Modifications of statistics under dimer diffusion

Georges Sitja

Modifications of statistics under dimer diffusion

Georges Sitja
CNRS, UMR 7325, Aix-Marseille University, Cinam, Campus Luminy, Case 913,
F-13288 Marseille 09, France.

Keywords: *size distribution, nanosized cluster, auto-organization, dimer diffusion, surface diffusion*

The diffusion, "explosion" and "evaporation" of dimers and the subsequent coalescence are treated in a formal way by identifying and solving the differential equations deduced from the respective behaviors of dimers in the different cases. This study leads to analytic formulas allowing to calculate, in a simple and fast way, the size statistics obtained after the coalescence of the dimers or their constituents once the dimers have completely disappeared. These formulas are of capital interest to characterize systems in which the dimers initially present disappear.

1 Introduction

On December 29th, 1959, Richard Feynman gave a speech [1] on the still unexplored properties of infinitely small objects. He specifies that by "small" he does not mean objects like millimeter-sized electric motors which already exist in 1960, but of structures whose characteristic sizes are of the order of few nanometers. Although the main part of his speech speaks about miniaturization, he foresees the quite singular properties of the nanometric objects, making appear quantum effects. He also makes an allusion to the problem that could constitute the diffusion for the stability of such objects.

Nowadays, the study of the properties of nanometric clusters of atoms has become common in various fields, such as optics[2][3], magnetism[4], and heterogeneous catalysis[5]. The search for new properties has led to the study of smaller and smaller metal particles, approaching, or even reaching the ultimate size of the individual atom[6]. The use of nano-structured supports makes that, nowadays, many scientific studies are conducted on cluster networks at critical sizes where diffusion is likely to occur[7][8][9][10][11]. When depositing atoms on a nano-structured surface, the Poisson distribution[12] can be achieved by choosing proper growth parameters (substrate temperature, deposition rate). This is interesting, not only because of the satisfaction provided by the knowledge of the sample state, but also because it's extremely useful to deduce the properties of the clusters. Indeed, by measuring the activity of an assembly of particles and by varying the quantities of deposited atoms, the knowledge of the Poisson distribution allows us to deconvolute the signal to obtain the responses of the clusters as a function of their size. However, if during the experiments, because of the temperature or the chemical environment, diffusion occurs, the initial well-characterized system disappears leadingand is replaced by a new and inoperable system. The experimental conditions can be sufficient to initiate the diffusion of monomers[8], dimers[13], but also larger clusters[9]. At present, the most common technique to determine the size statistics of very small clusters is to use an STM (Scanning Tunneling Microscope) which provides safely the ratio between occupied and unoccupied sites, i.e. the occupancy rate.

In the litterature, the most advanced formula to determine the occupancy rate after a deposition assuming that monomers and dimers have diffused is given in an article by Liu et al.[14]. This formula gives the occupancy rate R_{occ} assuming the initial distribution is known, but does not tell us anything about the final size distribution.

$$R_{occ} = 1 - P_1 - P_2 \tag{1}$$

Where P_n are the known initial probabilities of having a site occupied by a cluster of n atoms. However, this formula is eminently false since it does not take into account the formation of new particles of size greater

than 2 during the coalescence of dimers with monomers or dimers with dimers. One can easily convince oneself that by thinking about a deposit where only monomers and dimers are present, the occupancy rate will be 0 after the diffusion. It is strange that such kind of bogus formula can be pulled out of a hat. It is also amazing that this bogus formula is used to question the experimental measurements of another research team.

2 Motivations

This study is motivated by the following concrete problem (which will serve as an experimental reference when it will not be specified otherwise): One makes a deposit of atoms (for example, by condensation of a flow of atoms coming from an evaporator) on a surface having nucleation centers distributed on a network. In this case, the environment of each nucleation center is identical and the probability of capture of an atom diffusing on the surface is the same whatever the nucleation center considered. The first question that can be asked is "What is the size distribution of the clusters formed as a function of the mean number of atoms per site deposited? In fact, the answer to this question is known: If an average of x atoms per nucleation site have been deposited, the probability of having a cluster of n atoms follows the Poisson law[15]:

$$P_n(x) = \frac{1}{n!}x^n e^{-x} \tag{2}$$

The second question is: What is the new size distribution if during the experiments the monomers or the dimers start to diffuse and disappear to leave only clusters of size greater than 2 atoms. The diffusion of monomers is already addressed in an exact way[16], and I will focus here on the dimer diffusion.

