J Stat Phys (2008) 132: 207–234
DOI 10.1007/s10955-008-9539-6

On the Most Compact Regular Lattices in Large Dimensions: A Statistical Mechanical Approach

Giorgio Parisi

Received: 3 October 2007 / Accepted: 2 April 2008 / Published online: 14 May 2008

Abstract In this paper I will consider the computation of the maximum density of regular lattices in large dimensions using an approach based on statistical mechanics. The starting point will be some theorems of Rogers, which are virtually unknown in the community of physicists. Using his approach one can find many similarities (and differences) with the problem of computing the entropy of a system of hard spheres. The relation between the two problems is investigated in detail. Some conjectures are presented: further investigation is needed in order to check their consistency.

Keywords Lattice packing · Minkowski theorem · Most compact lattices · Virial expansion

1 Introduction

The aim of this paper is to study a well known and celebrated problem, i.e. that is the maximum density of hard spheres when they are packed on a regular lattice. In two and three dimensions the solution is well known; the lattices with maximal packing density are the hexagonal and the ffc lattices respectively.

In generic dimensions the result for the maximal packing density in not known [1]: a lower bound on the maximal packing density has been established by Minkowski [2] and it has only been marginally improved by later studies.

In this paper, after a brief review of the established mathematical results, we will show how the usual techniques of statistical mechanics may be used in this problem. We will not be able to find the maximal packing lattice density in high dimensions, but we believe that our work may be an useful step in this direction, especially for pointing out the connections between this problem and the statistical mechanics of disordered systems, like spin glasses and structural glasses. As we shall see later, there are very interesting relations between this problem and the thermodynamic properties of a system of hard spheres.

G. Parisi (✉)
Dipartimento di Fisica, Sezione INFN, SMC of INFM-CNR, Università di Roma "La Sapienza",
Piazzale Aldo Moro 5, 00185 Rome, Italy
e-mail: Giorgio.Parisi@roma1.infn.it

In Sect. 2 we shall introduce the basic concepts and definitions which will be used later, in particular we will introduce the Rogers measure over the space of all possible lattices. In Sect. 3 we will present the main mathematical results, due to Rogers [4–7], in particular we will show how to compute the moments of a function of the lattice using the Rogers measure. In Sect. 4 we perform some simple computations in order to become more familiar with Rogers results. In Sect. 5 we show how the whole approach can be simplified if we introduce connected moments. In Sect. 6 we present a crucial technical conjecture that is needed to make further progresses. Finally in the last section we start a comparison between the packing problem and the thermodynamic of hard spheres; we present a conjecture for the relation between these two problems in the infinite dimension limit. Further work is needed to find out if this conjecture is consistent. Finally we present our conclusions and some comments on possible developments. An Appendix is devoted to the computation of some integrals in the infinite dimensional limit.

2 Some Definitions

Let us start by defining the problem and by establishing our notations.

We must parametrize the regular D-dimensional lattices having cells of unit volume. This can be done by considering the set of unimodular real square matrices $D \times D$ (i.e. those matrices with determinant equal to 1). To each matrix Λ of this set we can associate a regular lattice, that is given by all the points of the form

$$x_i = \sum_k \Lambda_{i,k} n_k \equiv (\Lambda n)_i, \tag{1}$$

where the n_k take all the possible integer values (both positive and negative). This correspondence is not one to one: the same lattice can be obtained by infinitely many different matrices.

The condition of unimodularity (i.e. $\det(\Lambda) = 1$) implies that the volume of the fundamental cell of the lattice is equal to one.

If we consider non overlapping spheres centered on the points of the lattices, the maximum allowed *diameter* ($\mathcal{R}(\Lambda)$) is given by

$$\mathcal{R}(\Lambda)^2 = \min_n |\Lambda n|^2 = \min_n \sum_{i,k} n_i A_{i,k} n_k, \tag{2}$$

where the positive matrix A is given by

$$A_{i,k} = \sum_j \Lambda_{i,j} \Lambda_{k,j}, \tag{3}$$

and the minimum is done over all the possible choices of the integer D-dimensional vector n, with the exclusion of the origin. Indeed the quantity defined in (2) is the minimum distance of a point of the lattice from the origin. The *radius* of the largest sphere, centered in one point and not intersecting the other points, is the *diameter* of the largest non-overlapping spheres, centered on the points of the lattice.

Our aim is to compute

$$\mathcal{R}_M = \max_\Lambda \mathcal{R}(\Lambda). \tag{4}$$

We remark that the computation of the minimum in (2) for generic Λ is an NP-hard problem for large dimensions D. Of course for some matrices (e.g. diagonal ones) the computation is quite simple. On the other end, when the off-diagonal elements are large, the computation may becomes quite complex. This fact may suggest that some of the techniques used in spin glasses (where one studies the statistical mechanics of some NP-complete problems [3]) may be relevant also in this case.

In order to analyze better this problem we introduce the notation

$$\langle f(|x|)\rangle_\Lambda \equiv \sum_n f(|x(n)|) \equiv \sum_n f(|\Lambda n|), \tag{5}$$

where the sum is done over Z^D, origin excluded.

The quantity $\mathcal{R}(\Lambda)$ can computed as

$$\begin{aligned} \mathcal{R}(\Lambda)^2 &= -\lim_{\beta\to\infty} \ln(F(\beta)_\Lambda), \\ F(\beta)_\Lambda &= \frac{\langle \exp(-\beta|x|^2)\rangle_\Lambda}{\beta}. \end{aligned} \tag{6}$$

In other words $|x|^2$ is the Hamiltonian,

$$Z(\beta)_\Lambda \equiv \langle \exp(-\beta|x|^2)\rangle_\Lambda \tag{7}$$

is the partition function and $\mathcal{R}(\Lambda)^2$ is the ground state energy.

In spin glasses the typical problem would be the computation of the average value (over Λ) of the free energy associated to the partition function in (7). This is not the problem we face here. In this language the maximum packing density problem is a minimax problem, i.e. it consists in finding the matrix Λ (or equivalently A) such the ground state energy is as large as possible. A similar problem has been recently addressed for spin glasses in Ref. [8].

We can formulate the same problem in an alternative way if we consider the quantity

$$K(R)_\Lambda = \langle \theta(R - |x|)\rangle_\Lambda, \tag{8}$$

i.e. the number of points, excluded the origin, inside a sphere of radius R. It is obvious that

$$\begin{aligned} K(R)_\Lambda &= 0 \quad \text{for } R < \mathcal{R}(\Lambda), \\ K(R)_\Lambda &> 0 \quad \text{for } R > \mathcal{R}(\Lambda). \end{aligned} \tag{9}$$

If at given Λ we were able to compute $\langle\cdot\rangle_\Lambda$, we would immediately obtain the wanted result after a maximisation over Λ of the quantity $K(R)_\Lambda$. We have already remarked that the computation of these expectation values for a given lattice is a nasty problem that cannot simply solved analytically. Often in random problems we introduce an ensemble of problems and we try to compute the average over this ensemble. Usually this goal can be reached analytically.

In order to formulate the statistical problem we must firstly introduce a measure $d\mu(\Lambda)$ on the space of all possible lattices; we will specify later the form of this measure. We define

$$\overline{F(\Lambda)} = \int d\mu(\Lambda) F(\Lambda). \tag{10}$$

If the measure $d\mu(\Lambda)$ does not vanish near the maximum of $\mathcal{R}(\Lambda)$, we can extract the value of $\mathcal{R}_M$ from the properties of the appropriate averages. We have two different possibilities to perform this last step:

- We can use the formula

$$\mathcal{R}_M^2 = -\lim_{\beta\to\infty}\lim_{n\to-\infty}\frac{\ln\left(\overline{Z(\beta)^n_\Lambda}\right)}{n\beta}. \tag{11}$$

  We need to compute the average partition function of n replicas of the model in the slightly unusual limit $n \to -\infty$. Although there are techniques to perform this computation [8], in this paper we will try to use an alternative approach more similar to the standard replica method that it is known to works very well in the case $n = 0$.
- We can alternatively compute the moments of $K(R)_\Lambda$, this quantity being defined in (8):

$$K^{(s)} = \overline{K(R)^s_\Lambda} \equiv \sum_{k=0,\infty} P(k, R)k^s. \tag{12}$$

  The function $P(k, R)$ is the probability that a lattice (randomly chosen with the measure $d\mu(\Lambda)$) has k points inside a sphere of radius R. As far as the origin ($n = 0$) is excluded, the function $P(k, R)$ is different from zero only for even k (if a point of the lattice is inside a sphere also the opposite point is inside the same sphere). When $R > \mathcal{R}_M$ there are no lattices with no points inside a sphere of radius R. It is evident that

$$\begin{aligned} P(0, R) &> 0 \quad \text{for } R < \mathcal{R}_M, \\ P(0, R) &= 0 \quad \text{for } R > \mathcal{R}_M. \end{aligned} \tag{13}$$

  The program consists in reconstructing the function $P(k, R)$ from its moments: in this way one can find the value of $\mathcal{R}_M$. In this paper we will explore this second approach.

