALGOL and in FORTRAN in the Numerical Algorithms Group suite of programs, but here the inverse has been calculated by an iterative method (Gupta 1971). If S_k is the kth approximation to A^+ then $S_{k+1}=S_k(2I-AS_k)$. Iteration is continued until the trace of $(2I-AS_k)$, which converges to the rank of A, is sufficiently close to an integer.

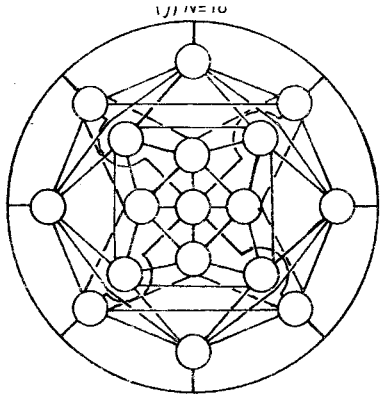
In a physical sense the configuration of points on the hypersphere floats, since no particular points are fixed as defining axes. As no points are fixed, none is treated

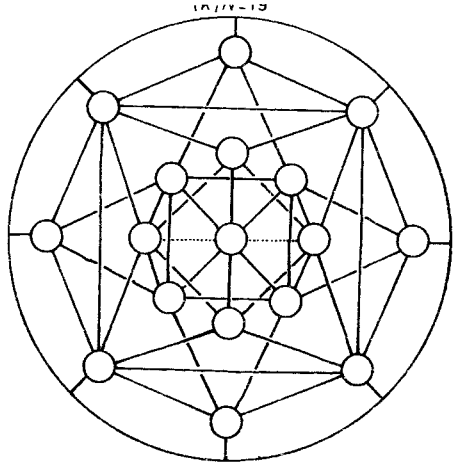
Table 1. The packing of points on the surface of a four-dimensional hypersphere. D is the maximum value found for the least distance between pairs of points among the N points on the surface. It is quoted in degrees.

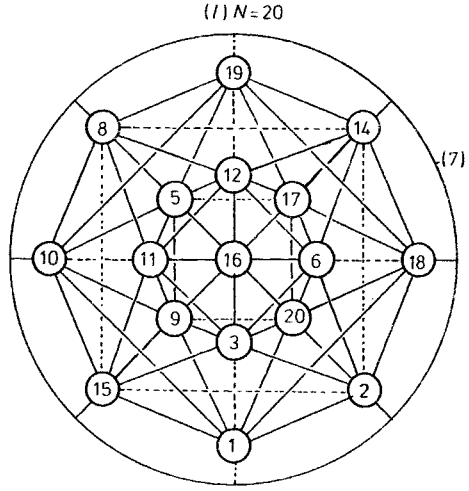
N	D	D (predicted)	Coordination
3	$120=\cos^{-1}(-\frac{1}{2})$	120·512	
4	$109{\cdot}471=\cos^{-1}(-\frac{1}{3})$	109·492	
5	$104{\cdot}478=\cos^{-1}(-\frac{1}{4})$	101·644	4-dimensional simplex 4 at $\cos^{-1}(-\frac{1}{4})$
6	90	95·650	Cross polytope less two points
7	90	90·860	Cross polytope less one point
8	90	86·904	Cross polytope 6 at 90; 1 at 180
9	>80·64	83·558	
10	$80{\cdot}406=\cos^{-1}(\frac{1}{6})$	80·675	Midpoints of edges of simplex 6 at $\cos^{-1}(\frac{1}{6})$; 3 at $\cos^{-1}(-\frac{2}{3})$
11	>76·669	78·152	
12	$75{\cdot}522=\cos^{-1}(\frac{1}{4})$	75·918	Ring of 4 octahedra 6 at $\cos^{-1}(\frac{1}{4})$; 1 at 90; 2 at 120; 2 at $\cos^{-1}(-\frac{3}{4})$
13	>71·98	73·919	
14	>71·440	72·115	
15	>69·227	70·476	
16	>67·191	68·976	
17	>65·196	67·596	
18	>64·939	66·320	
19	>64·117	65·136	
20	>64·070	64·032	
24	$60=\cos^{-1}(\frac{1}{2})$	60·256	
25	>55·58	59·441	
120	36	35·237	600-cell

(a) N=9
(b) N=10
(c) N=11
(d) N=12
(e) N=13
(f) N=14
(less close-packed configuration)
(g) N=14
(closest - packed configuration)
(h) N=15
(pent 5 is connected to points 9,4,7,3,6,1)
1
2
3
4
6
7
8
9
10
11
12
13
14









Superimposed points - connected
13: 2 6 8 11 14 15 3 12 (7)
4: 1 5 9 10 17 18 19 20 (16)

Figure 1. The figures represent Schlegel diagrams of some of the more symmetrical packings which were encountered. The fully regular polytopes in four dimensions are well known but the less regular ones are relatively unexplored and obviously very numerous. The symmetry in the configurations shown was not introduced but arose naturally as the refinements converged to symmetrical configurations which may or may not be absolutely the maximum separations for each given value of N.

In the Schlegel diagrams full lines denote contacts with the minimum angular distance and broken lines mark somewhat greater distances. The outer circle represents one of the points. In some cases certain points have been removed to clarify the diagrams and their contacts are referred to by numbers. Table 1 lists the maximum value of the separations found for each value of N. The predicted values of D are calculated from $ND^3 = (180/\pi)^3 4\pi^2/(2)^{1/2}$, which is the value corresponding to the density of the face-centred-cubic packing.