3 How to get rid of dimers?

Before starting the calculations, it is good to think about the different scenarios that eliminate dimers from the surface. Since the beginning of this article, the term "diffusion" is used regularly, because it is the most natural thing that can happen to monomers and dimers. For monomers it is obvious: they have no other possibility than to diffuse if they want to disappear (except to dissolve them in the substrate, which we will not consider here: we will work with a constant quantity of matter, or in other words, with constant number of atoms). The dimers can diffuse on the surface in the manner of monomers before going to meet another cluster already present. However, even if it seems natural to think that this mechanism describes what happens on the surface, one should not neglect other possibilities carrying out a theoretical study: the dimers could for example "decompose" leading to the diffusion of the monomers resulting from this "decomposition". There is also the last possibility (which I think unlikely): the dimer can "decompose", and the interaction of one of the monomers with the surface can be such that only one of the released atoms starts to diffuse.

Hereafter, I will call "diffusion" the process by which a dimer moves as a whole on the surface. The "dislocation" followed by the diffusion of two individual atoms will be called "explosion" and finally, the unlikely case where one of the atoms resulting from the "decomposition" remains fixed and that the other diffuse will be called "evaporation".

We note here that, in the case where the dimer diffuses, nothing prohibits that the monomers are stable on the surface. This case may seem really surprising, but in the calculations, this possibility can be considered without adding additional complications.

4 Study assumptions.

Six assumptions will be made for this study:

1. Initial probabilities are known.

2. Only dimers disappear.
3. The diffusion of dimers or constituents can be decomposed in two of tree steps according the scenario:
   (a) If the dimer diffuses:
      i. A dimer is removed from the set of dimers.
      ii. The taken dimer is then placed randomly at the surface (i.e. on a nucleation site)
   (b) If the dimer explode:
      i. A dimer is removed from the set of dimers
      ii. The first monomer is then placed randomly at the surface on a site holding a cluster.
      iii. The second single atom is then also placed randomly on a site holding a cluster.
   (c) If the dimer evaporates:
      i. A dimer is removed from the set of dimers and a monomer is added to the set of monomers.
      ii. A single atom is placed on a site on the surface.
4. The number N of nucleation sites is very large : $1/N \ll 1$
5. The mean free path of a diffusing atom is large in comparison of the distance of nucleation centers.
6. Finally, the size of a cluster is negligible in regard of the distance of nucleation center. This means that the capture probability for an atom does not depend of the size of the cluster already present on the nucleation center, and that condition 5 will be fulfilled.

5 Dimer Diffusion

We will suppose here that the dimers leave their position and that, without breaking up, they diffuse until they reach a nucleation site (already occupied or not).

Let us consider a starting situation in which we have the probability P_0 of having an empty site, the probability P_1 of having a site with a monomer, P_2 of having a site with a dimer... and P_n of having a site with a cluster of n atoms. A priori, P_1 should be equal to 0, but, as said above, we are going to keep it, it won't make the calculations more complicated, and it could allow us to "solve" the improbable but not impossible problem where monomers, because of their enhanced interaction with the substrate, could be more efficiently trapped on the surface defects than the dimers.

The diffusion of the dimer can be split into two steps: 1 - a dimer is removed from the surface; 2 - the dimer is deposited on a randomly selected site. It will be noticed that the first half step of the diffusion will impact only the number of empty sites and the number of sites containing a dimer. The second step will affect the number of all other size classes.

Starting from Study assumptions, the calculations given in Appendix 1 - Calculations for dimer diffusion lead to the following formulas summarizing the size statistics after dimer diffusion:

$$\begin{aligned}
\underline{P_0} &= 1-(1-P_0)e^{-x_a} \\
n\ odd\ :\ \underline{P_{n\neq 0}} &= e^{-x_a}\left[\frac{1}{(n/2)!}x_a^{n/2}(P_0-1)+\sum_{k=0}^{n/2-1}\frac{1}{k!}x_a^k P_{n-2k}\right] \\
n\ even\ :\ \underline{P_n} &= e^{-x_a}\left[\sum_{k=0}^{(n-1)/2}\frac{1}{k!}x_a^k P_{n-2k}\right] \\
&with \\
x_a &= \frac{P_2}{1-P_0}
\end{aligned} \tag{3}$$

P_n being the initial probabilities before diffusion, and $\underline{P}_n$ the probabilities once all dimers have diffused and coalesced.

