Appendix A

Avoided volume

The volume part ω_{NCC} of the total microcanonical weight (see eq. (3.12a) on page 48) describes the accessible volume to N_f non–overlapping fragments positioned within a spherical container of volume V. For low pressures ($\sim$ large volumes) analytical forms (exact and approximation) of ω_{NCC} can be worked out (see sec. A.1). At very high densities, near the critical packing fraction (see below) a free–volume theory can be formulated (see e.g. [AW62] and refs. quoted therein). Unfortunately for intermediate pressures these approximations are not valid and there is no available ansatz suitable for MMMC even for simple models of hard spheres, see e.g. [LUD01]. This issue is yet an unsolved, mathematical problem, that generates a lot of literatures in mathematics [WIL98, CS89, ARH98] but also for concrete applications in crystallography [WIL91, BS97], in chemical physics (via the equation of state, see sec. A.1.2), nuclear physics [BBI$^+$95, RRS88, RAD01], molecular biology (see e.g. [KSS01]), etc.

A.1 Analytical expressions — EOS ansatz

A.1.1 Exact expressions

For clarity, let us first recall the definition of NCC

$$NCC \doteq \frac{\int_V \dots \int_V d\mathbf{r}_1 \dots d\mathbf{r}_{N_f}}{\int_V \dots \int_V d\mathbf{r}_1 \dots d\mathbf{r}_{N_f} \Theta\left(\mathbf{r}_1, \dots, \mathbf{r}_{N_f}\right)\ \eta\left(\mathbf{r}_1, \dots, \mathbf{r}_{N_f}\right)} \tag{A.1a}$$

$$= \frac{V^{N_f}}{\int_V \dots \int_V d\mathbf{r}_1 \dots d\mathbf{r}_{N_f} \Theta\left(\mathbf{r}_1, \dots, \mathbf{r}_{N_f}\right)\ \eta\left(\mathbf{r}_1, \dots, \mathbf{r}_{N_f}\right)}, \tag{A.1b}$$

where Θ forbids the overlapping of a cluster with the system boundary

$$\Theta\left(\mathbf{r}_1, \dots, \mathbf{r}_{N_f}\right) \doteq \begin{cases} 0 & \text{if at least one cluster overlaps with the boundary,} \\ 1 & \text{else;} \end{cases} \tag{A.2}$$

η forbids the overlapping between two clusters

$$\eta\left(\mathbf{r}_1, \dots, \mathbf{r}_{N_f}\right) \doteq \begin{cases} 0 & \text{if at least two clusters overlap,} \\ 1 & \text{else;} \end{cases} \tag{A.3}$$

NCC is the inverse probability to find a set of positions so that the clusters fit into the system volume.

For a single particle of radius r_1 in a container of volume $V = \frac{4\pi}{3}R^3$, the inverse probability to find a position that fits in V is

$$NCC = \frac{V}{\frac{4\pi}{3}(R - r_1)^3} = \left(1 - \frac{r_1}{R}\right)^{-3}, \tag{A.4}$$

where $\frac{4\pi}{3}(R - r_i)^3$ is the accessible volume to (the center of mass of) the particle, it is called the "eigen–accessible volume". Let us define the packing fraction $\kappa \doteq \frac{V_0}{V}$, where V_0 is the sum of the fragment volumes (here V_0 is simply $\frac{4\pi}{3}r_1^3$). NCC diverges at $R = r_1$, i.e. the critical packing fraction κ_c is equal to 1.

The one–particle NCC given by eq. (A.4) is called in the following "eigen–NCC".

When there are two spheres of radius r_1 and r_2 the divergence of κ is located at $R = r_1 + r_2$

$$1/2 \leq \kappa_c = \frac{r_1^2 + r_2^2}{(r_1 + r_2)^2} \leq 1.$$

NCC can also be computed analytically for the two spheres case. First consider that the cluster number 1 is fixed far from the boundary. The accessible volume to the second cluster is its eigen accessible volume $\frac{4\pi}{3}(R - r_2)^3$ minus the avoided volume due to cluster 1, $\frac{4\pi}{3}(r_1 + r_2)^3$. An integration over all the positions allowed for 1 yields to a first estimate of V_{acc}, the total accessible volume

$$V_{acc} = \left[\frac{4\pi}{3}(R - r_1)^3\right]\left[\frac{4\pi}{3}(R - r_2)^3 - \frac{4\pi}{3}(r_1 + r_2)^3\right]. \tag{A.5}$$

However when 1 is close to the boundary the forbidden volume for 2 produced by the wall and by 1 overlap partially (dotted region in fig. A.1 on the facing page). This common part is therefore counted twice in eq. (A.5). The volume of this common part is equal to the spherical segment one of a sphere of radius $r_1 + r_2$ and height $r_1 r_2 \sin\gamma$ (outer spheres in fig. A.1) minus the spherical segment volume of a sphere of radius $R - r_2$ and height $R - r_2 - r - r_1 r_2 \sin\gamma$. At fixed r the volume which is counted twice is

$$\begin{aligned} V_{for}(r) = {} & \frac{\pi}{3}(r_1 + r_2)^3 (1 - \cos\gamma)^2 (2 + \cos\gamma) \\ & - \frac{\pi}{3}(R - r_2 - r - (r_1 + r_2)\cos\gamma)^2 (2(R - r_2) + r + (r_1 + r_2)\cos\gamma), \end{aligned} \tag{A.6}$$

where

$$\cos\gamma = -\frac{1}{2}\frac{r^2 + (r_1 + r_2)^2 - (R - r_2)^2}{r(r_1 + r_2)}. \tag{A.7}$$

An integration of $4\pi r^2 V_{for}(r)\, dr$ between the radiuses $r = R - r_1 - 2r_2$ and $r = R - r_1$ yields

$$V_{for} = \frac{16\pi^2}{9} r_2^3 \left(3r_1^2 r_2 - 3R\, r_2^2 + 3R^2 r_2 + 3r_1\, r_2^2 + 6R^2 r_1 - 6r_1^2 R - 9r_1 r_2 R + r_2^3\right). \tag{A.8}$$

The total accessible volume is the sum of the right hand sides of eqs. (A.5) and (A.8). Finally, NCC is the inverse of this sum times V^2

$$NCC = \frac{R^6}{\left(R^3 + 3r_1 r_2 R - r_1^3 - r_2^3\right)(R - r_1 - r_2)^3}. \tag{A.9}$$

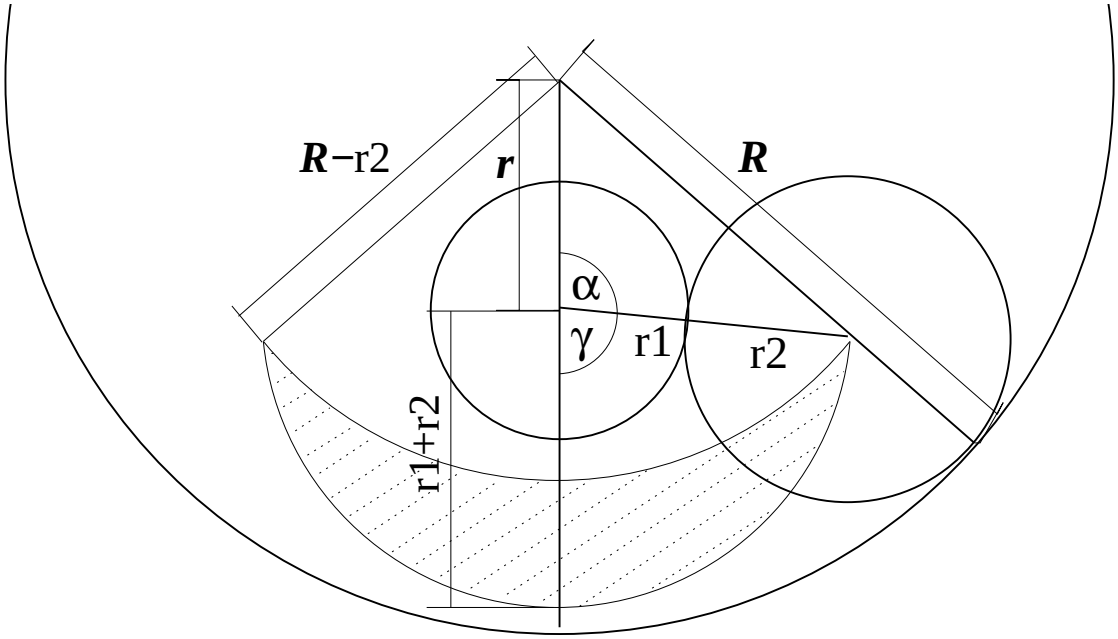


Figure A.1: Illustration of the avoided volume for two particles near the boundary. The forbidden region (dotted region) is counted twice in eq. (A.5) on the preceding page. In this two dimensional representation the surface counted twice is a crescent. Its area is equal to the outer disk segment area (radius of $r_1 + r_2$, angle 2γ) minus the inner one area (radius $r + r_1 + r_2$, angle $\arccos\left[\frac{(R-r_2)^2+r^2-(r_1+r_2)^2}{2r(R-r_2)}\right]$).

Using eq. (A.9), one can build a "two body" estimate of NCC for a system of N_f hard spheres

$$NCC_2 \doteq \left(\prod_{i<j}^{N_f} NCC(i,j)\right)^{\frac{1}{N_f-1}}. \tag{A.10}$$

where $NCC(i,j)$ is given by eq. (A.9). The exponent $\frac{1}{N_f-1}$ takes into account the fact that the avoided volume of a given particle has been counted $N_f - 1$ times in eq. (A.10). The NCC given by eq. (A.10) is a sort of geometric average. On the plus side, NCC_2 has a critical packing fraction which varies which the mass distribution μ and is smaller than one. One the minus side, from an algorithmic point of view, the updating scheme of NCC_2 when two fragment sizes are changed is of the order of $\mathcal{O}(N)$.

A.1.2 Approximations

The results presented in the previous subsection are the only exact ones used in MMMC. For $N > 3$ approximations are needed.

There exist a large literature on *infinite* ($N \to \infty$, $N/V = cst$) *unbounded* diluted ($\sim$ small pressures) gas of hard spheres (see e.g. [RFL59, SSM87, PCA]). Some results from these works are used in the following to estimate NCC for *finite bounded* systems.

The part in NCC due to the interactions with the wall vanishes in the infinite limit, so in the following only the part in NCC due to the interactions between spheres is worked

out. In order to add some boundary effect, as a first approximation, NCC is divided by $\frac{R^{3N}}{\prod_{i=1}^{N}(R-r_i)^3}$. The latter term is the product of the clusters eigen–NCC (see eq. (A.4)).