Before going on with these computations, we must chose the measure $d\mu(\Lambda)$. Three possibilities come to our mind. All of them are reasonable, however, depending on the technique we use, the computation of the average over Λ may be easier.

- There is a natural definition (due to Siegel [10]) of the measure over unimodular matrices, restricted to the fundamental region, the fundamental region being defined in such a way that each lattice may be represented in one and only one way by a matrix in the fundamental region. Computations with this measure are technically rather difficult and they will not considered here.
- We can consider the set of matrices Λ (introduced by Rogers) that at fixed ω depend on $D - 1$ α-variables. The action of such a matrix on a vector is defined as

$$\begin{aligned} (\Lambda n)_i &= \omega n_i \quad \text{for } i < D, \\ (\Lambda n)_D &= \eta\left(n_D + \sum_{i=1,D-1}\alpha_i n_i\right), \end{aligned} \tag{14}$$

  where the unimodularity condition implies that

$$\omega^{D-1}\eta = 1. \tag{15}$$

The ω-dependent measure is obtained by taking a flat measure in the interval 0–1 for each of the $D-1$ α-variables. Eventually the limit $\omega \to 0$ is taken. It is already non-trivial to prove that each lattice may be represented under the previous form. The proofs of this and other difficult points can be found in the paper of Rogers [4–7] and they will not reproduced here.

The computations with the Rogers measure are much simpler (many results are known). Luckily enough the two measures (Rogers and Siegel) are equivalent for our purposes: the quantities

$$\overline{\langle f \rangle^s_\Lambda}, \tag{16}$$

that will play a crucial role in our study, are the same if evaluated with the Rogers measure or with the Siegel measure.[1]

- As suggested by Kurchan and Mézard [9] we can consider the Gaussian measure:

$$d\mu(\Lambda) \propto \prod_{i,k} d\Lambda_{i,k} \exp(-\gamma \mathrm{Tr}(\Lambda\Lambda^*))\delta(\det(\Lambda) - 1), \tag{17}$$

where γ is an arbitrary parameter. In this case the some of the computations can be done with the same techniques used in many physical problems, e.g. spin glasses [3]. The same measure may be useful in numerical simulations.

As we shall see later, a simple computation shows that in the case of the Rogers measure we have that

$$\overline{\langle f \rangle} = \int d^D x f(x). \tag{18}$$

If we apply this result to the case $f(x) = \theta(R - |x|)$ we find that

$$\overline{K(R)} \equiv \sum_{k=0,\infty} P(k, R)k = V_D(R), \tag{19}$$

where $V_D(R) \equiv R^D \pi^{D/2} \Gamma(D/2+1)^{-1}$ is the volume of the D-dimensional sphere of radius R, that obviously depends also on D. The function $K(R)_\Lambda$ may take only even integer values and consequently $P(0, R)$ must be different from zero in the region where

$$V_D(R) < 2, \tag{20}$$

that is also a consequence of the celebrated Minkowski theorem in large dimensions.[2]

The proof of (20) is rather simple. Indeed using (18) we have that:

$$V_D(R) = \overline{K(R)} = \sum_{k=2,\infty} P(k, R)k \geq 2 \sum_{k=2,\infty} P(k, R) = 2(1 - P(0)). \tag{21}$$

Let us call R_c the D-dependent value of R such that

$$V_D(R_c) = 1, \tag{22}$$

[1] In the case of the Rogers measure the overline denotes the integral over all the α_i and the limit ω to zero.

[2] A more precise consequence of the original Minkowski theorem is that $P(0, R)$ is different from zero when $\zeta(D)V_D(R) < 2$ where $\zeta(D)$ is the Riemann zeta function ($\zeta(\infty) = 1$). As a consequence we have that $\mathcal{R}_M \geq R^*_D$ where $\zeta(D)V_D(R^*_D) = 2$.

its value being given by the condition

$$\frac{R_c^D \pi^{D/2}}{\Gamma(D/2+1)} = 1. \tag{23}$$

It is convenient to measure all the lengths in units of R_c when the dimension D go to infinity. At this end we define:

$$r = R/R_c, \quad r_M = \mathcal{R}_M/R_c. \tag{24}$$

When the dimensions D go to infinity, we face the problem of finding the limit of r_M that we suppose to exist. The following bounds are known for infinite D:

$$1 \le r_M \le 1.322. \tag{25}$$

The lower bound is the Minkowski theorem, while the upper bound (the Kabatiansky-Levenshtein bound) comes from a totally different approach that we cannot discuss in details for reasons of space [11, 12]. A simpler upper bound has derived by Levenshtein [13]:

$$r_M \le \frac{e}{2} \approx 1.359. \tag{26}$$

This bound is weaker of the previous, but the proof is more direct. It is still simpler to prove the Rogers bound [4–7]:

$$r_M \le \sqrt{2} \approx 1.414. \tag{27}$$

3 Some Known Results

3.1 Rogers Main Theorem

The aim of this section is to recall some known results of the values of the moments of the function f,

$$\overline{f^s} \equiv \overline{\left(\sum_n f(\Lambda n)\right)^s}, \tag{28}$$

using the Rogers measure (14) in the case where $s < D$. We shall see later how this annoying constraint ($s < D$) may be removed. The reader should note that here (and in the following) the vector signs are not indicated: both n and Λn are D dimensional vectors and Λ is a $D \times D$ dimensional matrix. For simplicity we will also assume that the function f is even.

From (18) we have that

$$\overline{f} \equiv \overline{\sum_n f(\Lambda n)} = \int_{-\infty}^{\infty} d^D x f(x). \tag{29}$$

We now segue into the evaluation of the higher moments of f. Fortunately the appropriate computations have been done in rigorously way by Rogers [4–7]. His results are relatively simpler in position space, although they can be formulated in momentum space.

The strategy for computing the moments defined in (28) consists of a few steps:

- We classify all sets of s vectors n_k ($k = 1, \ldots s$) according to their linear dependence.

- We perform the average and the sums inside each class.
- We write the final result the sum over all possible classes.

The crucial theorem is based on the following two lemmas.

- One can prove that

$$\overline{\prod_{k=1,s}\left(\sum_{n^k}^{LI} f(\Lambda n^k)\right)} = \left(\int d^D x f(x)\right)^s, \tag{30}$$

where $\sum^{LI}$ denotes the sum over all the sets of s vectors n^k belonging to Z^D that are linearly independent (as usual the origin, i.e. $n^k = 0$, never appears). The formula is valid only for $s \leq D$[3] and it has a very simple meaning. By changing the lattice each of the s points n^k may be carried in any point of the space independently from the other one, provided that they are linearly independent.

- In order to obtain the final result need also to consider the average of similar sums restricted to the case where the vectors n are linearly *dependent*. More precisely we introduce s vectors n that are linear combination of $s-h$ linear independent vectors and their linear dependence is specified by a matrix M. In other words the s vectors n span a $(s-h)D$ dimensional space and they satisfy the following h (vectorial) linear conditions:

$$\sum_{k=1,s} M_{j,k} n^k = 0 \quad \text{for } j = 1, h, \tag{31}$$

where the matrix M has integer elements and it is irreducible (there is no integer matrix M' such that $pM' = M$ with p integer). The previous case corresponds to $h = 0$.

In this case we find that

$$\overline{\prod_{k=1,s}\left(\sum_{n_i^k}^{(M)} f(\Lambda n^k)\right)} = \mathcal{N}(M) \int \prod_{k=1,s} d^D x^k f(x^k) \prod_{j=1,h} \delta^D\left(\sum_{k=1,s} M_{j,k} x^k\right), \tag{32}$$

where the sum ($\sum^{(M)}$) is done on vectors that satisfy the condition (31) (they depend on the matrix M and have $s-h$ linear independent components). The quantity $\mathcal{N}(M)$ is a normalization factor that is equal to 1 in many cases; for our purposes it may be taken equal to 1.

We now have all the elements to implement the strategy to compute the moments. We consider all possible linear dependence of the vectors n and we transform the sum over all the values of n to a sum over all possible linear dependencies. We can now state the main theorem of Rogers [4–7]:

$$\overline{\langle f \rangle^s} = \sum_{h=0,s-1} \sum_{M} \mathcal{N}(M) \int \prod_{k=1,s} d^D x^k f(x^k) \prod_{j=1,h} \delta^D\left(\sum_{k=1,s} M_{j,k} x^k\right), \tag{33}$$

where the sum is done over all the sets of $h \times s$ matrices M corresponding to different linear conditions.[4]

[3] It is obvious that we cannot find more than D linearly independent vectors in dimensions D.

[4] In order to verify if two matrices M correspond to different linear conditions, Rogers writes the matrices M in a *canonical* form and he presents his main theorem in this form. However we do not need his theorem under this more refined form.