6 Dimer explosion

While less plausible than the case treated in section 5, this case has to be solved, at least for the completeness of the present study. The differences between the size histograms calculated here and those calculated previously, when compared with experiments carefully carried out for this purpose, could allow to privilege this explosion mechanism over the diffusion mechanism, and provide indications on the hierarchy of energies involved in this or that system.

We suppose here that a dimer breaks up because of the temperature (or the chemical environment), and that the two monomers formed will diffuse very quickly to stick on a pre-existing nucleus.

It is important to realize that we cannot have stable monomers, and that if the two released monomers meet, it will be as if nothing had happened. We will therefore not consider this possibility.

Here the diffusion will take place in three steps: 1 - one dimer is removed from the surface; 2 - the first atom is dropped on a site containing a sufficiently stable cluster (at least a dimer); 3 - the second atom is also dropped on a sufficiently stable cluster.

The calculations detailed in Appendix 2 - Calculations for dimer explosion lead to the following formulas:

$$
\begin{aligned}
\underline{P}_0 &= P_0 + x_a \\
\underline{P}_1 &= 0 \\
\underline{P}_2 &= 0 \\
\underline{P}_{n\geq 3} &= \left[K_n + \sum_{i=1}^{n-2} K_{n-i} \frac{(-2)^i}{i!} [\ ln(A - x_a)\ ]^i \right] (A - x_a)^2 - 2^{(n-2)}(A - x_a) \\
&y \\
&with \\
x_a &= A - \frac{A^2}{A + P_2} \\
K_{(n\geq 3)} &= \frac{P_n + 2^{n-2}A}{A^2} - \sum_{i=1}^{n-2} K_{n-i} \frac{(-2)^i}{i!} [\ ln(A)\ ]^i \\
A &= 1 - P_0 \\
and\ K_2 &= \frac{A + P_2}{A^2}
\end{aligned}
\tag{4}
$$

With always P_n representing initial probabilities before the explosion, and $\underline{P}_n$ the probabilities once all dimers have vanished.

7 Dimer evaporation

One can consider the situation in which one atom of the cluster is more strongly bound to the surface than to the other atom of the dimer. In this case, the thermal energy necessary to break a dimer can be too weak to activate the diffusion of one of the two atoms of the dimer: One of the atoms starts its random walk and the other remains fixed on the original site.

In this case, the diffusion splits in two steps: a - A dimer disappears and a monomer appears, b - The atom diffusing on the surface stops on a random site, be it empty, or occupied by a cluster of any size.

From the calculations described in Appendix 3 - Calculations for dimer evaporation, we obtain the following formulas reflecting the evaporation of dimers.

$$\underline{P_0} = P_0 e^{-x}$$

$$\underline{P_1} = 1 - [1 - P_1 - x_a P_0] e^{-x_a}$$

$$\underline{P_{n\geq 2}} = \left[\sum_{i=0}^{n} P_{n-i} \frac{x_a^i}{i!} \quad - \quad \frac{x_a^{(n-1)}}{(n-1)!} \right] e^{-x_a} \tag{5}$$

with

$$x_a = \frac{(1-P_1) - \sqrt{(1-P_1)^2 - 2P_0P_2}}{P_0} \; if \; P_0 \neq 0$$
$$x_a = \frac{P_2}{1-P_1} \; if \; P_0 = 0$$

Again P_n representing initial probabilities before the evaporation, and $\underline{P_n}$ the probabilities once all dimers have vanished.

8 Discussion

It is almost obvious that the dimers will "disappear" after the monomers have diffused and, unless a physical or chemical process is available to obtain a known size distribution on a surface with $P_1 = 0$, it will be necessary to use the formulas described in the paper on the diffusion of monomers[16] to handle the initial distribution before using the formulas given in this work.