For an infinite gas of hard spheres one defines Z as [PAT72]

$$Z \doteq \frac{p}{T}, \tag{A.11a}$$

$$\equiv \left.\frac{\partial S}{\partial V}\right|_E. \tag{A.11b}$$

$Z = \frac{p}{T}$ is the equation of state of the gas ($\frac{ZN}{V}$ is the compressibility factor). The volume dependent part of the entropy of a hard sphere gas is

$$S = \ln \Omega = \ln \int_V \cdots \int_V \eta\left(\mathbf{r}_1, \cdots \mathbf{r}_{N_f}\right) \mathrm{d}\mathbf{r}_1 \cdots \mathrm{d}\mathbf{r}_{N_f} \tag{A.12}$$

where η is equal to zero when two clusters overlap, and to one otherwise, see eq. (A.3) on page 109. Of course Ω depends on the mass (size) distribution $\mu = \{N_1, N_2, \ldots, N_{N_f}\}$ of particles. NCC as already defined is

$$NCC\left(\mu, V\right) = \frac{V^{N_f}}{\Omega\left(\mu, V\right)}. \tag{A.13}$$

In the perfect gas approximation the spheres do not interact therefore Ω is simply V^{N_f} ($NCC = 1$), and the equations (A.11) and (A.13) lead to the following well–known equation of state

$$\frac{N_f}{V} = \frac{p}{T}. \tag{A.14}$$

For a gas of equal size hard spheres (of radius r) there exists another well-known simple approximation, the van der Waals approximation [DGLR89, LL94]. Since the minimal distance between two particles is $2r$, the forbidden volume due to these two spheres is approximatively $\frac{4\pi}{3}(2r)^3$. The total number of pairs is $\approx \frac{N_f^2}{2}$, hence the total avoided volume per particle is

$$V_{avoid} \approx \frac{1}{N_f}\frac{N_f^2}{2}\frac{4\pi}{3}(2r)^3 = 4V_0, \tag{A.15}$$

where $V_0 = N_f \frac{4\pi}{3} r^3$. Consequently

$$\Omega \approx \left(V - 4V_0\right)^{N_f}, \tag{A.16}$$

and eq. (A.11) yields

$$Z = \frac{N_f}{V - 4V_0}. \tag{A.17}$$

In the last two examples, Z were computed from NCC (or Ω). Conversely, one can use Z to determine NCC. Indeed NCC can be linked to Z in the following way

$$NCC(\mu, V) = \frac{V^{N_f}}{\exp\left(S\left(\mu, V\right)\right)} \qquad \text{by (A.12) and (A.13),} \tag{A.18}$$

$$= \frac{V^{N_f}}{\exp\left(C\left(\mu, V_0\right) + \int_{V_0}^{V} Z\left(\mu, V\right) dV\right)} \qquad \text{by (A.11),} \tag{A.19}$$

where $C(\mu, V_0)$ is the constant of integration. At $V = V_0$ the following relation holds

$$C(\mu, V_0) = N_f \ln V_0 - \ln NCC(\mu, V_0). \tag{A.20}$$

At fixed mass distribution and in the limit $V_0 \to \infty$, the forbidden volume is constant while the accessible one diverges, i.e. $\lim_{V_0 \to \infty} NCC = 1$; therefore

$$\lim_{V_0 \to \infty} \frac{C(\mu, V_0)}{N_f \ln V_0} = 1. \tag{A.21}$$

Since in the following, all the quantities are worked out at *fixed* mass distribution μ and for sake of simplicity the μ–dependence is not explicitely written.

An usual approach to estimate Z is to develop it in powers of the density $\rho = \frac{N_f}{V}$

$$Z = \rho + B_2\rho^2 + B_3\rho^3 + \ldots \tag{A.22}$$

where B_i is called the i^{th} virial coefficient [DGLR89]. Again there exists a large literature addressing the computation of these coefficients (see e.g. [EAG98] and refs. quoted therein).

Another approach is based on the virial equation, e.g. for a mixture of hard spheres

$$Z = \frac{p}{T} = \frac{N_f}{V} - \frac{2\pi N_f^2}{3TV^2} \sum_{ij} x_i x_j \int_0^\infty \frac{\partial u_{ij}(r)}{\partial r} g_{ij}(r)\, r^3 dr \tag{A.23}$$

where $x_i = \frac{N_i}{N_f}$ is the relative density of the species i; $u_{ij}(r)$ is the interaction potential between two particles from species i and j whose center of mass are separated by distance r; $g_{ij}(r)$ is the pair correlation function between species i and j. For a given g_{ij} one can integrate (at least numerically) the equation of state [THI63, WER63, RFL59].

For MMMC95, the Mansoori–Carnahan–Starling–Leland [MCSL71] formula for Z is used

$$Z = \frac{6}{\pi}\left\{\frac{\xi_0}{(1-\xi_3)} + \frac{3\xi_1\xi_2}{(1-\xi_3)^2} + \frac{3\xi_2^3}{(1-\xi_3)^3} - \frac{\xi_3\xi_2^3}{(1-\xi_3)^3}\right\} \tag{A.24}$$

where $\xi_k = \frac{\pi}{6}\sum_i \rho_i d_i^k$, $d_i = 2r_i$ and $\rho_i = \frac{N_i}{V}$ is the density of species i.

In order to compute NCC one has to calculate $\int Z(V)dV$, see eq. (A.19) on the facing page. Because of the V dependence of Z, terms of the following form have to be integrated

$$W(p, q, \alpha_1, \alpha_2) = \frac{\frac{\alpha_1}{V^p}}{\left(1 - \frac{\alpha_2}{V}\right)^q} = \frac{\alpha_1 V^{q-p}}{(V - \alpha_2)^q}. \tag{A.25}$$

First the substitution $Y = V - \alpha_2$ is made and the integration of W becomes

$$\int W\, dV = \int \frac{\alpha_1 (Y + \alpha_2)^{q-p}}{Y^q} dY. \tag{A.26}$$

The following cases are of interest

- for $q - p = 0$ (see the first three terms of eq. (A.24))

$$\begin{aligned} \int W\, dV &= \int \frac{\alpha_1}{Y^q} dY \\ &= \begin{cases} \alpha_1 \ln Y & \text{when } q = 1, \\ \frac{\alpha_1}{1-q} Y^{1-q} & \text{else,} \end{cases} \end{aligned} \tag{A.27}$$

- for $q-p=-1$ and $q=3$ (see the last term of eq. (A.24) on the preceding page)

$$\begin{aligned}\int W dV &= \int \frac{\alpha_1}{(Y+\alpha_2)\,Y^3} dY \\ &= -\frac{\alpha_1 \ln(Y+\alpha_2)}{\alpha_2^3} + \frac{\alpha_1 \ln Y}{\alpha_2^3} + \frac{\alpha_1}{\alpha_2^2 Y} - \frac{\alpha_1}{2\alpha_2 Y^2} \qquad (A.28)\\ &= \frac{\alpha_1 [\ln Y - \ln(Y+\alpha_2)]}{\alpha_2^3} + \frac{\alpha_1}{\alpha_2 Y}\left(\frac{1}{\alpha_2} - \frac{1}{2Y}\right). \qquad (A.29)\end{aligned}$$

Now one can define

$$\gamma_k = \xi_k V = \frac{\pi}{6}\sum_{i=1} N_i^A (2r_i)^k, \qquad (A.30)$$

where i is an index for the species. Combining the results yields

$$\begin{aligned}\hat{S}(V) \equiv \int Z dV &= \frac{6}{\pi}\left\{\gamma_0 \ln(V-\gamma_3) - \frac{3\gamma_1\gamma_2}{(V-\gamma_3)} - \frac{3\gamma_2^3}{2(V-\gamma_3)^2}\right.\\ &\quad \left. + \frac{\gamma_2^3 [\ln V - \ln(V-\gamma_3)]}{\gamma_3^2} + \frac{\gamma_2^3}{(V-\gamma_3)}\left[\frac{1}{2(V-\gamma_3)} - \frac{1}{\gamma_3}\right]\right\}\\ &= \frac{6}{\pi}\left\{\gamma_0 \ln(V-\gamma_3) + \frac{\gamma_2^3 [\ln V - \ln(V-\gamma_3)]}{\gamma_3^2}\right.\\ &\quad \left. - \frac{\gamma_2}{(V-\gamma_3)}\left(3\gamma_1 + \frac{V}{(V-\gamma_3)}\frac{\gamma_2^2}{\gamma_3}\right)\right\}\\ &\equiv \frac{6}{\pi}\gamma_0 \ln(V-\gamma_3) + \hat{S}_1(V). \qquad (A.31)\end{aligned}$$

From the definition of $\hat{S}_1(V)$ one immediately deduces

$$\lim_{V\to\infty} \hat{S}_1(V) = 0, \qquad (A.32)$$

which is consistent with $\lim_{V_0\to\infty} \hat{S}(V_0) = N_f \ln V_0 = \lim_{V_0\to\infty} \frac{6}{\pi}\gamma_0 \ln(V_0-\gamma_3)$.

The denominator of eq. (A.19) on page 112 can be written as

$$\begin{aligned}\exp\left(C(V_0) + \int_{V_0}^{V} Z(V_1)\,dV_1\right) &= \frac{V_0^{N_f}}{NCC(V_0)} \frac{\exp\left(\hat{S}(V)\right)}{\exp\left(\hat{S}(V_0)\right)}\\ &= \frac{V_0^{N_f}}{NCC(V_0)} \frac{(V-\gamma_3)^{\frac{6}{\pi}\gamma_0}}{(V_0-\gamma_3)^{\frac{6}{\pi}\gamma_0}} \frac{\exp\left(\hat{S}_1(V)\right)}{\exp\left(\hat{S}_1(V_0)\right)} \qquad (A.33)\end{aligned}$$

Using eqs. (A.32) and (A.21) and the fact that $\gamma_0 = \frac{\pi}{6}N_f$, the limit of the denominator of eq. (A.19) when $V_0 \to \infty$ is

$$\lim_{V_0\to\infty} \exp\left(C(V_0) + \int_{V_0}^{V} Z(V_1)\,dV_1\right) = (V-\gamma_3)_f^N \exp\left(\hat{S}_1(V)\right). \qquad (A.34)$$

Finally NCC is

$$NCC(V,\mu) = \left(\frac{V}{V-\gamma_3}\right)^{N_f - \frac{6}{\pi}\frac{\gamma_2^3}{\gamma_3^2}} \exp\left[\frac{6}{\pi}\frac{\gamma_2}{V-\gamma_3}\left(3\gamma_1 + \frac{V}{V-\gamma 3}\frac{\gamma_2^2}{\gamma_3}\right)\right]. \qquad (A.35)$$

The first and the second derivatives of $\ln NCC$ with respect to the volume are also needed, e.g. to compute the microcanonical pressure (see sec. 3.3)

$$\frac{\partial \ln NCC}{\partial V} = \frac{6}{\pi N_f} \frac{1}{V-\gamma_3} \left[\frac{1}{V} \left(\frac{\gamma_2^3}{\gamma_3} - \gamma_0\gamma_3 \right) - \frac{\gamma_2}{V-\gamma_3} \left(3\gamma_1 + \gamma_2^2 \left(\frac{1}{\gamma_3} + \frac{2}{V-\gamma_3} \right) \right) \right]. \quad \text{(A.36)}$$

$$\frac{\partial^2 \ln NCC}{\partial V^2} = \frac{6}{\pi N_f} \frac{1}{V-\gamma_3} \left[-\frac{1}{V} \left(\frac{1}{V} + \frac{1}{V-\gamma_3} \right) \left(\frac{\gamma_2^3}{\gamma_3} - \gamma_0\gamma_3 \right) + 2\frac{\gamma_2}{(V-\gamma_3)^2} \left(3\gamma_1 + \gamma_2^2 \left(\frac{1}{\gamma_3} + \frac{3}{V-\gamma_3} \right) \right) \right]. \quad \text{(A.37)}$$

The volume V_c at which NCC diverges is $V_c = \gamma_3$, which corresponds to a critical packing fraction of 1.