3.2 The Case $s = 2$

Let us show how this strategy works in the case $s = 2$. Here we have to compute

$$\overline{\sum_{n_1,n_2} f(n_1) f(n_2)}. \tag{34}$$

Now we have two possibilities:

- The two vectors n are linearly independent. We obtain the following contribution

$$\int d^D x_1 d^D x_2 f(x_1) f(x_2). \tag{35}$$

- We consider the case where the two vectors n are linearly dependent. In this case we can write the constraint in an unique way as

$$q_1 n^1 + q_2 n^2 = 0, \tag{36}$$

  if we restrict ourselves to the case of positive q_1 and $(q_1, q_2) = 1$ (i.e. the pair q_1 and q_2 is irreducible).[5] We finally finds the following contribution

$$\int d^D x d^D y f(x) f(y) \delta^D(q_1 x + q_2 y) = \int d^D x f(q_1 x) f(q_2 x). \tag{37}$$

Putting everything together we find the final expression

$$\overline{\sum_{n_1,n_2} f(n_1) f(n_2)} = \int d^D x_1 d^D x_2 f(x_1) f(x_2) + \sum_{q_1,q_2}^{I} \int d^D x f(q_1 x) f(q_2 x), \tag{38}$$

where the sum $\sum^I$ is restricted over the irreducible pairs with q_1 positive.

We have already seen how this construction works for $s = 2$, in the next subsection we will consider in details the case $s = 3$.

3.3 The Case $s = 3$

If we apply Rogers formula in this case we have only three possibilities: $h = 0$, $h = 1$ and $h = 2$ (h being the number of linear constraints).

- $h = 0$. In this case the matrix M does not exist and the corresponding contribution is simply given by

$$\int d^D x_1 d^D x_2 d^D x_3 f(x_1) f(x_2) f(x_3). \tag{39}$$

- $h = 1$. In this case we have a linear constraint of the form

$$q_1 x_1 + q_2 x_2 + q_3 x_3 = 0. \tag{40}$$

[5] The expression (a, b) denotes as usually the maximum common divisor of a and b.

We have two possibilities.

– One of the three q is equal to zero. Here we have three equal contribution. If we set $q_3 = 0$, we must impose that q_1 and q_2 have no common factors (i.e. $(q_1, q_2) = 1$). We finally find a contribution equal to

$$3 \sum_{q_1,q_2}^{I} \int d^D x_1 f(q_1 x_1) f(q_2 x_1) \int d^D x_3 f(x_3). \tag{41}$$

– All the q are non zero. The previous condition requires that q_1, q_2 and q_3 have no common factors. We finally find

$$\sum_{q_1,q_2,q_3}^{I} |q_3|^{-D} \int d^D x_1 f(x_1) \int d^D x_2 f(x_2) f\left(\frac{q_1 x_1 + q_2 x_2}{q_3}\right), \tag{42}$$

where the expression $\sum_{q_1,q_2,q_3}^{I}$ indicates the we sum over the values of q_1, q_2 and q_3 that have no common factors.

- $h = 2$. In this case we can write the two linear constraints as

$$q_1 x_1 = q_2 x_2 \quad \text{and} \quad g_1 x_1 = g_2 x_3. \tag{43}$$

We finally finds that the contribution from this case is given by

$$\sum_{q_1,q_2,g_1,g_2}^{I} |q_2 g_2|^{-D} \int d^D x f(x) f\left(\frac{q_1 x}{q_2}\right) f\left(\frac{g_1 x}{g_2}\right). \tag{44}$$

The final result is given by the sum of all the previous contributions.

In the same way we can write explicit formulae that contain more and more terms when s increases.

4 Some Simple Computations

In this section we present some simple computations, in order to familiarizes ourselves with the previous results.

If we look to the previous formulae it is evident that there is are terms that go to zero very fast when $D \to \infty$ for many choices of the function f. Let us see what happens in some simple examples; we will also consider some simple approximations.

4.1 The Gaussian Case

Let us consider the case where the function $f(x)$ is given by

$$f(x) = (2\pi)^{-D/2} \exp(-\beta x^2/2). \tag{45}$$

Here the integrals can be easily done and we finds that

$$
\begin{aligned}
&\overline{\langle f\rangle} = \beta^{-D/2}\\
&\overline{\langle f^2\rangle} = \beta^{-D} + \beta^{-D/2}\sum_{q_1,q_2}^{I}(q_1^2+q_2^2)^{-D/2},\\
&\overline{\langle f^3\rangle} = \beta^{-3D/2} + 3\beta^{-D}\sum_{q_1,q_2}^{I}(q_1^2+q_2^2)^{-D/2}\\
&\qquad + \beta^{-D}\sum_{q_1,q_2,q_3}^{I}(q_1^2+q_2^2+q_3^2)^{D/2} + \beta^{-D/2}\sum_{q_1,q_2,g_1,g_2}^{I}(q_1^2g_1^2+q_1^2g_2^2+q_2^2g_1^2)^{-D}.
\end{aligned}
\tag{46}
$$

For large D we find:

$$
\begin{aligned}
&(\overline{\langle f\rangle})^{-2/D} = \beta^{-1},\\
&(\overline{\langle <f\rangle^2})^{-2/D} \approx \max(\beta^{-2},(2\beta)^{-1}),\\
&(\overline{\langle f\rangle^3})^{-2/D} \approx \max(\beta^{-3},(2\beta^2)^{-1},(9\beta)^{-1}).
\end{aligned}
\tag{47}
$$

The problems connected with the limit $D\to\infty$ are clear from the previous formulae. Depending on the value of β different terms are the leading ones (the case $\beta<1$, $1<\beta<2$, $2<\beta<9/2$ and $\beta>9/2$ are different). Sometimes it happens that the leading term from one contribution is smaller of the subleading terms of an other contribution. The terms with the q's and the g's equal to 1 are always the leading ones inside a given class.

4.2 The Theta Function

We consider here the case of the function:

$$
f(x) = \theta(R-|x|). \tag{48}
$$

Our plan is to compute the leading term of each moment of f when the dimension goes to infinity, resuming the resulting series and extracting information on the probability $P(k,D)$. The relation $f^2(x)=f(x)$ will be useful in simplifying the result.

Let us use the previous formulae for computing the first three moments. We find in the same way as before that

$$
\begin{aligned}
&K^{(1)} = \overline{\langle f\rangle} = V_D(R),\\
&K^{(2)} = \overline{\langle f\rangle^2} = V_D(R)^2 + V_D(R)\sum_{q_1,q_2}^{I}(\max(q_1,|q_2|)^{-D},\\
&K^{(3)} = \overline{\langle f\rangle^3}\\
&\quad = V_D(R)^3 + 3V_D(R)^2\sum_{q_1,q_2}^{I}\max(q_1,|q_2|)^{-D}\\
&\qquad + V_D(R)\sum_{q_1,q_2,g_1,g_2}^{I}\left(\frac{\max(1,|q_2/q_1|,|g_2/g_1|}{q_2g_2}\right)^{-D}\\
&\qquad + \sum_{q_1,q_2,q_3}^{I}(q_3)^{-D}\int d^Dx_1\int d^Dx_2(\theta(R-|x_1|)\theta(R-|x_2|)\theta(q_3R-|q_1x_1+q_2x_2|).
\end{aligned}
\tag{49}
$$

We have terms that scale as different powers of $V_D(R)$. Inside each class the leading terms are those the variable q and g are equal to ± 1. If we keep only these terms we find the simple result

$$\begin{aligned} K^{(1)} &= \overline{\langle f \rangle} = V_D(R), \\ K^{(2)} &= \overline{\langle f^2 \rangle} = V_D(R)^2 + 2V_D(R), \\ K^{(3)} &= \overline{\langle f^3 \rangle} \\ &= V_D(R)^3 + 6V_D(R)^2 + 4V_D(R) \\ &\quad + 4 \int d^D x_1 \int d^D x_2 \theta(R - |x_1|)\theta(R - |x_2|)\theta(R - |x_1 + x_2|). \end{aligned} \tag{50}$$

The last integral is exponentially small for large D with respect to the other terms (see also the Appendix): it is the probability of having three overlapping spheres of the same radius. This crucial results follow from two general facts[6] that are valid when $D \to \infty$:

- The measure on a D-dimensional sphere is concentrated on its surface.
- Two generic vectors (of length 1) have vanishing scalar product; therefore the length the sum of two generic vectors of length less or equal to R is $2^{1/2}R$ with probability 1 when D goes to infinity.

Also in this case (depending if R is larger or smaller that R_c) different terms dominate the result for each moment when the dimension becomes large.

4.3 Summing the Leading Terms

Our aim it to obtain a closed formula for the leading terms for each moment of the function f in order to extract the asymptotic behaviour. Unfortunately controlling the leading term of each moment does not imply that we control the leading term of the probability distribution of the values of the function f in an uniform way.

Generalizing the analysis of the first three moments we conclude that at given value of h the leading terms, when the dimensions go to infinity, come from h linear conditions of the form

$$\sum q_i x_i = 0 \tag{51}$$

with $q_i \in \pm 1$.