The case $N = 17$ is very unsymmetrical and has not been shown because no tidy Schlegel figure could be found. For $N = 14$ the less symmetrical configuration is slightly more closely packed than the more symmetrical. For the case $N = 16$ the network of contacts is differently represented to show the symmetry better. Points 11, 1, 13 and 14 are taken to be along the axis of a cylinder, on which the other points lie. The cylinder is shown unrolled. The axial points mentioned are each connected to the three points in the layer above and to the three points in the layer below. The configuration is not totally symmetrical but could be made so by moving 10 so that it touched 12 and 9 insteady of 6 and 7. Full lines represent contacts at 64·19° and broken lines near contacts at 78·17°.

specially. The method is very similar to the physical model-building procedure of moving circles around on a sphere until they do not overlap. The generalised inverse is a simple way of applying the technique of eigenvalue filtering practised in the geometrical refinement of crystal structures. In some cases the representative point moving in $3N$-dimensional space may be trapped in a configuration which is a local maximum and not a global maximum. Setting the value of D well ahead enables the procedure to jump out of some such traps.

4. Results

Table 1 gives the maximum separations achieved for various values of N. The results up to $N = 8$ are already known to be the best possible while $N = 24$ and $N = 120$ are believed to be the best possible. Figures $1(a)$–(l) show the corresponding Schlegel diagrams for the connectivity of the assemblies, full lines denoting contacts and broken lines showing somewhat more remote approaches of one sphere to another. Probably the concept of the Schlegel diagram could be generalised to higher dimensions, but this has not been done and thus lines in the diagram overlap. In most cases the configuration converged to a symmetrical arrangement. There can be no guarantee that this local maximum represents absolutely the best packing. For example, for $N = 14$ two different configurations were found at 71·036° and 70·579° which were within half a degree of each other. In most cases refinement was stopped when it seemed clear that a maximum was being approached. The fit between the observed values and those calculated from the face-centred-cubic packing density is reasonably close.

5. Relationship to the three-dimensional packing of spheres

In packing points at equal distance on the surface of a four-dimensional hypersphere we are effectively packing three-dimensional spheres in a three-dimensional space which is curved, that is, where the metric is non-linear.

Suppose that we are examining the packing of equal circles in a plane where the hexagonally close-packed lattice configuration predominates (so that no less regular configuration approaches it in density). This predominance can be removed by changing the metric so that distances no longer add arithmetically and thus the best packing is not a lattice of regular repetitions. This change of metric can be accomplished by curving the space in which the circles are embedded. This gives the Fejes problem, the solution to which for a given N is by no means obvious.

Similarly, if we are examining the packings of spheres in three-dimensional space, lattice packings are the densest known, although they do not necessarily predominate. It was thought that non-lattice packings could be sampled by changing the metric, that is, by curving the three-dimensional space in which the spheres are embedded. This gives the four-dimensional Fejes problem.

The particular value of packing density found experimentally for the closest random (non-lattice) packing of equal spheres is $0{\cdot}6366 \pm 0{\cdot}0004$ (Gotoh and Finney 1974), which is within experimental error equal to $2/\pi$. We may enquire whether this value is a chance coincidence or whether it has some fundamental significance. It is, of course, the mean value of the cosine and occurs in alternating current theory.

The packing density (fraction of the three-dimensional surface of the four-dimensional hypersphere filled by the N three-dimensional spheres whose centres are at least D apart) can be calculated from the expression $\delta V = (r^3 \sin^2\theta \sin\phi)\, dr\, d\theta\, d\phi\, d\psi$. Integrated over the whole surface $\int dV$ gives $2\pi^2$ as expected for the hypersphere.

If we take as an index of maximum packing density the face-centred-cubic packing density which is $\pi/\sqrt{18} = 0{\cdot}74048$, we can calculate the expected maximum values of D for given values of N from $ND^2 = 4\pi^2/\sqrt{2}$ (here D is in radians). These values are given in column 3 of table 1 and are seen to agree reasonably with the course of those found by the relaxation method. Only the most symmetrical packings for $N = 5$, 8, 20 and 120 exceed the expected values. Tables of the coordinates θ, ϕ and ψ for each packing and of the distances between all pairs from which the exact configurations reported here can be reconstructed are available from the author.

Acknowledgments

Computing was done using the time-sharing facilities of Imperial College Computer Centre, University of London and I am grateful to the computer staff for their advice and to Linda Miller of this department for help with the diagrams. One of the referees gave detailed comments which will also be useful for the further progress of this topic.

References

Coxeter H S M 1962 *Trans. NY Acad. Sci., Ser. II* **24** 320–31
Fejes Toth L 1952 *Lagerung in der Ebene, auf der Kugel und im Raum* (Berlin: Springer)
Gotoh K and Finney J L 1974 *Nature* **252** 202–5
Gupta N N 1971 *IEEE Trans. Systems, Man and Cybernetics, January 1971* pp 89–90
Székely E 1974 *Ann. Univ. Sci. Budapest, Sect. Math.* **17** 157–75
Whyte L L 1952 *Am. Math. Mon.* **59** 606–11