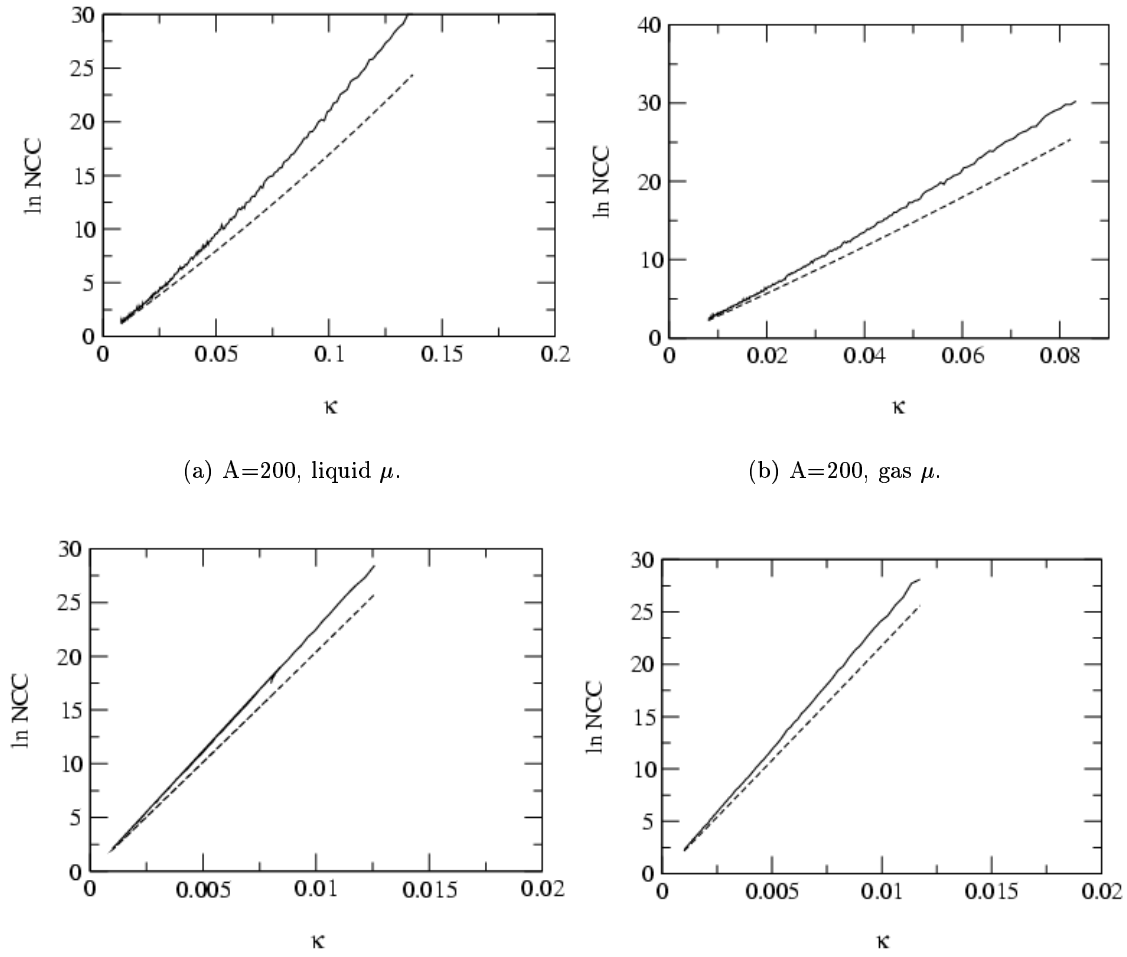
From an algorithmic point of view using eqs. (A.35) to (A.37) is very convenient. They depend only on the γ_k for which the updating scheme after each Monte–Carlo step is straightforward and fast in contrast to the claim of Raduta [RAD01].

On fig. A.2 on the next page are plotted different estimates of NCC using the EOS eq. (A.35) on the facing page (NCC_{EOS}) and the Monte–Carlo sampling presented in the next section NCC_{MC} as a function of the packing fraction $\kappa = \frac{V_0}{V}$. The total masses are $A = 200$ and $A = 1000$. For each A two characteristic mass distributions were taken from MMMC runs; one from the liquid side ($N_{f\,A=200} = 48$, $N_{f\,A=1000} = 194$) and the other from the gas side (near the multifragmentation region; $N_{f\,A=200} = 194$, $N_{f\,A=1000} = 545$). The agreement between the two estimates is good and they are equivalent in the limit $\kappa \to 0$ ($V \to \infty$). This agreement is largely enough for low pressure runs. Moreover what does matter in a Metropolis sampling is the relative differences between the weights of two consecutive states of the Markovian chain. So, even though the ratio NCC_{MC}/NCC_{EOS} might be rather big, one can assume that for a proposed move $c \to c'$ (see app. B.2) $NCC_{EOS}(c)/NCC_{EOS}(c') \approx NCC_{MC}(c)/NCC_{MC}(c')$ (of course this relation does not hold near the critical packing fraction). For intermediate pressures the pressure term due to the avoided volume p_{NCC} cannot be neglected. It is a function of the derivative of $\ln NCC$ with respect to V. One can see in figs. A.2 that the slope of $\ln NCC$ is quite accurately reproduced by NCC_{EOS}.

NCC_{EOS} underestimate NCC since the equation of state is based only on two body correlations.

A.2 Numerical estimates

For large packing fraction the estimate given by eq. (A.35) on the preceding page is not enough accurate. Although there exist (semi–)analytical estimates of Z for large κ (generally close to κ_c) there are usually worked out for very particular mass distributions (mono–modal, bi–modal, and, if at all, with Gaussian dispersions around the maxima), see e.g. [EAG98]. This is in sharp contrast with the need of computing the avoided volume of thousands of *different* mass distributions in one single MMMC run. Moreover these distributions are not always as simple as the one studied in the above mentioned literature. Hence the only way to estimate NCC is to use Monte–Carlo schemes.



(a) A=200, liquid μ.

(b) A=200, gas μ.

(c) A=1000, liquid μ.

(d) A=1000, gas μ.

Figure A.2: Comparison between NCC given by eq. (A.35) on page 114 (dashed lines) and its "exact value" (Monte-Carlo estimates; solid lines) as a function of the packing fraction κ for different total masses A and mass distributions μ (see text).

A.2.1 Simple Monte–Carlo scheme

A straightforward estimate of the ratio eq. (A.1a) on page 109 consists of placing randomly each clusters and check for overlapping. A positioning (event) is considered as successful if all the N_f spheres have been placed successfully. Defining N_s as the number of successful events and N_t as total number of trial events, then an estimate of NCC is given by

$$NCC^{-1} = \frac{N_s}{N_t}. \tag{A.38}$$

But since, on one hand, NCC can be very large (see e.g. figs. A.2 where $NCC \sim 10^{10}$), and, on the other hand, a rather good estimate is needed in order to compute accurately the derivative of NCC with respect to V [a], and since the relative statistical error in evaluating

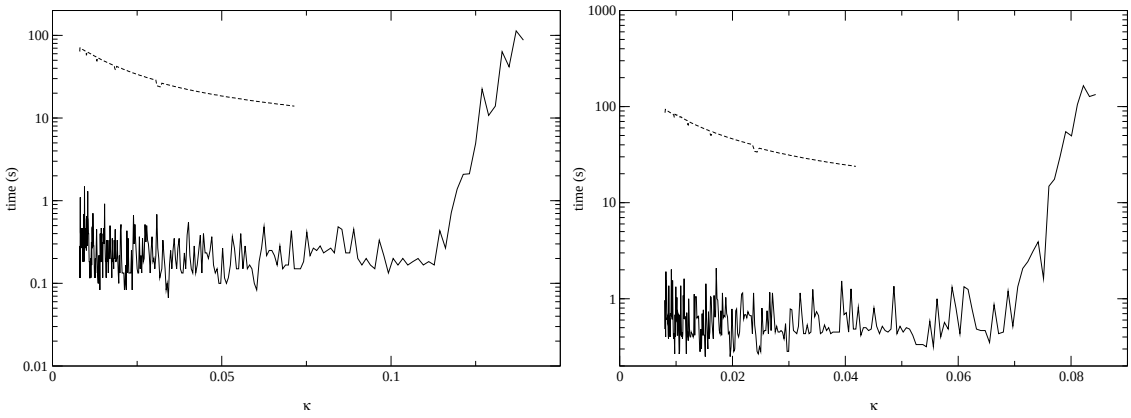
[a] $\frac{\partial NCC}{\partial V}$ and $\frac{\partial^2 NCC}{\partial V^2}$ are needed to estimate the pressure and the inverse microcanonical temperature

NCC^{-1} is of the order of $1/\sqrt{N_t}$, the simplest Monte–Carlo scheme is impractical [RAD01] (see figs. A.2 on the facing page).

A.2.2 Advanced Monte–Carlo scheme

For the simplest Monte–Carlo scheme a failed positioning of cluster k implies a complete resampling of all the $k-1^{\text{th}}$ first "successful" positions. Indeed if one keeps the $k-1^{\text{th}}$ first positions and resamples only the k^{th} one many times one would introduce biases because of the correlations between the different events and therefore an estimate of NCC according to eq. (A.38) on the preceding page would be inaccurate.

Rodgers and Baddley in [RB91] introduced a technic to correct NCC from these correlations. An algorithm based on this technic by avoiding a lot of resampling is faster than the simplest one by almost two orders of magnitude. For example on figs. A.3 are plotted the CPU–time (in seconds) needed to compute NCC for different packing fractions and mass distributions (these are the same μs as in figs. A.2(a) and A.2(b) on the facing page for $A = 200$)



(a) liquid (b) gas

Figure A.3: Comparison of the CPU–time needed to compute NCC, for different mass distributions ($A = 200$, see text) as a function of the packing fraction κ. The dashed lines correspond to the simple Monte–Carlo algorithm (at fixed number of attempts), and the plain lines to the one using the technic presented in [RB91] (at fixed precision).

The results are presented at constant number of trial for the simplest algorithm and at constant precision for the advanced algorithm. The time for the former decreases with decreasing volume because the sampling of the positions for one given trial stops earlier [b] as the volume decreases. At constant precision the CPU–time would have been an increasing function of κ.

at constant pressure β_p, see secs. 3.3 and 3.4, and app. D.