If we select these contributions the result can be given in terms of h integrals of the same function. For $s > 0$ in the case of an even function (i.e. $f(-x) = f(x)$) we get:

$$\overline{\langle f \rangle^s} \approx \sum_{h=1,s} \prod_{i=1,h} \left(\sum_{\nu_i = 1,s} C(s,h,\nu) 2^{\nu_i - 1} \int d^D x f(x)^{\nu_i} \right), \tag{52}$$

where the sum is done over all the sets of h integers ν_i such that

$$\sum_i \nu_i = s, \tag{53}$$

[6]The argument is well known and plays a crucial role in the evaluation of the leading term of the Mayer expansion for an hard sphere gas in infinite dimensions (see [15–18]).

and $C(s,h,\nu)$ is a crucial combinatorial factor that is equal $h!^{-1}$ times the number of ways in which we can divide s objects in h group of ν_i elements:

$$C(s,h,\nu) = \frac{s!}{h!\prod_{i=1,h}\nu_i!}. \tag{54}$$

These formulae simplify if we compute the generating function of the moments. After some simple algebra we find that

$$\overline{\exp(y\langle f\rangle)} \approx \exp\left(\frac{1}{2}\int dx(\exp(2yf(x))-1)\right). \tag{55}$$

If we apply this formula to the case of the theta function, defined in (48), we find that

$$\begin{aligned}\overline{\exp(yK(R)_\Lambda)} &\approx \exp\left(\frac{V_D(R)}{2}(\exp(2y)-1)\right)\\ &= \exp\left(-\frac{V_D(R)}{2}\right)\sum_{k=0,\infty}\frac{1}{k!}\left(\frac{V_D(R)}{2}\right)^k\exp(2ky).\end{aligned} \tag{56}$$

At the end we get a simple result:

$$P(2k,R) \approx \exp\left(\frac{-V_D(R)}{2}\right)\frac{1}{k!}\left(\frac{V_D(R)}{2}\right)^k, \tag{57}$$

while $P(2k+1,R)=0$ (The number of points of a lattice inside a sphere is always even!).

This Poisson distribution implies that

$$P(0,R) \approx \exp\left(\frac{-V_D(R)}{2}\right). \tag{58}$$

Let us find the consequences of this approximate result. We have called R_c the radius of a sphere of unit volume. If we take a large value of D at fixed ratio $r \equiv (R/R_c)$ we find that $V(rR_c)$ goes to zero or to infinity depending if r is smaller or greater than 1.

- If $r<1$, $P(0,rR_c)$ goes to 1 apart from exponentially small corrections. With probability going to 1 lattices do not have points at distance smaller that R_c.
- In the interesting case $r>1$ this computation give an exponentially small, but non-vanishing, contribution for $P(0,rR_c)$.

If we stick to this result, $P(0,rR_c)$ is always non zero and that we can find always a lattice that contains no points at distance less than rR_c. This result cannot be correct because it can be proved that no such lattice exist for large D as soon $r>1.322$ [11, 12].

This failure arise from neglecting exponentially small terms. They can be neglected only if these terms do not become larger of the exponential *small* expression we have obtained for the probability distribution by summing exponentially *large* terms. In other words we cannot neglect the subleading terms if there are strong cancellations among the leading terms and the sum of the leading terms becomes smaller of some subleading term. It was proved by Rogers that these cancellations do not happen in the relatively small region where $r^D < D$ (a similar region appears in the analysis of [20–23]) and this leads to a marginal improvement of the Minkowski bound.

Up to now we have transcribed the results of Rogers in a slightly different language skipping the proof of his theorems. Our aim is to extend our command of the expression for $P(k, R)$ by considering more terms and to control better the integrals the correspond to the neglected terms. This will be done in the next sections.

4.4 The Appearance of the Random Energy Model

Before going to more precise computations, it is convenient to study also the partition function retaining only the terms considered in the previous subsection. We recall the definition of the partition function at fixed Λ, that is given in (7):

$$Z(\beta)_\Lambda \equiv \langle \exp(-\beta |x|^2) \rangle_\Lambda . \tag{59}$$

If we apply the formulae of the previous subsections we find after some algebra that

$$\overline{Z(\beta)^s} = 2^s \sum_{h=1,s} \prod_{i=1,h} \sum_{n_i=1,s} C(s,h,n) 2^{-1} \left(\frac{2\pi}{\beta n_i} \right)^{-D/2}, \tag{60}$$

where (as before) the sum is done over all the sets composed by h integers n_i such that $\sum_i n_i = s$.

The previous formula does not look particularly illuminating, however a close look shows that this formula is well known to people studying the random energy model (REM) [14]. Using this knowledge we can rewrite it in a more illuminating way. At this end let us consider a new model, where the partition function can be written as

$$\mathcal{Z}(\beta) = \sum_{k=1,N} \exp(-\beta E_k), \tag{61}$$

where the *energies* E_k are N random independent quantities distributed with the probability distribution $P^{(N)}(E)$. Standard arguments shows that

$$\overline{\mathcal{Z}(\beta)^s} = \sum_{m=1,s} \sum_{n_i=1,s} C(s,m,n) \prod_{i=1,m} p_{n_i}^{(N)}, \tag{62}$$

where

$$p_n^{(N)} = N \int dE\, P^{(N)}(E) \exp(-n\beta E) = \int dE p^{(N)}(E) \exp(-n\beta E). \tag{63}$$

We have also defined

$$p^{(N)}(E) = N P^{(N)}(E). \tag{64}$$

In other words $p^{(N)}(E)dE$ is the probability of finding an energy in the interval $[E, E+dE]$.

The limit N going to infinity can be done without difficulties if the function $p^{(N)}(E)$ depends on N in such a way that this function remains finite in this limit, i.e. if the following limit exists:

$$p(E) \equiv \lim_{N\to\infty} p^{(N)}(E). \tag{65}$$

The normalization condition ($\int p^{(N)}(E)dE = N$) implies that

$$\int p(E) = \infty. \tag{66}$$

In order to identify $Z(\beta)$, defined in (60), with $\mathcal{Z}(\beta)$, the function $p(E)$, defined in (65) must be given by

$$p(E) = \frac{D}{4} E^{(D-1)/2} V_D. \tag{67}$$

Indeed, if we neglect the factor 2^s in (60),[7] the moments of the partition function defined in (60) coincide with those of a random energy model where

$$p_n \equiv \lim_{N\to\infty} p_n^{(N)} = 2^{n-1} \left(\frac{2\pi}{\beta n_i}\right)^{-D/2} = \frac{1}{2}\int d^D x \exp(-n\beta x^2/2). \tag{68}$$

Therefore we have that

$$p(E) = \frac{1}{2}\int d^D x \delta(E - x^2/2). \tag{69}$$

In other words we can consider N points x_k in the D-dimensional space that are uniformly randomly distributed with density $1/2$ inside a sphere with as radius going to infinity with N. Within the previous approximation the partition function is the limit $N \to \infty$ of

$$Z(\beta) = \sum_{k=1,N} \exp\left(-\beta x_k^2/2\right). \tag{70}$$

This implies that the previous approximation corresponds to a flat uncorrelated distribution of the points Λn. The probability of finding an arbitrary large sphere of the space with no points, is always non zero, although it is very small for a quite large sphere. This conclusion is perfectly consistent with the results of the previous subsection, where it was shown that there is no upper limit for the radius.

5 The Importance of Being Connected

5.1 The Introduction of Connected Moments

Before considering further contributions we need simplify the previous computations. At this end it is important to introduce the connected moments and to present some general consideration.

Let us consider an integer valued function f and its moments defined as

$$\langle f^s \rangle = \sum_k P(k) k^s \equiv f_s, \tag{71}$$

where $P(k)$ is the probability that the function f is equal to k.

It is usual to introduce the generating function

$$G(z) = \sum_{s=0,\infty} f_s \frac{z^s}{s!} = \sum_{s=0,\infty}\sum_{k=0,\infty} P(k)\frac{(zk)^s}{s!} = \sum_{k=0,\infty} P(k)\exp(zk) = \exp(C(z)), \tag{72}$$

[7] The factor 2^s, that is irrelevant for most conclusions may be easily obtained by adding a fixed degeneracy (i.e. 2) of the levels.

where $C(z)$ is the generating function of the connected moments:

$$C(z) = \sum_{s=1,\infty} \langle f^s \rangle_c \frac{z^s}{s!}. \tag{73}$$

Let us assume that we can write

$$C(z) = B \sum_{k=1,\infty} p(k)(\exp(zk) - 1), \tag{74}$$

where the quantities $p(k)$ are non-negative and the quantity B is fixed by the condition that

$$\sum_{k=1,\infty} p(k) = 1. \tag{75}$$

Using the theorem of composed probability we readily find that

$$P(k) = \exp(-B) \sum_{n=0,\infty} \left(\frac{B^n}{n!} \prod_{i=1,n} \left(\sum_{k_i=1,\infty} p(k_i) \right) \delta \left(\sum_{i=1,n} k_i - k \right) \right). \tag{76}$$

In other words the distribution $P(k)$ can be obtained using the following procedure:

- We first extract a number n with a Poisson distribution with average B.
- We extract n independent numbers k_i with probability $p(k)$.
- The quantity k is given by $\sum_{i=1,n} k_i$.

