Random-sequential-packing simulations in three dimensions for spheres

Douglas W. Cooper
IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, New York 10598
(Received 5 February 1988)

The goal of one three-dimensional random-parking (occupation) -limit problem is to determine the mean fraction of space that would be occupied by fixed equal-size spheres, created at random locations sequentially until no more can be added. No analytical solution has yet been found for this problem. Our earlier simulations, done for ratios of cubic-region side length (L) to sphere radius (R) up to $L/R=20$, predicted a parking-limit volume fraction (as volume becomes very large) of 0.37–0.40, using regression equations that indicated the approach to this limit as a function of the ratio of the volume of spheres tried to the volume of the region and the ratio L/R. Our results for random parking in a volume with penetrable walls can be adjusted with a multiplicative correction factor to give the results for the same volume with impenetrable walls. An improved algorithm, almost an order of magnitude faster than our earlier one, was used to extend our simulations to $L/R=40$ and confirm the original predictions, for a series of six runs totaling 9×10^6 attempts at sphere placement. The results supported a narrower estimate of the parking-limit volume fraction: 0.385 ± 0.010.

Much recent attention has been given to the two-dimensional parking (occupation)-limit problem, determining the area fraction that would be covered at saturation by the random placement of disks on a plane, such that none overlaps.[1–4] The results are applicable to various coverage problems, including adsorption on surfaces. The three-dimensional problem for spheres, determining the fraction of volume that would be filled at saturation by the random creation of equal-size spheres, has received less attention, although it has been the subject of two investigations.[5,6] Applications include modeling condensation, modeling coagulation and clustering,[7] modeling fluid structures,[8] and the statistical problem of sampling a three-dimensional space without replacement. Tory and Jodrey[4] recently reviewed the literature on this and other packing problems, placing them into a joint context.

The problem of "parking" spheres of radius R in a cubic volume of side L can be described mathematically by two dimensionless groups: L/R and n_3, the number of spheres tried times the volume of each sphere divided by the volume of the cube. Although the one-dimensional parking problem has been solved analytically,[9,10] the problem in higher dimensions, for disks, spheres, etc., has not. It has been shown theoretically and by simulation[11,12,6] that the parked volume fraction F approaches a "parking limit" F^*, such that $(F^*-F)\propto(1/n_3)^{1/3}$. For the method of simulation used here it has been shown[6] theoretically and by simulation that the parked volume fraction approaches the parking limit for small R/L as $(R/L)^l$, for large n_3.

Combining these two relationships and determining the best-fit coefficients by linear regression on a set of simulations means, we found[6]

$$F=0.383-0.202(1/n_3)^{1/3}+1.29(R/L)\ . \tag{1}$$

The standard errors of estimate of these coefficients were[6] 0.007, 0.025, and 0.06, respectively, with coefficient of determination (r^2) of 0.99 for a set of nine values of $n_3=41.9, 83.8, 167.5$ and $L/R=10, 14.14, 20$.

We wanted to extend these simulations, which typically ended with having hundreds of spheres placed, to larger values of L/R, such as $L/R=40$, requiring thousands of spheres placed. For a given L/R ratio, the time used was proportional to the number of attempts. Six runs of 320 000 placement attempts per run (2.56×10^6 attempts in all) at $L/R=20$ took about an hour on our IBM 3081 computer system, using a compiled form of the BASIC computer language (similar to FORTRAN in running time). We expected $L/R=40$, with eight times the volume, to take almost an order of magnitude longer, which was unacceptably long. We needed to develop a faster program so that we could study $L/R=40$, and thus test the regression we had obtained from $L/R=10-20$.

With our previous algorithm[6] we chose the coordinates of each trial sphere at random within $x,y,z=0-L$. Spheres already placed were tested to see whether they were within $2R$ of the sphere being tried; if so, the trial sphere was rejected. If there were no spheres that would intersect the trial sphere, the trial sphere was added to the volume. The volume fraction was calculated as the number of spheres placed times the volume per sphere divided by the cubical volume. Periodic boundary placement was not attempted; rather, the edge effects (R/L effects) were eliminated by extrapolation to $R/L=0$. The previous algorithm was changed to the new, accelerated algorithm by having the new program divide the region into octants and search by octants. This was facilitated by making the placements on $x,y,z=-L/2$ to $+L/2$. If no intersections were found in the octant in which the trial sphere was generated, then the other octants were searched as well. The time consumed by sorting the spheres into octants was much less than the time saved in searching what was often only one-eighth the

TABLE I. Summary of results of six independent $L/R=40$ simulations and the regression-equation (1) predictions.