[b] For one given trial, the center of mass are sampled sequentially. The smaller the volume the higher is the probability for this chain to be stopped early.

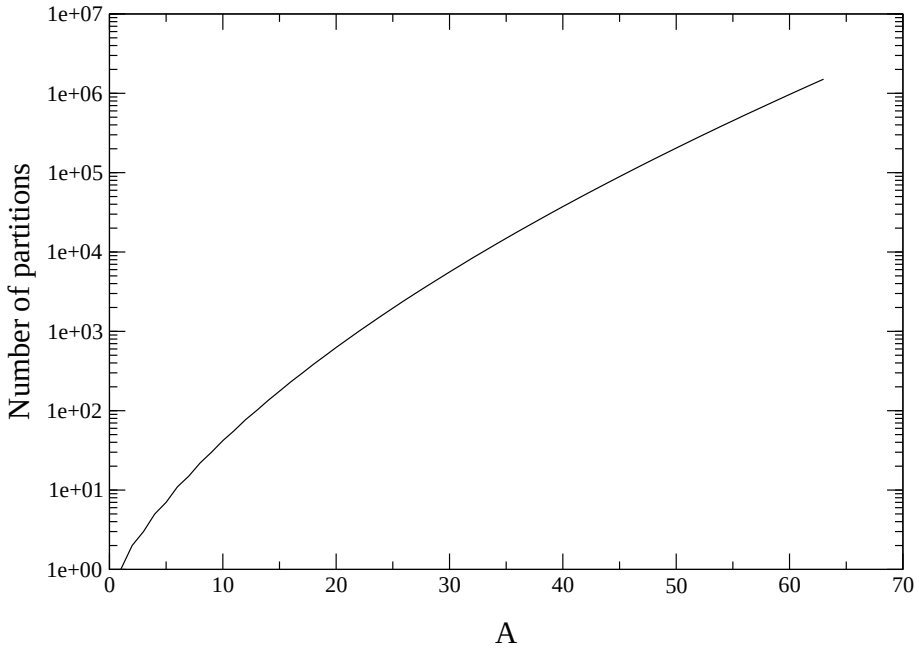


Figure A.4: Number of possible mass distributions as a function of the total mass A.

A.2.3 Further improvements

Using the results from [RB91] leads already to a drastic gain in CPU–time. However a typical run at high pressure still takes too much CPU–time. During a typical MMMC run (one point in the (E, V)–plane) $\sim 2.10^6$ events are generated. For each events $NCC(\mu, V)$, $NCC(\mu, V - \Delta V)$ and $NCC(\mu, V + \Delta V)$ are estimated. By considering that much less precision is needed than the one asked to the data plotted in figs. A.3 on the page before (assuming that the statistical errors would be smoothed out due to the averaging) one ends up with run times of the order of $2.10^6 \times 3 \times 0.05$ seconds ≈ 3 days!

In the following some improvements to the advanced algorithm used in MMMC95 are briefly reviewed.

- Several millions of events (mass distributions) are generated during one MMMC run. Nevertheless there are not all different, on the contrary the value of μ (the mass distribution) usually fluctuates around some mean mass distribution. Therefore after equilibration MMMC has to compute NCCs which have already be computed and that many times. The solution is straightforward: the NCCs are stored and reused whenever needed. The concrete implementation of this simple idea is less straightforward. Indeed a mass distribution is a set of integers each standing for a cluster mass. Moreover the typical number of mass distributions to be stored is huge, see e.g. fig. A.4 where the number of partitions for $A = 60$ is already $\approx 10^6$. Hence a straightforward storage would need a huge amount of RAM space for $A = 200$.

  For simplicity the integers are sorted by decreasing order. There exists no simple direct *and* efficient way to sort these sets. However the storage of sets of sorted

integers is a well–known problem in computing science. It can be solved by using the B–tree technic [BM72, BY89]. Briefly, consider the sets of mass distributions (here $A = 16$) in table A.1. These mass distributions can be represented in a tree-

(a)	8	3	3	2	
(b)	8	3	2	2	1
(c)	8	3	3	1	1
(d)	8	4	4		

Table A.1: Set of mass distributions #1.

form as in fig. A.5. All the sets in tab. A.1 shares the same biggest mass, i.e. the

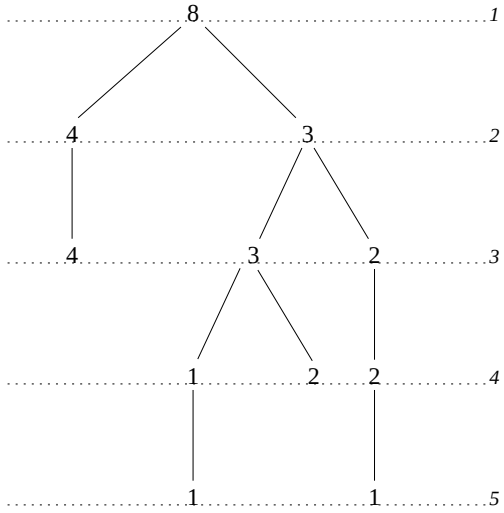


Figure A.5: Representation of the mass distributions of tab. A.1 on a tree. The numbers on the left side represent the height of the knots, i.e. the rank of the integer in the mass distribution.

same root in fig. A.5. The depth of a knot represents the rank of the attached integer in the mass distribution. The NCCs are stored at the level of the leaves (tips of the branches). There are as many leaves as mass distributions.
The advantage of the B–tree storage is clear. The knots might be shared by many mass distributions saving a lot of RAM. The implementation of this kind of storage by preventing from re-evaluating again and again NCCs of the most probable mass distributions saves a large amount of CPU–time (again of an order of magnitude). One can go a little bit further and instead of performing independent runs in the (E, V)–plane, one can use the NCCs at say (E_1, V_1) to perform a run at $E_1 + \Delta E, V_1$, since NCC *does not* depend on the system energy.

- Consider the two mass distributions in table A.2 on the following page; they have the same number of fragments and only their second and sixth clusters are different. Far from the critical packing fraction, it is reasonable to assume that $NCC(a) \approx NCC(b)$ [c]. In other word NCC is a "smooth" function of the mass distribution μ.

[c] This assumption has been verified many times in simulations.

(a)	100	4	3	3	3	2	1
(b)	100	3	3	3	3	3	1

Table A.2: Set of mass distributions #2.

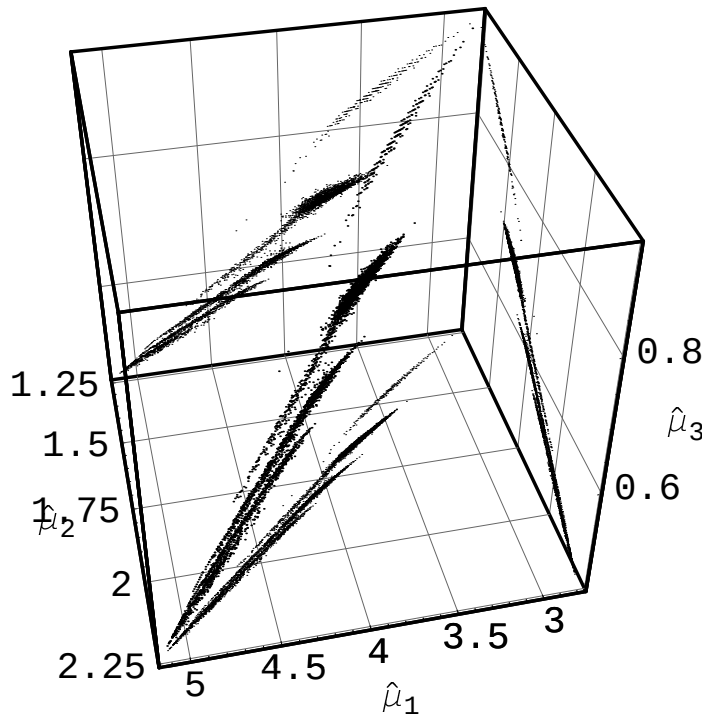


Figure A.6: Mass distributions on the $(\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_4)$ space at fixed $\hat{\mu}_0$, and the projections of the points on the planes $(\hat{\mu}_1, \hat{\mu}_2)$, $(\hat{\mu}_1, \hat{\mu}_4)$ and $(\hat{\mu}_2, \hat{\mu}_4)$. The total mass is $A = 200$ and the number of fragment is $N_f = \hat{\mu}_0 = 12$. This figure does not contain all the possible mass distributions satisfying $\hat{\mu}_0$.

One can use this result to interpolate some NCCs between already known values. The set of integers μ does not provide a good basis for this interpolation, so one has to use some continuous parameterization that reduces the number of coordinates. The parameterization chosen in the present work is based on the set

$$\hat{\mu} = \{\hat{\mu}_0, \hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_4\}, \tag{A.39}$$

where $\hat{\mu}_k \doteq \frac{1}{A^{k/3}} \sum_{i=1}^{N_f} N_i^{k/3}$. $\hat{\mu}_0$ is simply the number of fragments N_f; $\hat{\mu}_3$ is not used because it is a constant of μ, indeed

$$\hat{\mu}_3 \doteq \frac{1}{A} \sum_{i=1}^{N_f} N_i = 1. \tag{A.40}$$

In fig. A.6 each point corresponds to one given μ in the new coordinates $\hat{\mu}$. The total mass and the number of fragments (clusters) are resp. $A = 200$ and $N_f = 12 = \hat{\mu}_0$. Large values of $\hat{\mu}_1$, $\hat{\mu}_2$ and $\hat{\mu}_4$ correspond to a monodisperse mass distribution (twelve

clusters of mass $\approx \frac{200}{12}$), and small values to μ with one big clusters and eleven monomers.

As one can see in fig. A.6 that the positions of the points are strongly correlated. The coordinates (A.39) on the preceding page are largely sufficient for interpolation (NCC itself is a smooth function of $\hat{\mu}$). The storage of the $\hat{\mu}$ is technically cumbersome (because the coordinates are not discrete and the density of points is not constant), but again it can be solved using some B–tree algorithms.

Appendix B

Technical "details"

B.1 Introduction

The main method used for numerical applications in this thesis is based on the Metropolis sampling [MRR+53]: one builds a Markovian chain, i.e. from a state c a new state c' is sampled. This new state is accepted with a probability transition $P(c \to c')$ given by

$$P(c \to c') = \min\left(1, \frac{\omega(c')}{\omega(c)}\right), \tag{B.1}$$

where $\omega(c)$ is the statistical weight of c. This probability transition satisfies the detailed balance equation which is the cornerstone of the Metropolis sampling

$$p(c \to c')\,\omega(c) = p(c' \to c)\,\omega(c'). \tag{B.2}$$

The mean value of an observable F is given by

$$\langle F \rangle = \frac{1}{N}\sum_{j=1}^{N} F(c_j). \tag{B.3}$$

Eq. (B.1) is the simplest form for $P(c \to c')$; for practical reasons (either lack or a priori information) eq. (B.1) or eq. (B.3) can or have to be modified (see below).