We finally find that

$$P(0) = \exp(-B) = \exp\left(\lim_{z \to -\infty} C(z) \right). \tag{77}$$

The problem of computing the value of $P(0)$ is reduced to the evaluation of the generating function of the connected moments in a particular limit.

On the other hand, let us assume that there exist an analytic function $f_c(s)$ such that for positive integer s

$$f_c(s) = \langle f^s \rangle_c, \tag{78}$$

and the function $f_c(s)$ does not have a nasty behavior at infinity in the complex plane. If this function does not have singularities in the region $\Re e\, s \geq 0$, by substituting (78) in (73), one can prove that

$$\lim_{z \to -\infty} C(z) = -f_c(0). \tag{79}$$

This can be done using the representation

$$C(z) = (2\pi i)^{-1} \int_{-i\infty}^{i\infty} \Gamma(-s) f_c(s) (-z)^s, \tag{80}$$

where the integration path crosses once the real line at a point in the interval 0–1. For large negative z the integration path may be deformed to the left and the leading contribution comes from the pole at $s = 0$.

The task of computing $P(0)$ is reduced to the problem of finding an analytic expression for the function $f_c(s)$ and of computing it at $s = 0$. Our strategy will consists in summing subsets of all the possible contributions and to evaluate the result at $s = 0$.[8]

If we use the computation of the moments done in the previous section, where only the leading terms have been taken into account, we find a very simple result for the connected moments:

$$f_c(s) = \int d^D x f(x)^s. \tag{81}$$

If we take the function f equal to $\theta(R - |x|)$, we find that

$$f_c(s) = \int d^D x \theta(R - |x|)^s = V_D, \tag{82}$$

that is obviously an analytic function of s in the whole complex plane.

5.2 How to Compute the Connected Moments

It is evident from the previous discussion that the direct computation of the connected moments is extremely interesting. The final expression for the connected moments certainly contains less terms than the expression for the moments: this is not a surprise to people working in statistical mechanics.

We discuss now how we may compute the other terms in the connected moments. Let us consider what happens up to the third moment. If we start from the formulae of the Sect. 3, using the same notation, we find:

$$\langle f \rangle_c = \int d^D x f(x),$$

$$\langle f^2 \rangle_c = \sum_{q_1,q_2}^{I} \int d^D x_1 d^D x_2 f(x_1) f(x_2) \delta(x_1 q_1 + x_2 q_2), \tag{83}$$

$$\langle f^3 \rangle_c = \sum_{q_1,q_2,q_3}^{I} \int d^D x_1 d^D x_2 d^D x_3 f(x_1) f(x_2) f(x_3) \delta(q_1 x_1 + q_2 x_2 + q_3 x_3) \tag{84}$$

$$+ \sum_{q_1,q_2,g_1,g_2}^{I} \int d^D x_1 d^D x_2 d^D x_3 f(x_1) f(x_2) f(x_3) \delta(g_1 x_1 + g_2 x_2) \delta(q_1 x_1 + q_2 x_3). \tag{85}$$

The only terms that remains in this limit are connected (in the same way as connected diagrams): they cannot be split into products of terms that contain integrals that can be done independently.

In order to derive a similar result for the higher moments and to stream down the discussion let us set all factor $\mathcal{N}(M)$ to 1. In this case the general formula of Sect. 3 can be written as

$$\langle f^s \rangle = \sum_G \int \prod_{i=1,n} (d^D x_i f(x_i)) G(\{x\}), \tag{86}$$

[8] The reader may be puzzled by this result that has a strong replica flavour.

where the sum is done over the appropriate set of function $G(\{x\})$. These functions are an appropriate product of delta functions that enforce the linear dependence of some of the x's.

A function G is connected if it goes to zero when one or more of the x go to infinity together. In other words we must have that

$$\lim_{\lambda\to\infty} G(x_1+\lambda c_1, x_2+\lambda c_2, \ldots, x_n+\lambda c_n) = 0 \tag{87}$$

for any set of vectors c where at least one c is different from zero and all c's are not equal (this is essentially the clustering property of connected correlation functions in a pure state in statistical mechanics).

If G is not connected, it can be written as the product of its connected pieces. If we take care of all the multiplicity factors and we use the standard manipulations of statistical mechanics we find that

$$\langle f^s\rangle_c = \sum_{G_c}\int \prod_{i=1,n}\left(d^D x_i\, f(x_i)\right) G_c(\{x\}), \tag{88}$$

where the sum is done only over those G that are connected (we call them $G_c(\{x\})$).[9]

6 A Crucial Conjecture

6.1 A Serious Difficulty

Before going further we must face a delicate problem that we have postponed up to now.

The program we have put forward may seem witless. We have seen that using Rogers theorems in dimensions D we can write (33) of the form

$$\overline{f^s} = F(s, D), \tag{89}$$

that is valid only for $s < D$ (as we have seen the r.h.s of the previous equation contains some sums over integers and some integrals).

Rogers theorems (that are at the basis of our computations) do not give useful information for $s \geq D$ for the very good reason that in this case both sides of equation (33) are infinite. Indeed there are very asymmetric lattices where the function $\langle f\rangle_\Lambda$ is very large (e.g. there are many points very near to the origin). Conversely the sum over the matrices M appearing in l.h.s. of (33) is not always convergent. Indeed it is easy to check that the formula (38) for $\overline{f^2}$ is divergent in two dimensions as soon as $f(0) \neq 0$.

How could we write a meaningfully expression in finite dimension for $\overline{\exp(y\langle f\rangle)}$, if we can compute only a few moments of the $\langle f\rangle$? We would like to put forward a reasonable conjecture that should allow us to overcome this difficulty. Before doing it, let us look better to the previous formulae for the connected moments and see what happens if we refine the previous computations by including some additional terms.

[9] Some people may consider the previous formula self-evident. Indeed let us consider a function f that is the characteristic function of a domain of volume V. General considerations may be used to argue that $\langle f^s\rangle_c$ should be proportional to V and this is possible only if only connected G are present in the r.h.s of the previous equation.

6.2 A More Refined Computation

Let us consider for simplicity only the connected moments. The non-connected one can be obtained from the connected ones. We want now to analyze in details the origin of the divergences that we have seen in the previous subsection.

Let us start by considering the second moment in the Gaussian case where the function $f(x)$ is equal to $\exp(-x^2)$. In this case we have explicit formulae. Doing the appropriate integrals one obtain that

$$\langle f^2\rangle_c = \sum_{q_1,q_2}^{I} \left(\frac{\pi}{q_1^2+q_2^2}\right)^{D/2}. \tag{90}$$

As we will see in dimensions larger than two the sums are convergent.

In the same we obtain

$$\langle f^3\rangle_c = \sum_{q_1,q_2,q_3}^{I} \left(\frac{\pi^2}{q_1^2+q_2^2+q_3^2}\right)^{D/2} + \sum_{q_1,q_2,g_1,g_2}^{I} \left(\frac{\pi}{q_1^2g_1^2+q_2^2g_1^2+q_1^2g_2^2}\right)^{D/2}. \tag{91}$$

In dimensions three only the first term is divergent. The other term is divergent in dimensions two.

This is a general feature. The most divergent term is always the one that correspond to a single linear constraint. We can therefore start by considering only this contribution. We want to arrive to a final simple expression, that can be reached in a few steps.

Within the approximation of taking the linear constraint we have

$$\overline{\langle f\rangle^s}_c \approx \sum_{q_1\cdots q_s}^{I} \int \prod_{k=1,s} dx_k^D f(x_k)\delta\left(\sum_{i=1,s} q_i x_i\right), \tag{92}$$

where by the superscript I we indicate the fact that there is no common factor among all the q. In our case we find

$$f_c(s) = \overline{\langle f\rangle^s}_c \approx \pi^{-D/2} \sum_{q_1\cdots q_s}^{I} \left(\frac{\pi^{s-1}}{\sum_{i=1,s} q_i^2}\right)^{D/2}. \tag{93}$$

We note that, if H is an homogeneous function of degree $-\nu$,

$$\sum_{q_1\cdots q_s} H(q) = \sum_{p=1,\infty} \sum_{q_1\cdots q_s}^{I} H(pq) = \sum_{p=1,\infty} p^{-\nu} \sum_{q_1\cdots q_s}^{I} H(q) = \zeta(\nu) \sum_{q_1\cdots q_s}^{I} H(q). \tag{94}$$

We finally find that (within the approximation of keeping only the linear constraint)

$$f_c(s) = \frac{\Gamma(D/2)}{\pi^{D/2}\zeta(D)} \int \frac{dt}{t} t^{D/2} G(t)^s, \tag{95}$$

where

$$G(t) = \pi^{1/2} \sum_{k=1,\infty} \exp(-tk^2). \tag{96}$$

Few remarks are in order.