Number tried	Volume ratio n_3	Number placed	Number predicted by Eq. (1)	Volume fraction placed	Volume fraction predicted
640000	41.9	5452	5455	0.3568	0.3570
		5423		0.3549	
1280000	83.8	5651	5639	0.3699	0.3690
		5648		0.3697	
2560000	167.5	5809	5785	0.3802	0.3786
		5808		0.3801	

volume. This approach is a compromise between searching the entire volume each time and cutting up the volume into smaller cubes, about $2R$ on edge, then searching the 27 smaller cubes that might have sphere centers, which would be analogous to an approach used by Haile *et al.*[13] for a related problem. Haile *et al.* reported an increase in speed of a factor of 8 due to their approach, compared to searching the entire region. We achieved almost this acceleration, in the limit of large numbers of spheres tried, with less programming complexity. For example, at $L/R=20$, the attempted placement of 80000 spheres was four times as fast with the accelerated program; numerical experiments with placing fewer spheres showed that the ratio of the speed of the accelerated program to that of the original program increased as the number of attempted placements increased. As the probability of successfully placing a sphere tends to zero, the probability of that happening in the first octant tried tends to 1, making the search of other octants rare, accelerating the process by almost a factor of 8.

It is likely that even greater acceleration can be achieved by having the simulation algorithm look for "holes" in which to place spheres and then place them there randomly. This has been done in two dimensions by several investigators.[14,15,3] However, the random selection of which hole to fill and the random selection of where to place the object in that hole are difficult to achieve without subtle bias; thus, small differences between the procedure adopted by Tanemura[3] and that adopted by Lotwick[15] produced substantial differences in the two-dimensional parking limits inferred.[15] The method we used is truly random in its choice of possible placement positions.

The results of the new set of simulations, $L/R=40$, are shown in Table I. The first column shows the number of attempts at placement per run; n_3 is listed next, the ratio of volume tried to volume of region; the third column shows the number of spheres successfully placed; the fourth column shows the number predicted to be placed, based on regression equation (1); the fifth column shows the volume fraction placed; the sixth column shows the predicted volume fraction, based on regression equation (1).

Figure 1 shows (a) the regression lines for $R/L=0, \frac{1}{40}$, and $\frac{1}{20}$; (b) the previous results ($\times$) for $R/L=\frac{1}{20}$, part of the data on which the regression (1) was based; (c) our new $R/L=\frac{1}{40}$ results ($\triangle$) from Table I; (d) the results ($\triangle$) of our two previous $R/L=\frac{1}{20}$ simulations of 2.56×10^6 attempts each; (e) the path of one of our previous $R/L=\frac{1}{20}$ simulations; and (f) the path (●) of one of our new $R/L=\frac{1}{40}$ simulations of 2.56×10^6 attempts. (The paths are from runs chosen at random, not from selected runs.) Some of the symbols are masked due to overlapping. Figure 1 shows close agreement between our previous regression equation and our new simulation results. It shows F to be nearly linear with respect to $(n_3)^{-1/3}$ for $n_3<1$.

To improve our estimate of the parking-limit fraction, the results of $L/R=40$ were combined with our previous simulation results. A set of means were obtained from the simulation data for the combinations of $n_3=41.9, 83.8, 167.5$ and $L/R=10, 14.14, 20, 40$. A least-squares linear regression on the 12 means yielded

$$F=0.385-0.209(1/n_3)^{1/3}+1.29(R/L)\ . \tag{2}$$

The standard errors of estimate of these coefficients were

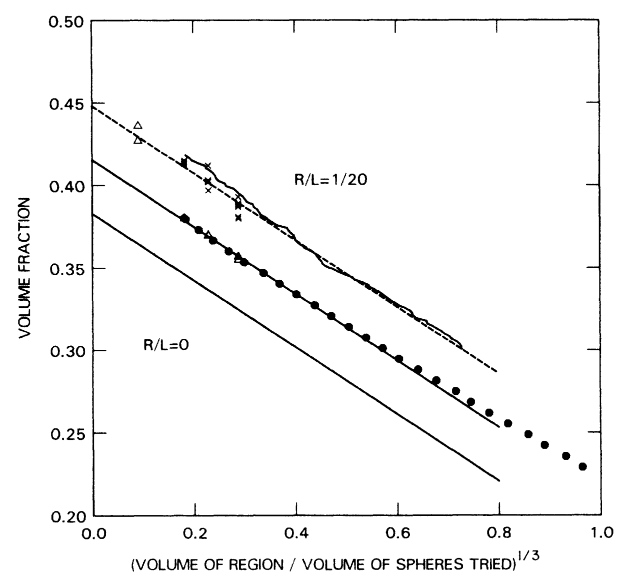

FIG. 1. Volume fraction parked (packed) vs $(1/n_3)^{1/3}$, the cube root of the ratio of the volume of the region to the total volume of all spheres tried: regression lines for $R/L=0, \frac{1}{40}$, and $\frac{1}{20}$; previous results ($\times$ and Δ) for $R/L=\frac{1}{20}$; new $R/L=\frac{1}{40}$ results ($\triangle$); path (irregular line) of a previous $R/L=\frac{1}{20}$ simulation; path (●) of a new $R/L=\frac{1}{40}$ simulation.