B.2 Monte–Carlo sampling

A priori probability

There is an implicit assumption made from eq. (B.2) to eq. (B.1); namely the *a priori probability* $\mathcal{A}(c \to c')$ to sample c' "from" c satisfies [Kra98]

$$\mathcal{A}(c \to c') = \mathcal{A}(c' \to c). \tag{B.4}$$

Eq. (B.4) is not necessarily always satisfied. Hence it is sometimes technically difficult to ensure that the move $c \to c'$ is chosen with the same a priori probability than $c' \to c$. One can also force the Markovian chain to go in a given direction in the parametric space by using some a priori knowledge.

Now the detailed balance equation has to be reevaluated and explicitely written with the a priori probability. The probability $p\left(c \to c'\right)$ is split up into two separate parts

$$p\left(c \to c'\right) = \mathcal{A}\left(c \to c'\right) P\left(c \to c'\right), \tag{B.5}$$

where $P\left(c \to c'\right)$ is the acceptance probability of the move proposed with $\mathcal{A}\left(c \to c'\right)$. The full detailed balance is

$$\mathcal{A}\left(c \to c'\right) P\left(c \to c'\right) \omega(c) = \mathcal{A}\left(c' \to c\right) P\left(c' \to c\right) \omega(c'). \tag{B.6}$$

Now, one form for the acceptance probability is

$$P\left(c \to c'\right) = \min\left(1, \frac{\omega(c')}{\omega(c)} \frac{\mathcal{A}\left(c' \to c\right)}{\mathcal{A}\left(c \to c'\right)}\right). \tag{B.7}$$

mmmc95

In MMMC95 there is only one move, namely

$$\underbrace{\left(M, E_M^*\right) + \left(N, E_N^*\right)}_{c} \to \underbrace{\left(P, E_P^*\right) + \left(Q, E_Q^*\right)}_{c'} \tag{B.8}$$

where M, N, P and Q are clusters, their mass are respectively indicated by their name, $M + N = P + Q$. E_M^*, E_N^*, E_P^* and E_Q^* are their respective internal excitation energy. $M \in \mu(c)$, $N \in \mu(c')$, $N \in \{0\} \cup \mu(c)$ and $Q \in \{0\} \cup \mu(c')$, where "0" is a "virtual" vacuum fragment. It has no mass and no excitation energy and does not contribute to the microcanonical weight factor eq. (3.8) on page 47.

This move spans all the ones used in MMMC77 [GRO97], e.g. "split fragment" is now

$$\left(M, E_M^*\right) + 0 \to \left(P, E_P^*\right) + \left(Q, E_Q^*\right),$$

but also moves that were not proposed (implying a breaking of the detailed balance condition, see sec. 3.4.1), e.g.

$$(1, 0) + (1, 0) \to (2, 0). \tag{B.9}$$

A move is performed in two steps. First a new mass distribution is generated then the excitation energies (E_P^*, E_Q^*) are sampled.

B.2.1 Partitioning factor

As in this subsection mass distributions are only considered, the internal excitations energies of eq. (B.8) can be omitted

$$\underbrace{M + N}_{c} \to \underbrace{P + Q}_{c'}. \tag{B.10}$$

The only weight that might be changed by (B.10) is the partitioning weight $\omega_{sym} = \frac{1}{\prod_{i=1}^{A} \xi_c(i)!}$, where $\xi_c(i)$ is the number of cluster of size i in the event c. In order to use eq. (B.1) on the page before in the numerical code eq. (B.4) must be satisfied which is an impossibly difficult task. So one must use eq. (B.7) where the a priori probability is explicitely taken into account

$$P\left(c \to c'\right) = \min\left(1, \frac{\omega_{sym}(c')}{\omega(_{sym}c)} \frac{\mathcal{A}\left(c' \to c\right)}{\mathcal{A}\left(c \to c'\right)}\right). \tag{B.11}$$

In its turn $\mathcal{A}(c \to c')$ is split up into two parts, i.e. $\mathcal{A}(c \to c') = p(M,N)p(\to P,Q|M,N)$, where $p(M,N)$ is the probability to choose M and N and $p(\to P,Q|M,N)$ is the probability to sample P and Q once M and N are chosen.

In order to increase the pass acceptance and to decrease the correlation time, M is chosen among the clusters in $\mu(c)$ with a probability proportional to its mass. N is chosen among the clusters in $\{0\} \cup \mu(c) \setminus M$ in an equiprobable way, i.e.

$$p(M,0) = \frac{M\xi_c(M)}{A}\frac{1}{N_f} = \frac{M\xi_c(M)}{AN_f} \tag{B.12a}$$

$$p(M,N \neq M) = \frac{M\xi_c(M)}{A}\frac{\xi_c(N)}{N_f} = \frac{M\xi_c(M)\xi_c(N)}{A\,N_f} \tag{B.12b}$$

$$p(M,M) = \frac{M\xi_c(M)}{A}\frac{\xi_c(M)-1}{N_f} = \frac{M\xi_c(M)\,(\xi_c(M)-1)}{A\,N_f} \tag{B.12c}$$

Eq. (B.12a) correspond to the case when a "real" and the vacuum fragments are chosen. Eq. (B.12b) is the general case when two fragments of different masses are chosen whereas in Eq. (B.12c) both fragments have the same mass.

$M+N$, $M \geq N$ are repartitioned in the following way. An integer number i is chosen in the range $\left[-N, \text{floor}\,\frac{M+N}{2} - N\right]$, where floor x is the biggest integer less or equal to x, with a probability proportional to $\frac{1}{|i-N|+1}$. The new mass distribution is given by

$$Q = N + i, \tag{B.13a}$$

$$P = M + N - Q, \tag{B.13b}$$

for an illustration see fig. B.1 on the following page. This repartitioning favors small changes in the mass distribution. Moreover it is easy to verify that it ensures $p(\to P,Q|M,N) = p(\to M,N|Q,P)$. Thus, in eq. (B.11) only the ratio $\frac{p(P,Q)}{p(M,N)}$ has to be estimated.

Now everything can be collected in order to compute the probability transition. As an example, let us consider the move

$$M + N \to P + Q,$$

with $M \neq N$, $M \neq P$, $N \neq Q$, $Q \neq 0$ and $N \neq 0$, i.e. no vacuum fragment is involved and all the fragments have different masses. The second argument in the min function in eq. (B.7) on the preceding page becomes

$$\frac{\omega_{sym}(c')}{\omega(_{sym}(c)}\frac{\mathcal{A}(c' \to c)}{\mathcal{A}(c \to c')} = \frac{\xi_c(M)!\xi_c(N)!\xi_c(P)!\xi_c(Q)!}{\xi_{c'}(M)!\xi_{c'}(N)!\xi_{c'}(P)!\xi_{c'}(Q)!}\frac{P\xi_{c'}(P)\xi_{c'}(Q)}{M\xi_c(M)\xi_c(N)} \tag{B.14}$$

$$= \frac{P}{M}\frac{(\xi_c(M)-1)!\,(\xi_c(N)-1)!\xi_c(P)!\xi_c(Q)!}{\xi_{c'}(M)!\xi_{c'}(N)!\,(\xi_{c'}(P)-1)!\,(\xi_{c'}(Q)-1)!} \tag{B.15}$$

but the proposed move implies that $\xi_{c'}(M) = \xi_c(M) - 1$, $\xi_{c'}(N) = \xi_c(N) - 1$, $\xi_c(P) = \xi_{c'}(P) - 1$ and $\xi_c(Q) = \xi_{c'}(Q) - 1$, and eq. (B.15) becomes

$$\frac{\omega_{sym}(c')}{\omega_{sym}(c)}\frac{\mathcal{A}(c' \to c)}{\mathcal{A}(c \to c')} = \frac{P}{M}. \tag{B.16}$$

With this sampling, the symmetrization weight factor is *exactly* taken into account.

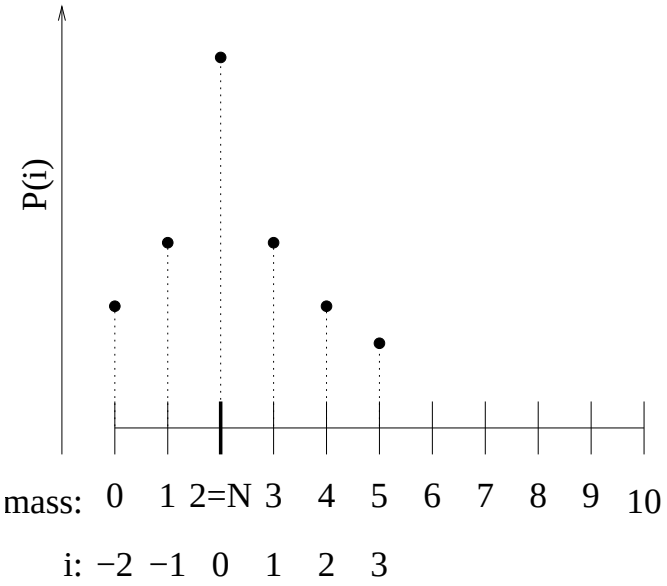


Figure B.1: Illustration of the repartitioning procedure. The initial cluster sizes are $M = 8$ and $N = 2$. The mass of the new smallest fragment Q is given by $Q = N + i = 2 + i$ where i is a number in the range $\left[-N, \text{floor}\,\frac{M+N}{2} - N\right] = [-2, 3]$, chosen with a probability proportional to $\frac{1}{|i-N|+1}$.

Example

The simple case of a small system with $A = 6$ is considered. All the possible partitions are listed in table B.2.1.0 on the facing page along with their respective ω_{sym} which can be easily computed.

Using the values in tab. B.2.1.0 one can compute the mean values of several observables

- $< N_{fr} >$mean number of fragments;
- $< M_1 >$, $< M_2 >$and $< M_3 >$, mean size of the first, second and third biggest fragment;
- $< N >$, probability that a fragment of size N is present in an event (excluding monomers),

and compare them with the results given by MMMC95. The results[a] are listed in table B.2.

The numerical result are in very good agreement with the analytical results. For MMMC77 the results are really bad for such a small system total mass since the algorithm used to estimate the partitioning weight in MMMC77 is valid for large system total mass A [zZG93].

B.2.2 Excitation energy sampling

The algorithm for the excitation energy sampling has been developed having in mind the following constraints

[a]The results are the averages over 10^6 events.

Configurations	ω_{sym}
6	1
$5+1$	1
$4+2$	1
$4+1+1$	$1/2! = 1/2$
$3+3$	$1/2! = 1/2$
$3+2+1$	1
$3+1+1+1$	$1/3! = 1/6$
$2+2+2$	$1/3! = 1/6$
$2+2+1+1$	$1/(2!2!) = 1/4$
$2+1+1+1+1$	$1/4! = 1/24$
$1+1+1+1+1+1$	$1/6! = 1/720$

Table B.1: List of all the possible mass distributions for $A = 6$ along with their respective symmetrization weight ω_{sym}.

quantities evaluated	analytical results	MMMC95
$< N_{fr} >$	$9276/4051 \simeq 2.290$	2.291
$< M_1 >$	$16501/4051 \simeq 4.073$	4.071
$< M_2 >$	$5791/4051 \simeq 1.430$	1.431
$< M_3 >$	$1651/720 \simeq 0.4076$	0.4080
$< 6 >$	$24/209 \simeq 0.1148$	0.1147
$< 5 >$	$24/209 \simeq 0.1148$	0.1144
$< 4 >$	$36/209 \simeq 0.1722$	0.1724
$< 3 >$	$52/209 \simeq 0.2488$	0.2488
$< 2 >$	$73/209 \simeq 0.3493$	0.3498

Table B.2: Comparison of analytical and numerical results for different observables.