- In this case the function $f_c(s)$ can be written under the form

$$f_c(s) = \int dz p(z) z^s, \tag{97}$$

where

$$p(z) \propto \int \frac{dt}{t} t^{D/2} \delta(G(t) - z). \tag{98}$$

It is crucial to note that the expression for $p(z)$ is well defined also in the case where some of the moments of the function f are divergent. The divergence of the moments of the function f is related to the asymptotic behaviour of the function $p(z)$ at large z.

- We have that for small t

$$G(t) = \pi^{1/2} \left(t^{-1/2} + \tilde{G}(t) \right), \tag{99}$$

where $\tilde{G}(t)$ is a C^∞ function at $t = 0^+$ (it is a well known fact that can be easily proved using the Poisson formula). Therefore the function $f_c(s)$ (as defined in (95)) has simple poles on the real axis at integer values of s starting from $s = D$. The first pole correspond to a decrease of the function $p(z)$ as z^{-D-1} at large z. The other poles are located at $s = D + 1 + 2n$, with integer non-negative n, as can be seen from the representation:

$$f_c(s) \propto \int \frac{dt}{t} t^{(D-s)/2} \left(1 + t^{1/2} \tilde{G}(t) \right)^s. \tag{100}$$

- For dimensions $D > s$ the quantity $f_c(s)$ (as defined in (95)) is an analytic function of D that has some poles for some values of D, i.e. when $D = s$ or $D = s - 1 - 2n$. Moreover $f_c(s)$ is an analytic function of both variables s and D.

6.3 A Preliminary Conjecture

We conjecture that these analyticity properties in D and s are true in general also if we retain the terms with more that one linear constraint. This conjecture implies that each given moment can be computed at sufficient high dimensions, where it is convergent, and that we can evaluate it as analytic continuation in D and s at the point we need it. The analytic continuation of the moments in the dimensions D at fixed s and the analytic continuation of the moments in the value of s at fixed dimensions D should coincide. If in a given dimension they are both singular, the form of the singularity in s and in D should be related.

Moreover the functions $G(z)$ and $C(z)$, that have been introduced in (72), are well defined for negative z in terms of the function $P(k)$. We also conjecture that if we use the results coming from analytic continuation from high dimensions and we plug them in (80), we obtain the correct result for the function $C(z)$. Of course the function $C(z)$ would not be C^∞ at $z = 0$ and it would have a cut on the real positive axis, where its integral representation is not convergent. A similar situation has been conjectured in the framework of the perturbative expansion for non-renormalizable quantum field theories long time ago [19].

If these conjectures are true, Rogers formulae bring information on the function $P(k)$ also in the region where its moments are divergent: our program make sense.

These conjectures are quite strong. It could be possible to verify by explicit computations that the function $f_c(s)$ has the needed analytic properties in the region where Rogers formulae are convergent. On the other hand it is not obvious that the function $f_c(s)$ at given

dimension coincides with its analytic continuation from high dimensions (the function $f_c(s)$ can be obtained from the function $P(k)$ by inverting the representation (80)).

We do not know how these conjectures could be proved in a rigorous and systematic way. One could in principle check by explicit computations that the leading singularities at $z = 0$ (or equivalently the tails of the function $P(k)$ at large k) are correctly reproduced, however it is not clear how to automatize such a computation. It would be interesting to check if this conjecture is supported by a numerical study of the first moments in low dimensions. We will not attempt to further study this point here.

7 A Comparison with the Statistical Mechanics of Hard Spheres

7.1 The Statistical Mechanics of Hard Spheres

Here we try to get some tentative conclusions for the asymptotic behaviour of r_M in large dimensions. In order to do this we have to recall some statistical properties of a gas (or solid depending on the density) of hard spheres.

An object that it is quite familiar to physicist is the partition function of a system of hard spheres of diameter R:

$$Z(\mathcal{V}, N) = \int \prod_{i=1,N} dx_i \prod_{i,k=1,N} \theta(|x_i - x_k| - R). \tag{101}$$

The partition function depends on the total number of particles N and on the volume $\mathcal{V}$ of the box where the particle are confined. Without loss of generality we can assume that $\mathcal{V} = N$, i.e. the density equal to one. In the infinite volume limit we have that

$$Z(\mathcal{V}, N) \approx \exp(\mathcal{V}\, S_H(r)), \tag{102}$$

where $S_H(r)$ is the entropy density of hard spheres and r is the reduced diameter defined in (24). The entropy is a function of the reduce diameter r, it should diverge when $r = r_A$, i.e. the highest density packing (for $r > r_A$ we have $Z = 0$). It is hard problem to decide if in the infinite volume limit the highest density packing is a lattice packing and if $r_A = r_M$.

In dimensions 3 the most compact packing is a lattice packing and $r_A = r_M$. Therefore the partition function vanishes as soon as $r > r_M$ and

$$\lim_{r \to (r_M)^-} S_H(r) = -\infty. \tag{103}$$

A crucial phenomenon for a system of hard spheres in three dimensions is crystallization: there is a diameter r_C such that for $r > r_C$ the configurations of the hard particles are very similar to those of a regular crystal. The transition from the fluid to the crystal phase is a first order transition.

Generally speaking in higher dimensions we can have two possibilities:

1. The partition function of hard spheres is different from zero also in the region $r > r_M$ and diverges at a value of $r = r_A > r_M$. In this case r_A is the maximal diameter of a non-crystalline (i.e. amorphous) packing and it is bigger of the maximal diameter of a crystalline packing. This situation is quite different from the three dimensional situation.
2. As in the three dimensional case the partition function of hard spheres becomes zero at $r_M = r_A$. We can however distinguish among two cases.

a) There is a diameter r_C where the entropy has a singularity and the system crystallize.
b) There is no crystallization transition.

The situation become more complex if we consider also the possibility of having a glass transition in the non-crystalline phase [20–23] at the point r_G. This glass transition may be or in the metastable phase above r_C, as in three dimensions, or in the stable liquid phase below r_C. This glass phase transition is important because the virial expansion does not give information on the behaviour above r_G. The arguments of [20–23] predict that $r_G = 1$.

7.2 A First Look to the Virial Expansion

The effective density of the particles is proportional to r^D and excluded volume effects becomes larger when r increases. The usual virial expansion may be used to compute $S(r)$ as function of r and of the fugacity in the grain-canonical ensemble:

$$\mathcal{Z}(\mathcal{V}, z) = \sum_N \frac{z^N}{N!} Z(\mathcal{V}, N) \equiv \exp(-\mathcal{V} F(z)). \tag{104}$$

Eventually we can adjust the fugacity z in such a way to have density one. These computations are well described in the literature. Diagrams are classified according to increasing complexity.

Our aim is to compare the virial expansion for the partition function with the formulae that we have obtained for $\ln(P(0, r))$. Let us consider firstly the results that we obtain if we consider a particular class of diagrams, i.e. those of a chain. In this case one finds that

$$-S_H(r) = V(R) + \sum_{k=3,\infty} k^{-1} \int \prod_{i=1,k} d\nu(x_i) \delta\left(\sum_{i=1,k} x_i\right), \tag{105}$$

where the relation between R and r is given by (24) and

$$d\nu(x) = d^D x \theta(R - x). \tag{106}$$

We would like to compute the quantity $S_L(r) \equiv \ln(P(0, r))$. We can follow the same arguments as in the Sect. 4.3. One finds, among many other terms the following terms (neglecting factors two)

$$-S_L(r) = V(R) + \sum_{k=2,\infty} \int \prod_{i=1,k} \left(d\nu(x_i) \sum_{q_i}\right) \delta\left(\sum_{i=1,k} x_i q_i\right), \tag{107}$$

where the sum over the q is done with the condition:

- All q are different from zero.
- Conventionally q_1 is positive.
- The set of k variables q_i for $i = 1, k$ is irreducible, i.e. they do not have common factors.
- For $k = 2$ the terms with $q_1 = 1 = \pm q_2$ are absent.

At each given k the leading term in the sum comes from those term where $q_1 = 1$ and $q_i = \pm 1$. If we retain only these contributions, we find (neglecting factors two) exactly the same result that for the $S_H(r)$. Therefore, within the approximations of considering only the terms coming from one linear constraint and the for each moment taking only the leading

term, the expression for $S_L(r)$ coincide with that for $S_H(r)$ obtained summing a subset of diagrams of the virial expansion.

This result should not be surprising and it can be simply understood: we have seen that the distribution of the points of a random lattice are random if they do not satisfies linear constraints (see Sect. 4.4).[10] Moreover the probability distribution of a finite (not diverging with D) number of lattice points would be just a same of points that are not constrained to be on the lattice.