0.0046, 0.018, and 0.03, respectively. The coefficients are little different from those of (1) but do have smaller errors of estimate. Regression equation (2) had a coefficient of determination (r^2) of 0.9960, indicating that 99.6% of the variation in the means is included in the regression expression. The predictions from regression equation (2) were little different from those of regression equation (1); for example, at $n_3 = 167.5$ and $R/L = \frac{1}{40}$, $F = 0.3793$ was predicted by (2), and $F = 0.3786$ was predicted by (1).

The assumptions underlying linear regression include that the errors are normally distributed and that the variances do not depend on the values of the independent variables. The assumption of normally distributed errors for least-squares regression is more likely to be satisfied by regressing against the means, as we have done, because errors in the means are more normally distributed than the errors in the individual data. The assumption of equal variance for each of the independent variable values (homoscedasticity) is probably not correct, which means that although our use of unweighted least-squares regression will produce unbiased results, the coefficients will not have the minimum variance possible.[16] The small fraction of random variation $(1-r^2)$ here makes this effect small.

The probability of a successful placement equals the (marginal) change in the number placed per unit change in the number tried. The probability can be obtained by differentiating the regression equation or by counting the change in number placed per additional number tried. Using regression equation (1), the probability becomes $(0.202/3)(n_3)^{-4/3}$. At the end of the run displayed, 2.56×10^6 attempts at $R/L = \frac{1}{40}$, the last 8 of the 5808 successful placements took 87 664 attempts, thus averaging about 9 successful placements per 100 000 attempts. The regression equation (1) predicts 7 per 100 000, a close match.

For regions where the walls are impenetrable, our simulations are equivalent to having the distance between the walls be $L + 2R$. That means that the volume fraction for impenetrable walls will be $L^3/(L+2R)^3$ multiplied by that found here for penetrable walls. At $L/R = 40$, for example, the impenetrable-wall volume fraction will be 0.867 times the penetrable-wall volume fraction.

These results confirm the previous regression, Eq. (1), which gave estimates of F^* from 0.37 to 0.40. The revised regression, Eq. (2), gives a revised estimate of F^* (± two standard errors): $F^* = 0.385 \pm 0.010$. Dividing the region to be searched into octants accelerates the search up to almost a factor of 8, without adding greatly to the complexity of programming.

[1]G. Y. Onoda and E. G. Liniger, Phys. Rev. A **33**, 715 (1986).
[2]J. Feder, J. Theor. Biol. **87**, 237 (1980).
[3]M. Tanemura, Ann. Inst. Statist. Math. **31**, 351 (1979).
[4]E. M. Tory and W. S. Jodrey, in *Advances in the Mechanics and the Flow of Granular Materials,* edited by M. Shahinpoor (Trans Tech, Clausthal-Zellerfeld, West Germany, 1983), Vol. I, p. 75.
[5]K. Gotoh, W. S. Jodrey, and E. M. Tory, Powder Technol. **21**, 285 (1978).
[6]D. W. Cooper, J. Colloid Interface Sci. **119**, 442 (1987).
[7]K. Gotoh, M. Nakagawa, and H. Matsuoka, Particulate Sci. Technol. **3**, 27 (1985).
[8]B. Widom, J. Chem. Phys. **44**, 3888 (1966).
[9]A. Renyi, Publ. Math. Inst. Hung. Acad. Sci. **3**, 109 (1958).
[10]J. J. Gonzalez, P. C. Hemmer, and J. S. Hoye, Chem. Phys. **3**, 228 (1974).
[11]R. H. Swendsen, Phys. Rev. A **24**, 504 (1981).
[12]Y. Pomeau, J. Phys. A **13**, L193 (1980).
[13]J. M. Haile, C. Massobrio, and S. Torquato, J. Chem. Phys. **83**, 4075 (1985).
[14]E. M. Tory, W. S. Jodrey, and D. K. Pickard, J. Theor. Biol. **102**, 439 (1983).
[15]H. W. Lotwick, J. Statis. Comput. Simul. **15**, 295 (1982).
[16]C. Daniel and F. S. Wood, *Fitting Equations to Data,* 2nd ed. (Wiley-Interscience, New York, 1980), p. 54.