1. it must fulfill the detailed balance equation and therefore be reversible,
2. the new sampled excitation energies must as much as possible lead to a positive remaining energy eq. (3.5) on page 45. This is particularly constraining at small total energy, where, without this constraint nearly 60% of the proposed moves are rejected only because of negative remaining energy.
3. it should favor small steps, e.g, consider $M \to P + Q$ with $P >> Q$, then the algorithm should support $E^*_P \sim E^*_M$.

To simplify let us consider the following move

$$\underbrace{(M, E^*_M) + (N, E^*_N)}_{c} \to \underbrace{(P, E^*_P) + \big(Q, E^*_Q\big)}_{c'}, \tag{B.17}$$

with $P \geq Q \geq 3$, $M \geq N \geq 3$. The remaining (kinetic) energies in c and c' are

$$E_k(c) = \bar{E}(c) - E^*_M - E^*_N - E_{b\,M} - E_{b\,N}, \tag{B.18a}$$
$$E_k(c') = \bar{E}(c') - E^*_P - E^*_Q - E_{b\,P} - E_{bQ}, \tag{B.18b}$$

where E_b stands for the binding energies.

During the process (B.17) $\bar{E}$ is conserved, i.e. $\bar{E}(c) = \bar{E}(c')$. The first step is to check whether there is enough energy to create c', i.e if $E_k(c') = \bar{E}(c) - E_{bP} - E_{bQ} > 0$.

Now let us rewrite $E_i^* = \alpha_i \tilde{E}_{max,\,i}^*$, $i = M$, N, P and Q, where $\alpha_i \in]0, 1]$ and $\tilde{E}_{max,\,i}^*$ is minimum between

a. $E_{maxx,\,i}^* = (i-2)\epsilon_{max}^*$ the maximal allowed excitation energy for clusters i (see sect. 3.2 on page 40),

b. the available excitation energy for the move Eq. (B.17) (constraint #2).

The energies are sampled sequentially. First P (conversely M) therefore

$$\tilde{E}_{max,\,P}^* = \min\left((P-2)\epsilon_{max}^*, \bar{E}(c) - E_{b\,P} - E_{b\,Q}\right) \tag{B.19a}$$
$$\tilde{E}_{max,\,M}^* = \min\left((M-2)\epsilon_{max}^*, \bar{E}(c) - E_{b\,M} - E_{b\,N}\right) \tag{B.19b}$$

than Q (conversely N)

$$\tilde{E}_{max,\,Q}^* = \min\left((Q-2)\epsilon_{max}^*, \bar{E}(c) - E_{b\,P} - E_{b\,Q} - E_P^*\right) \tag{B.20a}$$
$$\tilde{E}_{max,\,N}^* = \min\left((N-2)\epsilon_{max}^*.\bar{E}(c) - E_{b\,M} - E_{b\,N} - E_M^*\right) \tag{B.20b}$$

Considering the move $c \to c'$, α_M and α_N are known. Their values are both shifted

$$\begin{aligned} \alpha_1 &= \mod\ (\alpha_M + \Delta\alpha_1) \\ \alpha_2 &= \mod\ (\alpha_M + \Delta\alpha_2)\,, \end{aligned}$$

where $\Delta\alpha_1$ and $\Delta\alpha_2$ are two random numbers sampled in the range $[-0.05, 0.05]$. The new α_P and α_Q are simply either $(\alpha_P, \alpha_q) = (\alpha_1, \alpha_2)$ or $(\alpha_P, \alpha_q) = (\alpha_2, \alpha_1)$, each case has a probability of $1/2$.

Now a computation of the a priori probability ratio yields

$$\frac{\mathcal{A}(c \to c')}{\mathcal{A}(c' \to c)} = \frac{\tilde{E}_{max,\,P}^*\,\tilde{E}_{max,\,Q}^*}{\tilde{E}_{max,\,M}^*\,\tilde{E}_{max,\,N}^*}. \tag{B.21}$$

B.3 Multicanonical algorithm

In this section the blocking mechanism used to estimate the Bg function for the self gravitating system studied in part III (eq. (6.21) on page 92) is presented. The estimate of Bg is obtained by an algorithm based is based on multicanonical technics [Bin97, Ber96, Lee93, FS89].

The usual multicanonical task is to compute the free energy as a function of the total energy [Par01]. For the gravitational system presented in part III one has to compute Bg as a function of the inertial momentum I and of the potential energy ϕ. The updating scheme presented in [Smi96] is used. One of the reason for this choice is that although it has been given for a one dimensional task it can be trivially extended to bi–variate problems. The other reason is that it is one of the few algorithms to give and use information about the statistical errors on the estimate of the weight Bg (for another recent algorithm see [Bor01]).

In a multicanonical scheme $Bg \equiv W$ is built iteratively. To improve the performances of the algorithm blocking mechanism has been added. After an iteration. if it is estimated

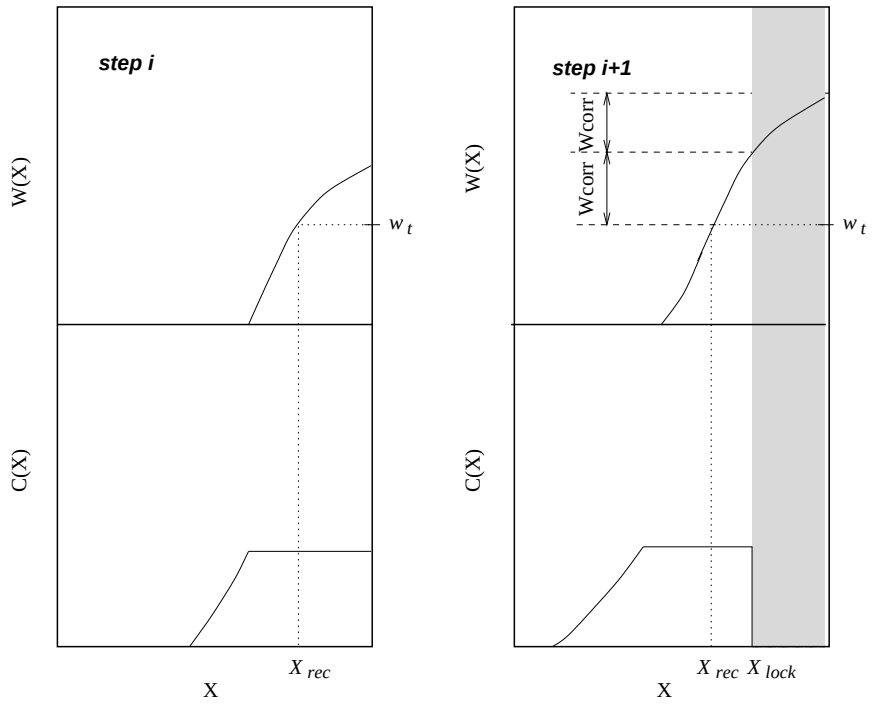


Figure B.2: Schematic illustration of the blocking mechanism. The visited state histogram $C(X)$ is plotted for the iteration steps i and $i+1$, the plotted weight $W(X) \equiv Bg(X)$ is the one obtained *after* the iterations i and $i+1$, respectively. The region where $W > W_t$ are tagged to be locked during the next iterations. W_t is a suitable threshold weight. The consequences of the locking prossecus can be seen in $C_{i+1}(X)$ where $C_{i+1}(X) = 0$ for $X > X_{i+1\,lock}$. The weight $W_{i+1}(X > X_{i+1\,lock})$ is corrected simply by $W_{i+1}(X > X_{i+1\,lock}) = W_i(X > X_{i+1\,lock}) + W_{corr}$ where $W_{corr} = \frac{1}{\Delta X} \int_{X_{i+1\,lock}-\Delta X}^{X_{i+1\,lock}} W_i(X)dX$, where ΔX is a suitable positive constant.

that enough information has been collected on a given region of the parametric space (I, ϕ) then this region is tagged as "locked" so that it will not be visited during next iterations (see fig. B.2). This mechanism enables the program to spread more quickly over the parametric space and save computation time compared to usual multicanonical algorithms.

Figure B.3 on the following page shows a slice of $Bg(I, \phi)$ for $I = 3$ at different iteration steps i for the gravitational system and $N = 20$ (Fig. B.3(a)). The histogram $C(I = 3, \phi)$ of the visited region is also plotted in order to illustrate the blocking mechanism (Fig. B.3(b)). As expected Bg is strongly peaked around the disordered region $\phi \approx -1$ (this value correspond to the mean of ϕ over randomly generated spatial configurations). After 10 iterations the ratio between the maximum and the minimum of Bg is $\approx \exp 120$. This ratio increases exponentially with N, e.g. at $N = 10$ its value is $\approx \exp 80$.

In fig. B.3 the final estimate of Bg is shown. The CPU–time to compute the whole Bg but also the observables used in chap. 6 (radial distribution, distance distribution) is around 40 hours on an ALPHA-workstation.

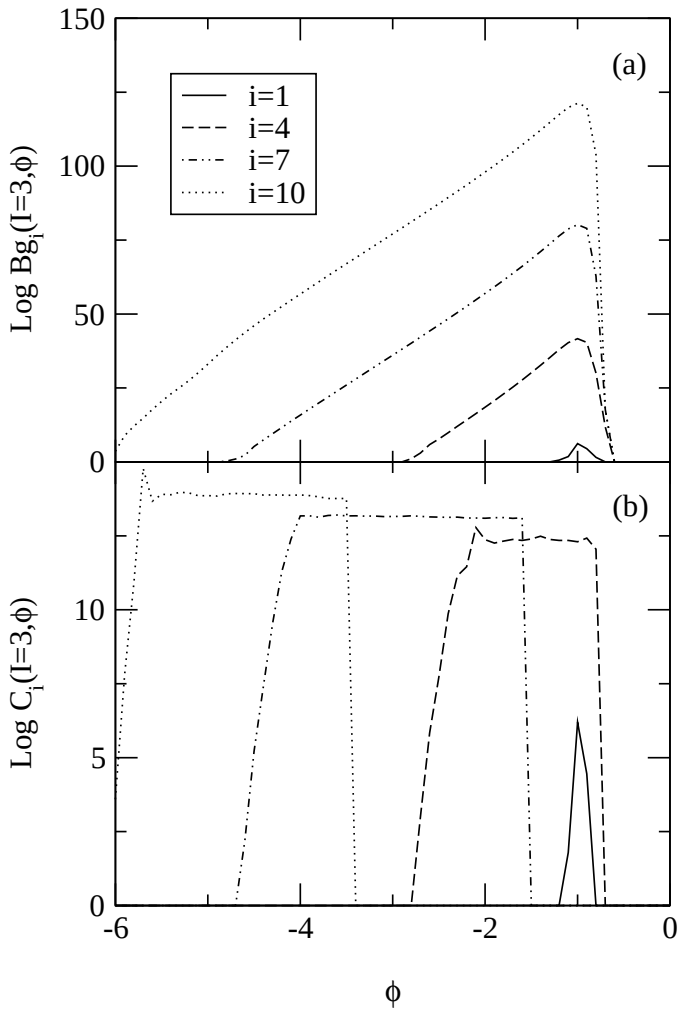


Figure B.3: Estimate of the density of state Bg, panel (a), and histogram of the visited states C, panel (b), for $I = 3$ at different iteration steps i of the multicanonical algorithm as a function of the potential energy ϕ. Panel (a) shows how Bg is built step by step. Bg is an extremely peaked function, the log of the ratio between its maximum and its minimum is about 120. Without the blocking mechanism (see text) C_i would have been non null for all value of ϕ visited during previous steps $j < i$. In panel (b) one sees that the algorithm does no longer visit "well–known" regions ($\phi \gtrsim -1.5$) already after four steps.