7.3 A Naive Conjectures

Let us consider the generic contribution to the virial expansion. Every diagrams of the virial expansion of $S_H(r)$ with k lines and L loops can be written be written in the form:

$$\int \prod_{i=1,k} dv(x_k) \prod_{l=1,L} \delta^D \left(\sum_{i=1,k} M_{l,i} x_i \right), \tag{108}$$

where the quantities M take the value 0 or 1. The form of the matrix M fixes the topology of the diagram.

It is possible to check that all the contributions of the previous form also appears in the computation of $S_L(r) = -\ln(P(0,r))$, and the terms with many loops correspond to terms with more than one linear constraint. All the diagrams of the virial expansion appear in the computation of $S_L(r)$, however this last expression contains much more terms that are not present in the virial expansion.

It is possible that the leading contributions in both problems are the same (as it happens in the case of the chain approximation) so one could may be tempted to conjecture that in the limit of infinite dimensions the two problems coincide and we have that

$$S_H(r) \approx S_L(r). \tag{109}$$

The previous equality would implies that in infinite dimensions the maximal density of lattice packing is equal to maximal density of packing without any lattice constraint. Although the conjecture may be true order by order in the virial expansion, we should be very careful because the sum of subdominant term may become dominant. However the presence of the glassy transition at $r = 1$ [20–23] may introduce extra complications, as we shall see in the next section.

7.4 A More Refined Conjecture and Open Problems

In order to further understand the problem we need to control both the hard sphere gas and the lattice partition function in large dimensions D and eventually to treat the difference among the two problems as a perturbation. The task of computing the properties of the hard sphere gas in the infinite dimensional limit is non trivial (when $r > 1$) [18, 20–26]. In some sense we have transformed an unsolved problem into an other unsolved problem. The situation is not so bad, as far as the computation of the thermodynamical properties of an hard sphere gas in the fluid phase may be not out of reach. One should firstly obtain some results in the fluid phase for $r > 1$, and study the properties of the glass transition is present. However the fluid phase may not exist for $r > 1$ because of the glass transition.

[10] The matrix Λ contains a divergent number of parameters when the dimensions D goes to infinity.

Before presenting some new definitions and conjectures let us recall that some progresses have been recently done. By resuming the contributions that correspond the hypernetted chain approximation one obtain an expressions for the entropy in the infinite dimensional case in the region $r < e/2$ [18]. Let us call it $S_{I_0}(r)$.

Of course the hypernetted chain approximation is not the reality, many diagrams (the so called the *bridge* diagrams) have been neglected. We can consider to correct the hypernetted chain approximation by introducing the simplest bridge diagrams. Let us suppose that we are able to solve the corresponding integral equations in the large dimensions limit. In this way we could compute a better approximation to the entropy that we call $S_{I_1}(r)$. We can consider more complex bridge diagrams and we can get in this way a new expression for the entropy. Let us suppose that this process converges: we call the final function $S_I(r)$.

This function is interesting because there are some chances that it can be computed, or estimated, by careful computations. I have not started such computation, however I think that the existence of the function $S_I(r)$ is a reasonable conjecture. If at some value of r the bridge diagrams become important, one could consider more complex equation. If only a finite number of bridge diagrams become relevant one after the other, this program could be implemented because many diagrams can be computed analytically in very large dimensions.

We conclude that it is convenient to consider three entropy functions for large dimensions D. Let us summarize the rather complex situation:

- $S_H(r)$ is the entropy for a gas of hard spheres as function of the diameter. It may have singularities at various values of r. In three dimensions it has a singularity at r_C that corresponds to crystallization. It seem that in three dimensions at $r_G > r_C$ in the metastable region there is a glass transition, characterized by a faster increase of the pressure.[11] It is possible than in higher dimensions the relative position of the glass transition and the crystal transition do change. It is also possible that in high dimensions the crystal like transition disappears, but this would be quite unlikely. The value of r_C is estimated to be equal to 1 [20–23].
- $S_L(r) = \ln(P(0, r))$, i.e. the logarithm of the probability of finding a lattice packing with reduced radius r. It should diverge toward infinity when we reach the maximum density for a lattice packing, i.e. r_M. The reader should notice that this function can be written as an integral over a finite dimensional space so for fixed D it should be a continuous function of the radius. Only in the infinite dimensions it could develop a real first order phase transition, however it could be possible to observe the premonitory signs of this transition by evaluating it numerically. It is also not clear if a glass transition at $r = 1$ is present also for this function.
- The two previous quantities are not easy to compute, although they are well defined. One could use here a third quantity that is maybe slightly less defined, but it would be much easier to compute, i.e. the function $S_I(r)$ that we have just introduced and it is the entropy computed using integrals equations that generalize the hypernetted chain approximation, as discussed above. In other words $S_I(r)$ coincides with the results of the hypernetted chain approximation as soon the bridge diagrams are neglected, but it can be much more complex.

[11] As far as this phenomenon happens in the metastable phase, things are not so well defined; however one can introduce slightly modified Hamiltonians that forbids crystallization [27].

In this contest there are a few points that should be investigated.

- The computation of $S_I(r)$ is the more urgent task in order to obtain analytic predictions. While the situations is more or less clear in the HNC, it is not clear if there is a value of the density where bridge diagrams must be considered. Having under control $S_I(r)$ would be very interesting.
- One should extend to the lattice case the analytic tools that are used to predict a glass transition in the hard sphere liquid [20–23], in order to see if they predict a phase transition also in $S_L(r)$.
- One should numerically evaluate the quantity $S_L(r)$, at least in low dimensions where numerical computations should be feasible, to see if there are signs of a first order transition by increasing the dimensions.

If we succeed in showing that $S_I(r)$ is well defined, one can put forward the conjecture that

$$S_I(r) = S_L(r). \tag{110}$$

A preliminary test of this conjecture can be done by verifying that the terms that we have neglected in the difference do not pile up to something finite.

In conclusion this paper has opened more problem that those it has solved, however I am confident that progress can be done on this line. The difference and the commonalities of the two problems we are investigating could be further clarified.

Acknowledgements It is a pleasure to thank Jorge Kurchan and Marc Mézard for their contribution in the first steps of this work and for the many discussions on the subject of this paper. I would also to thank Francesco Zamponi for many interesting discussions on this and related problems and Salvatore Torquato for interesting correspondence. I am grateful to Bernard Derrida for a careful reading of the manuscript.

Appendix: On the Large Dimensional Limit of Some Integrals

In the approach of this paper it is crucial to evaluate some integrals in the limits where the dimensions D of the space go to infinity. Most of the integrals we consider go to zero exponentially. This result is related to the fact that in large dimensions most vectors are orthogonal. Here will mainly discussing the exponential factors and we will neglect the prefactors.

Before considering the general case, let us start by studying a simple example:

$$A(D) = \int d\mu(x) d\mu(y) \delta((x+y)^2 - 1), \tag{111}$$

where

$$d\mu(x) = S(D)^{-1} 2\delta(x^2 - 1) d^D x. \tag{112}$$

The factor

$$S(D) = \frac{2\pi^{D/2}}{\Gamma(D/2)} \tag{113}$$

is the surface of the D dimensional sphere and it is such that

$$\int d\mu(x) = 1. \tag{114}$$

We go to polar coordinates and (neglecting prefactor that are at most powers of D) we find,

$$A(D) \approx \int d\cos(\phi)\sin(\phi)^D \delta\left(\cos(\phi) - \frac{1}{2}\right) \approx \left(\frac{3}{4}\right)^{D/2}. \tag{115}$$

The same result can be obtained if we consider the same integral with the measure

$$d\nu(x) = S(D)^{-1} d^D x \theta(1 - x^2), \tag{116}$$

still satisfying the condition

$$\int d\nu(x) = 1. \tag{117}$$

We can also consider the integral

$$A(D) = S(D)^{-1} \int d\nu(x) d\nu(y) \theta((x + y)^2 - 1). \tag{118}$$

This integral is proportional to the phase space for having three sphere with each sphere touching each of the remaining two.

In large dimensions the measure of a sphere is concentrated on its surface and the difference among a θ function or a δ function of the radius is irrelevant. Therefore the integral have a similar limit when the dimension goes to infinity and $A(D) \approx B(D)$.

The same result could be obtained by using the integral representation for the θ function

$$\theta(z) = \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \frac{d\alpha}{\alpha + i\epsilon} \exp(i\alpha z). \tag{119}$$

In this way we obtain that

$$\begin{aligned} B(D) &\approx \Gamma(D/2)^2 \int d\alpha_1 d\alpha_2 d\alpha_3 \exp(i\Sigma)(i\mathcal{D})^{-D/2} \\ &= \Gamma(D/2)^2 \int d\alpha_1 d\alpha_2 d\alpha_3 \exp(i\Sigma - D/2 \ln(i\mathcal{D})), \end{aligned} \tag{120}$$

where

$$\Sigma = \alpha_1 + \alpha_2 + \alpha_3, \quad \mathcal{D} = \alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_3\alpha_1.$$

The saddle point equations are:

$$i = \frac{D(\alpha_2 + \alpha_3)}{2\mathcal{D}} = \frac{D(\alpha_1 + \alpha_3)}{2\mathcal{D}} = \frac{D(\alpha_1 + \alpha_2)}{2\mathcal{D}}. \tag{121}$$

The solution to the saddle equations is given by

$$\alpha_1 = \alpha_2 = \alpha_3 = i\frac{D}{3}. \tag{122}$$

Substituting back we find the result (115).