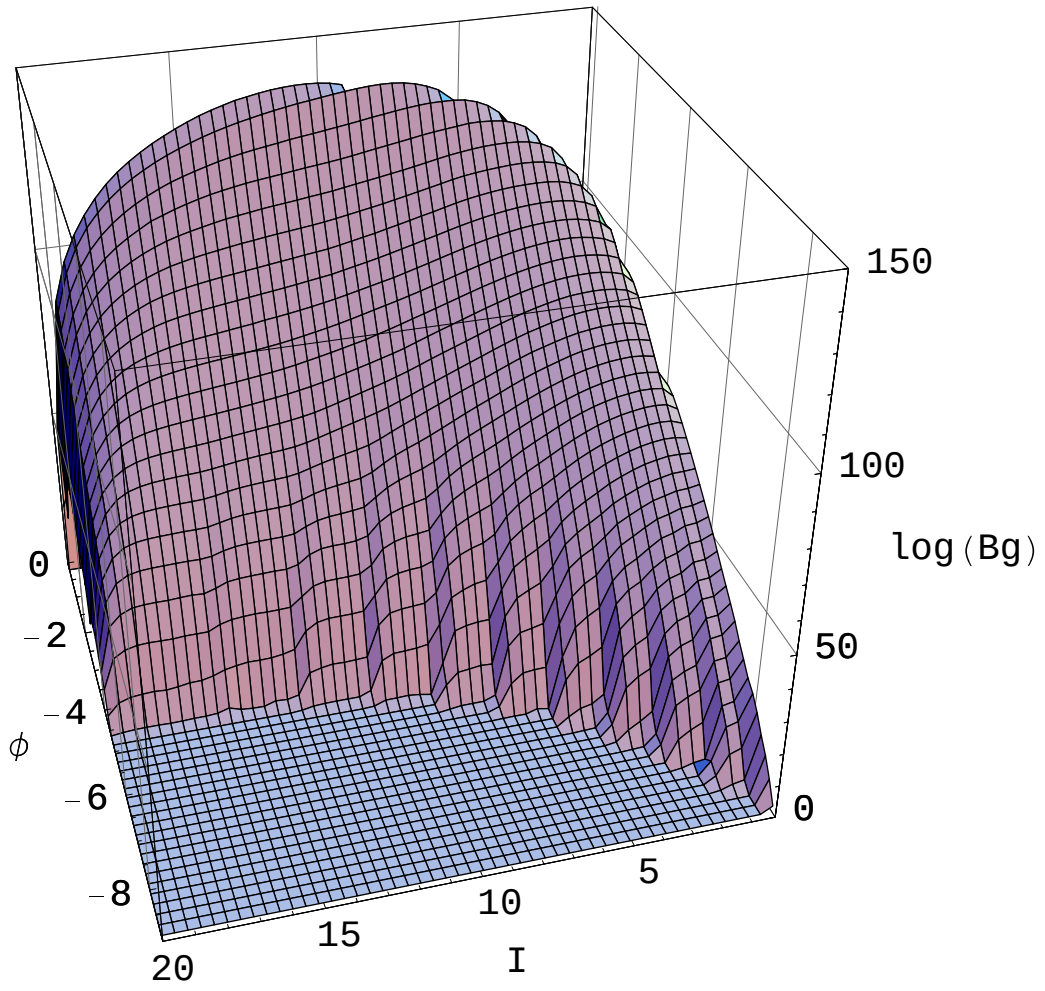


Figure B.4: Final estimate of Bg for $N = 20$ as a function of the inertial momentum I and the potential energy ϕ.

Appendix C

Momentum distribution

In this appendix $< \mathbf{p}_k >_{\mathbf{q}_k}$ the average momentum of particle k at fixed position is computed.

For simplicity k is set to 1. The α-component of $< \mathbf{p}_1 >_{\mathbf{q}_1}$ is

$$\begin{aligned} < p_1^\alpha >_{\mathbf{q}_1} &= \frac{\int \left(\prod_i d\mathbf{p}_i \prod_{i=2}^N d\mathbf{q}_i\right) p_1^\alpha \delta\left(E-\mathcal{H}\right) \delta^2\left(\sum_i \mathbf{p}_i\right) \delta\left(\sum_i \mathbf{q}_i \times \mathbf{p}_i - L\right) \delta^2\left(\sum_i \mathbf{q}_i\right)}{\int \left(\prod_i d\mathbf{p}_i \prod_{i=2}^N d\mathbf{q}_i\right) \delta\left(E-\mathcal{H}\right) \delta^2\left(\sum_i \mathbf{p}_i\right) \delta\left(\sum_i \mathbf{q}_i \times \mathbf{p}_i - L\right) \delta^2\left(\sum_i \mathbf{q}_i\right)} \\ &= \frac{\int \left(\prod_{i=2}^N d\mathbf{q}_i\right) \mathcal{P}_1^\alpha \delta^2\left(\sum_i \mathbf{q}_i\right)}{\int \left(\prod_{i=2}^N d\mathbf{q}_i\right) W(\mathbf{r}) \delta^2\left(\sum_i \mathbf{q}_i\right)}, \end{aligned} \tag{C.1}$$

where $\mathcal{P}_1^\alpha\left(E, L, \{\mathbf{q}\}\right) = \int \left(\prod_i d\mathbf{p}_i\right) p_1^\alpha \delta\left(E-\mathcal{H}\right) \delta^2\left(\sum_i \mathbf{p}_i\right) \delta\left(\sum_i \mathbf{q}_i \times \mathbf{p}_i - L\right)$, $\{\mathbf{q}\}$ is a short hand for $\{\mathbf{q}_1, \ldots, \mathbf{q}_N\}$ and $W(\mathbf{q})$ is the microcanonical weight at fixed spatial configuration $\{\mathbf{q}\}$, its value is $W(E, L, \{\mathbf{q}\}) = \mathcal{C} \frac{1}{\sqrt{I}} E_r^{N-5/2}$, where $\mathcal{C} = \frac{\prod_i m_i}{M} \frac{1}{(2\pi)^{N+9/2}} \frac{1}{\Gamma(N-3/2)}$ (see eq. (1.11) on page 7). $\mathcal{H} = \sum_i \frac{\mathbf{p}_i^2}{2m_i} - \phi(\mathbf{q})$ is the Hamiltonian where $\phi(\mathbf{q})$ is the potential.

The outline of the derivation of $\mathcal{P}_1^\alpha$ is the same as in [Lal99] for W.

First $\mathcal{P}_1^\alpha$ is Laplace transformed $E' = E + \phi \rightsquigarrow s$

$$\tilde{\mathcal{P}}_1^\alpha\left(s, L, \{\mathbf{q}\}\right) = \int_o^\infty dE' e^{-sE'} \mathcal{P}_1^\alpha\left(E, L, \{\mathbf{q}\}\right), \tag{C.2}$$

$$= \int \prod_i d\mathbf{p}_i \, p_1^\alpha \, \exp\Big\{ - s \sum_i \frac{\mathbf{p}_i^2}{2m_i} \Big\} \delta^2\left(\sum_i \mathbf{p}_i\right) \delta\left(\sum_i \mathbf{q}_i \times \mathbf{p}_i - L\right). \tag{C.3}$$

Using the integral form of the delta Dirac

$$\delta(x) = \int_{-\infty}^\infty \frac{dw}{2\pi} \exp\Big\{ iw \cdot x \Big\}, \tag{C.4}$$

for the conservation of the linear and angular momenta in eq. (C.3) yields

$$\tilde{\mathcal{P}}_1^{\alpha}(s,L,\{\mathbf{q}\}) = \int \prod_i d\mathbf{p}_i \frac{d\mathbf{w}_1}{(2\pi)^2}\frac{dw_2}{2\pi}\, p_1^{\alpha} \exp\Big\{ -s\sum_i \frac{\mathbf{p}_i^2}{2m_i} + i\mathbf{w}_1\cdot\sum_i \mathbf{p}_i + iw_2\sum_i \mathbf{q}_i\times\mathbf{p}_i - iw_2 L\Big\}, \quad \text{(C.5)}$$

where $\mathbf{w}_1 = (w_1^1, w_1^2)$ is a two dimensional vector. Now one can write explicitly the vectors components in eq. (C.5)

$$\tilde{\mathcal{P}}_1^{\alpha}(s,L,\{\mathbf{q}\}) = \int \prod_i d\mathbf{p}_i \frac{d\mathbf{w}_1}{(2\pi)^2}\frac{dw_2}{2\pi}\, p_1^{\alpha} \exp\Big\{ -s\sum_i\sum_{\beta=1}^{2} \frac{(p_i^{\beta})^2}{2m_i} + i\sum_{\beta=1}^{2} w_1^{\beta}\cdot\sum_i p_i^{\beta} + iw_2\sum_i\sum_{\gamma,\beta=1}^{2}\epsilon_{\gamma\beta}q_i^{\gamma}p_i^{\beta} - iw_2 L\Big\}, \quad \text{(C.6)}$$

where ϵ is the antisymmetric tensor of rank 2.

The integration I_1 over $\{p_i^{\beta}\}$ $i = 1,\ldots,N$ with $\beta \neq \alpha$ gives

$$\begin{aligned} I_1 &= \int \prod_i dp_i^{\beta} \exp\Big\{ -s\sum_i \frac{(p_i^{\beta})^2}{2m_i} + i\sum_i w_1^{\beta}p_i^{\beta} + i\sum_i w_2\sum_{\gamma}^{2}\epsilon_{\gamma\beta}q_i^{\gamma}p_i^{\beta}\Big\} \\ &= \frac{\left(\prod_i m_i\right)^{1/2}}{(2\pi s)^{N/2}} \exp\Big\{ -\frac{1}{2s}\sum_i m_i\left(w_1^{\beta} + w_2\sum_{\gamma=1}^{2}\epsilon_{\gamma\beta}q_i^{\gamma}\right)^2\Big\}. \end{aligned} \quad \text{(C.7)}$$

The integration I_2 over $\{p_i^{\alpha}\}$ $i = 2,\ldots,N$ gives

$$\begin{aligned} I_2 &= \int \prod_{i=2}^{N} dp_i^{\alpha} \exp\Big\{ -s\sum_{i=2}^{N} \frac{(p_i^{\alpha})^2}{2m_i} + i\sum_{i=2}^{N} w_1^{\alpha}p_i^{\alpha} + i\sum_{i=2}^{N} w_2\sum_{\gamma}^{2}\epsilon_{\gamma\alpha}q_i^{\gamma}p_i^{\alpha}\Big\} \\ &= \frac{\left(\prod_{i=2}^{N} m_i\right)^{1/2}}{(2\pi s)^{(N-1)/2}} \exp\Big\{ -\frac{1}{2s}\sum_{i=2}^{N} m_i\left(w_1^{\alpha} + w_2\sum_{\gamma=1}^{2}\epsilon_{\gamma\alpha}q_i^{\gamma}\right)^2\Big\}. \end{aligned} \quad \text{(C.8)}$$

Finally an integration over p_1^{α} yields

$$\begin{aligned} I_3 &= \int dp_1^{\alpha}\, p_1^{\alpha}\Big\{ -s\frac{(p_1^{\alpha})^2}{2m_1} + iw_1^{\alpha}p_1^{\alpha} + iw_2\sum_{\gamma}^{2}\epsilon_{\gamma\alpha}q_1^{\gamma}p_1^{\alpha}\Big\} \\ &= \frac{(m_1)^{1/2}}{(2\pi s)^{1/2}}\Big\{ -\frac{im_1}{s}\Big[w_1^{\alpha} + w_2\sum_{\gamma=1}^{2}\epsilon_{\gamma\alpha}q_1^{\gamma}\Big]\Big\} \\ &\quad \exp\Big\{ -\frac{m_1}{2s}\left(w_1^{\alpha} + w_2\sum_{\gamma=1}^{2}\epsilon_{\gamma\alpha}q_1^{\gamma}\right)^2\Big\} \end{aligned} \quad \text{(C.9)}$$

One collects the results from eq. (C.7) to eq.(C.9)

$$\tilde{\mathcal{P}}_1^{\alpha}(s,L,\{\mathbf{q}\}) = \int \frac{d\mathbf{w}_1}{(2\pi)^2}\frac{dw_2}{2\pi}\frac{\prod_i m_i}{(2\pi s)^N}\Big\{ -\frac{im_1}{s}\Big[w_1^{\alpha} + w_2\sum_{\gamma=1}^{2}\epsilon_{\gamma\alpha}q_1^{\gamma}\Big]\Big\}$$
$$\exp\Big\{ -\frac{1}{2s}\sum_i\sum_{\beta=1}^{2} m_i\left(w_1^{\beta} + w_2\sum_{\gamma=1}^{2}\epsilon_{\gamma\beta}q_i^{\gamma}\right)^2 - iw_2L\Big\}. \quad \text{(C.10)}$$

The first term in the argument of the exponential in eq. (C.10) can be expanded and simplified (using $\sum_i \mathbf{q}_i = 0$)

$$\tilde{\mathcal{P}}_1^{\alpha}(s,L,\{\mathbf{q}\}) = \int \frac{d\mathbf{w}_1}{(2\pi)^2}\frac{dw_2}{2\pi}\frac{\prod_i m_i}{(2\pi s)^N}\Big\{ -\frac{im_1}{s}\Big[w_1^{\alpha} + w_2\sum_{\gamma=1}^{2}\epsilon_{\gamma\alpha}q_1^{\gamma}\Big]\Big\}$$
$$\exp\Big\{ -\frac{M}{2s}\sum_{\beta=1}^{2}(w_1^{\beta})^2 - \frac{w_2 I}{2s} - iw_2L\Big\}, \quad \text{(C.11)}$$

where $M = \sum_i m_i$, and $I = \sum_i m_i\mathbf{q}_i^2$.

$\tilde{\mathcal{P}}_1^{\alpha}(s,L,\{\mathbf{q}\})$ is the sum of two multiple integrals. The one which contains an argument proportional to $w_1^{\alpha}\exp((w_1^{\alpha})^2)$ is null since this argument is an odd function of w_1^{α}. The remaining multiple Gaussian integrals over $\mathbf{w}_1$ and w_2 can straightforwardly be computed, and the result is

$$\tilde{\mathcal{P}}_1^{\alpha}(s,L,\{\mathbf{q}\}) = \frac{1}{(2\pi)^4}\frac{L}{MI}\frac{\prod_i m_i}{(2\pi)^N}\frac{\sum_{\gamma=1}^{2}\epsilon_{\gamma\alpha}q_1^{\gamma}}{s^{N-3/2}\sqrt{2\pi I}}\exp\Big\{ -s\frac{L^2}{2I}\Big\}. \quad \text{(C.12)}$$

The inverse Laplace transform of $\tilde{\mathcal{P}}_1^{\alpha}(s,L,\{\mathbf{q}\})$ gives [EMOT54]

$$\mathcal{P}_1^{\alpha} = \mathcal{C}Lm_1 I^{-3/2}\sum_{\gamma=1}^{2} q_1^{\gamma}\epsilon_{\gamma\alpha}\, E_r^{N-5/2}, \quad \text{(C.13)}$$

if $E_r > 0$, where $\mathcal{C} = \frac{\prod_i m_i}{M}\frac{1}{(2\pi)^{N+9/2}}\frac{1}{\Gamma(N-3/2)}$ and $E_r = E' - \frac{L^2}{2I} = E - \frac{L^2}{2I} - \phi(\{\mathbf{q}\})$. Using (C.13) in (C.1) one gets finally

$$\begin{aligned} < p_1^{\alpha} >_{\mathbf{q}_1} &= \frac{\int\left(\prod_{i=2}^{N} d\mathbf{q}_i\right) Lm_1 I^{-3/2}\sum_{\gamma=1}^{2} q_1^{\gamma}\epsilon_{\gamma\alpha}\, E_r^{N-5/2}\delta^2\left(\sum_i \mathbf{q}_i\right)}{\int\left(\prod_{i=2}^{N} d\mathbf{q}_i\right) I^{-1/2}\sum_{\gamma=1}^{2} q_1^{\gamma}\epsilon_{\gamma\alpha}\, E_r^{N-5/2}\delta^2\left(\sum_i \mathbf{q}_i\right)} \\ &= Lm_1 < I^{-1} >_{\mathbf{q}_1}\sum_{\gamma} q_1^{\gamma}\epsilon_{\gamma\alpha}. \end{aligned} \quad \text{(C.14)}$$

Finally

$$< \mathbf{p}_1 >_{\mathbf{q}_k} = Lm_1 < I^{-1} >_{\mathbf{q}_k}\sum_{\gamma,\alpha} q_k^{\gamma}\epsilon_{\gamma\alpha}\hat{\mathbf{e}}_{\alpha}, \quad \text{(C.15)}$$

where $\hat{\mathbf{e}}_{\alpha}$ is the α-component unit vector.

$< \mathbf{p}_k^2 >_{\mathbf{q}_k}$ can be derived in a similar way, and the result is

$$\begin{aligned} < \mathbf{p}_k^2 >_{\mathbf{q}_k} &= 2m_k\left(1 - \frac{m_k}{M}\right) < \frac{N-5/2}{E_r} >_{\mathbf{q}_k} \\ &\quad -m_k I_k < \frac{N-5/2}{I\,E_r} >_{\mathbf{q}_k} + I_k L^2 m_k < I^{-2} >_{\mathbf{q}_k}, \end{aligned} \quad \text{(C.16)}$$

where $I_k = m_k \mathbf{q}_k^2$.

Appendix D

Temperature at constant pressure

The entropy of a configuration in MMMC can be written as (see sec. 3.3)

$$S = \ln \int_{W'} d\mathbf{x}\, C(\mathbf{x}) \frac{V^{N_f}}{NCC} E_k^{\tilde{N}}. \tag{D.1}$$

The microcanonical inverse temperature β (at constant *volume*) is

$$\beta = \left\langle \frac{\tilde{N}}{E_k} \right\rangle, \tag{D.2}$$

and the microcanonical pressure is

$$p = \frac{1}{\beta} \left[\left\langle \frac{N_f}{V} \right\rangle - \left\langle \frac{\partial \ln NCC}{\partial V} \right\rangle \right]. \tag{D.3}$$

The microcanonical inverse temperature at constant *pressure* β_p is given by

$$\beta_p = \beta \left[1 - p \frac{\frac{\partial p}{\partial E}|_V}{\frac{\partial p}{\partial V}|_E} \right], \tag{D.4}$$

where

$$\frac{\partial p}{\partial E}|_V = -\beta^{-2} \frac{\partial^2 S}{\partial E^2}|_V \frac{\partial S}{\partial V}|_E + \beta^{-1} \frac{\partial^2 S}{\partial E \partial V}, \tag{D.5}$$

$$\frac{\partial p}{\partial V}|_E = -\beta^{-2} \frac{\partial^2 S}{\partial E \partial V} \frac{\partial S}{\partial V}|_E + \beta^{-1} \frac{\partial^2 S}{\partial V^2}|_E, \tag{D.6}$$

$$\frac{\partial S}{\partial V}|_E = \left\langle \frac{N_f}{V} \right\rangle - \left\langle \frac{\partial \ln NCC}{\partial V} \right\rangle \left\langle \frac{1}{1} \right\rangle, \tag{D.7}$$

$$\frac{\partial S}{\partial E}|_V = \left\langle \frac{\tilde{N}}{E_k} \right\rangle, \tag{D.8}$$

$$\frac{\partial^2 S}{\partial V^2}|_E = -\left(\frac{\partial S}{\partial V}|_E \right)^2 + \left\langle \frac{Nf(Nf-1)}{V^2} \right\rangle - \left\langle \frac{2N_f}{V} \frac{\partial \ln NCC}{\partial V} \right\rangle + \left\langle 2 \left(\frac{\partial \ln NCC}{\partial V} \right)^2 \right\rangle - \left\langle \frac{1}{NCC^2} \frac{\partial^2 NCC}{\partial V^2} \right\rangle, \tag{D.9}$$

$$\frac{\partial^2 S}{\partial E^2}|_V = -\left(\frac{\partial S}{\partial E}|_V\right)^2 + \left\langle \frac{\tilde{N}(\tilde{N}-1)}{E_k^2} \right\rangle, \tag{D.10}$$

and

$$\frac{\partial^2 S}{\partial E \partial V} = -\frac{\partial S}{\partial V}|_E \frac{\partial S}{\partial E}|_V - \left\langle \frac{\partial \ln NCC}{\partial V} \frac{\tilde{N}}{E_k} \right\rangle + \left\langle \frac{N_f}{V} \frac{\tilde{N}}{E_k} \right\rangle. \tag{D.11}$$