Alternatively, we can introduce the variables

$$z_i = \frac{\alpha_i}{\Sigma} \quad \text{for } i = 1, 3, \tag{123}$$

that satisfy the constraint

$$\sum_{i=1,3} z_i = 1. \tag{124}$$

We finally find

$$\Gamma(D/2)^2 \int d\Sigma \Sigma^{-D-1} \exp(i\Sigma) dz_1 dz_2 dz_3 \delta\left(\sum_{i=1,3} z_i - 1\right) \mathcal{D}^{-D/2}$$

$$= \frac{\Gamma(D/2)^2}{\Gamma(D)} \int dz_1 dz_2 dz_3 \delta\left(\sum_{i=1,3} z_i - 1\right) \mathcal{D}^{-D/2}$$

$$\approx 2^{-D} \int dz_1 dz_2 dz_3 \delta\left(\sum_{i=1,3} z_i - 1\right) \mathcal{D}^{-D/2}, \tag{125}$$

where

$$\mathcal{D} = z_1 z_2 + z_2 z_3 + z_3 z_1. \tag{126}$$

The saddle point is given by

$$z_i = \frac{1}{3}, \tag{127}$$

and we recover the previous result.

The same technique can be used to estimate the integrals

$$B_N = \int \prod_{i=1,N-1} d\nu(x_i) \theta\left(1 - \left(\sum_{i=1,N-1} x_i\right)^2\right). \tag{128}$$

We find

$$B_N \approx \left(\frac{N}{N-1}\right)^{N-1} \left(\frac{1}{N-1}\right)^{D/2} \approx_{N\to\infty} \left(\frac{e}{N}\right)^{D/2}. \tag{129}$$

Other integrals may be estimated in the same way. We may be interested to consider the cluster integral:

$$C = \int d\nu(x_1) d\nu(x_2) d\nu(x_3) \theta(1 - (x_1 + x_2)^2) \theta(1 - (x_2 + x_3)^2) \theta(1 - (x_3 + x_1)^2). \tag{130}$$

This integral is proportional to the phase space for having four sphere with each sphere touching each of the remaining three. We find that

$$C \approx \frac{\Gamma(D/2)^3}{\Gamma(3D/2)} \int \prod_{i=1,6} dz_i \delta\left(\sum_{i=1,6} z_i - 1\right) \mathcal{D}^{-D/2}, \tag{131}$$

where

$$\mathcal{D} = z_1 z_2 z_3 + z_1 z_3 z_4 + z_2 z_3 z_4 + z_1 z_2 z_5 + z_1 z_3 z_5 + z_1 z_4 z_5 + z_2 z_4 z_5 + z_3 z_4 z_5$$
$$+ z_1 z_2 z_6 + z_2 z_3 z_6 + z_1 z_4 z_6 + z_2 z_4 z_6 + z_3 z_4 z_6 + z_1 z_5 z_6 + z_2 z_5 z_6 + z_3 z_5 z_6. \tag{132}$$

Finally using the saddle point

$$z_i = \frac{1}{6}, \tag{133}$$

we find

$$C \approx \left(\frac{1}{2}\right)^{D/2}. \tag{134}$$

A related, but technically different problem arise in computing series of the form

$$G(r) \equiv \sum_{N=1,\infty} (-r)^{ND} A_N, \tag{135}$$

when D is large and A_N are defined in (128). As soon as $r > 1$ each term of the series is divergent for large N. However the alternating sign structure implies a certain amount of cancellation. This kind of computation can be found in [18], we sketch here for completeness.

At this end it is useful to observe that using the convolution theorem, starting from the definition in that (128) we find that

$$A_N = \int d^D p I(p)^N, \tag{136}$$

where $I(p)$ is the Fourier transform of the measure $dv(x)$. We thus find

$$G(r) = \int d^D p (1 + r^D I(p))^{-1}. \tag{137}$$

After a short computation we find that $I(p)$ is proportional to a Bessel function. It is positive at small p and it has a negative minimum at $p = p_m$, where p_m behaves as

$$p_m = aD + O(1), \quad a = \frac{1}{2}. \tag{138}$$

The value of the quantity $I(p_m)$ can be estimated to be equal to

$$I(p_m) = -b^{-D}, \quad b = \sqrt{e/2} \approx 1.166. \tag{139}$$

This result for the function $I(p)$ can be obtained directly from the known formulae on Bessel functions. It may be instructive to derive it from scratch. We can start from the representation for D-dimensional Fourier transform in polar coordinates

$$I(p) = \int_{-1}^{1} dx (1 - x^2)^{(D-1)/2} \exp(ipx). \tag{140}$$

Alternatively we can write

$$I(p) = \frac{\pi^{D/2}}{V(D)} \int \frac{d\lambda}{\lambda} \exp\left(-\frac{p^2}{4\lambda}\right) \lambda^{-D/2}. \tag{141}$$

In both cases one finds a purely imaginary saddle point in the variable x (or λ) that gives a purely real contribution. For example the value of x at the saddle point is given by

$$x = \frac{2z}{1 + \sqrt{1 - 4z^2}}, \tag{142}$$

where

$$Dz = p. \tag{143}$$

Therefore for positive $p < Da$ we finds that

$$I(p) = \exp(-Dg(z)), \tag{144}$$

where the function g is given by

$$g(z) = \frac{1}{2}\ln\left(\frac{1+\sqrt{1-4z^2}}{2}\right) - \frac{2z^2}{1+\sqrt{1-4z^2}}. \tag{145}$$

On the contrary as soon as $p > Da$ the saddle point is no more purely imaginary and the result is complex. In the region where $p = Da + t$, t being of order one when $D \to \infty$, one finds

$$I(p) = \exp(-Dg(a)h(t)), \tag{146}$$

where $h(t)$ is an oscillating function of order 1.

One can use these results to show that the chain approximation does not present difficulties for $r < B$, but it diverges as soon $r > B$ as the effect of the pile up of sub-asymptotic terms.

References

1. Conway, J.H., Sloane, N.J.A.: Sphere Packings, Lattices and Groups. Springer, New York (1998)
2. Minkowski, H.: J. Reine Angew. Math. **129**, 274 (1905)
3. Mézard, M., Parisi, G., Virasoro, M.A.: Spin Glass Theory and Beyond. World Scientific, Singapore (1987)
4. Rogers, C.A.: Acta Math. **94**, 249 (1955)
5. Rogers, C.A.: Proc. Lond. Math. Soc. **8**, 609 (1958)
6. Rogers, C.A.: Proc. Lond. Math. Soc. **3**, 305 (1956)
7. Rogers, C.A.: Packing and Covering. Cambridge University Press, Cambridge (1962)
8. Parisi, G., Rizzo, T.: in preparation
9. Kurchan, J., Mézard, M.: private communication
10. Siegel, C.L.: Ann. Math. **46**, 340 (1945)
11. Kabatiansky, G.A., Levenshtein, V.I.: Probl. Inf. Trans. **14**, 1 (1978)
12. Cohn, H., Elkies, N.: Ann. Math. **157**, 689 (2003)
13. Levenshtein, V.I.: Sov. Mat. Dokl. **20**, 417 (1979)
14. Derrida, B.: Phys. Rev. B **24**, 2613 (1981)
15. Frisch, H.L., Rivier, N., Wyler, D.: Phys. Rev. Lett. **54**, 2061 (1985)
16. Frisch, H.L., Rivier, N., Wyler, D.: Phys. Rev. A **36**, 2422 (1987)
17. Frisch, H.L., Percus, J.K.: Phys. Rev. E **60**, 2942 (1999)
18. Parisi, G., Slanina, F.: Phys. Rev. E **62**, 6554 (2000)
19. Parisi, G.: Nucl. Phys. B **100**, 368 (1975)
20. Parisi, G., Zamponi, F.: J. Chem. Phys. **123**, 144501 (2005)
21. Parisi, G., Zamponi, F., J. Stat. Mech. (2006)
22. Parisi, G., Zamponi, F.: J. Stat. Phys. **123**, 1145 (2006) P03017
23. Zamponi, F.: arXiv:cond-mat/0604622v2
24. Torquato, S., Stillinger, F.H.: Exp. Math. **15**, 307 (2006)
25. Stillinger, F.H., Torquato, S.: Phys. Rev. E **73**, 031106 (2006)
26. Scardicchio, A., Stillinger, F.H., Torquato, S.: arXiv:0705.1482 (2007)
27. Angelani, L., Di Leonardo, R., Parisi, G., Ruocco, G.: Phys. Rev. Lett. **87**, 055502 (2001)