8.1 Experimental verification

One of the first papers reporting the use of nano-structured surfaces for the organization of metal clusters gives enough details to make a comparison between an experimental result and the theoretical formulas obtained here. In this article published in 2006, the team of T. Michely uses a moiré produced by graphene on a surface (111) of an iridium (Ir) single crystal to organize Ir clusters. To know the deposited quantity they examine the 2D islands formed where no graphene is present, which allows to calibrate the deposit in a rather reliable way.
The table 1 allows to compare in a synthetic way the experimental results with the different theoretical predictions. As expected, the trivial formula 1 disagrees more as the deposit is low and as there are essentially monomers and dimers just after the deposit. The columns of the table labelled “dimer diffusion” and “dimer explosion” are calculated from the probabilities given by the formulas obtained in the reference [16], which are themselves calculated from the Poisson distribution corresponding to the values n of the average number of atoms per site. Here we see that the diffusion of dimers gives a result quite close to the explosion of dimers. We know anyway that in this case, it is the dimer diffusion that must be considered given the Ir-Ir and Ir-graphene interaction energies.
For each experimental point, we calculate the quantity of Ir that should have been deposited to obtain, with the theoretical formulas, the experimental coverage. This value is called "recalculated deposit". Here, it is the comparison of the recalculated deposit with the experimental deposit that is used to compare the experiment with the theory.

Deposit θ M.L.	n	experimental R_{occ}	formula 1	dimer diffusion 3	recalculated/ measured	dimer explosion 4
0.01	0.87	19.2%	5.8%	19.4%	0.988	21.1%
0.02	1.74	42.6%	25.3%	38.5%	1.109	40.9%
0.03	2.61	65.5%	48.4%	56.2%	1.195	58.3%
0.04	3.48	68.7%	67.5%	71.1%	0.954	72.3%
0.05	4.35	82.9%	80.9%	82.2%	1.014	82.8%

Table 1: Comparison of experimental occupancy rates R_{occ} with theoretical predictions for the different formulas. The formula 1 is the one reported in [14]. The initial distribution before any diffusion is assumed to be given by the Poisson law. To easily compare the theoretical and experimental results, the 6^{th} column gives the ratio between theoretical deposit and the experimental deposit.

The comparison of the "calculated/measured" ratio in column 6 of table 1 shows that over the five experimental values, a very good agreement for three of them is obtained. The 0.02 monolayer deposit is in disagreement of 10% and this disagreement rises to 20% for the 0.03M.C. deposit. Can we incriminate measurement imprecision during the deposition for two of these experiments ? In fact it is difficult to say, however the experimenters are often confronted with experimental hazards not always controllable. The small difference between the occupancy rate between the deposit for 0.03M.C. and the following one for 0.04M.C. would tend to show that there is indeed an experimental under evaluation of the quantity deposited during the third deposit. To this we must add that in 2006 the aim of this work was not to confirm any theory but to prove that the type of substrate used made it possible to obtain particle arrays with a known distribution (the Poisson distribution) allowing to characterize properly the sample. We can add that a wrong model would certainly give increasing (or decreasing) deviations with respect to the coverage, which is not the case here. The value for the deposit of 0.03ML distinguishes itself, but the other values are within the norm for a study that was not conducted with the aim of a rigorous comparison with a theory. One can hope that some experimenter will decide to accurately verify the conclusions of the formulas given here.

8.2 Diffusion Vs Explosion

While improbable (at least with atoms), one can envisage that experimentally a cluster would break before diffusing. One may wonder if the formulas allow distinguishing between one or the other of these scenarios. Since it is the dimers that will have a different behavior between both scenarios, one needs a distribution that before diffusion or explosion maximizes the number of dimers.

	Initial size distribution	dimer diffusion	dimer explosion
P_0	0.0	0.0	0.0
P_1	0.0	0.0	0.0
P_2	α	0.0	0.0
P_3	0.0	0.0	0.1931α
P_4	0.0	0.1839α	0.1736α
P_5	0.0	0.0	0.08876α
P_6	0.0	0.1226α	0.0324α
P_7	0.0	0.0	0.0093α
P_8	0.0	0.0459α	0.0022α
Relative occupancy	1	$e^{-1} \simeq 0.36787$	1/2

Table 2: Comparative evolution of a dimer assembly under the effect of diffusion or explosion

If there are only dimers, the difference between diffusion and explosion is very important as shown in table 2. If we can only access the occupancy rate, the ratio between the number of dimers deposited and the number of particles obtained gives a clear signal to discriminate one or the other mechanism: indeed, counting 37 clusters or 50 clusters in an area where there were 100 dimers should be easy even taking into account the uncertainty with only a single counting. If experimentally one can count the number of atoms in the clusters obtained, the lack or the existence of clusters with an odd number of atoms is an infallible argument.

However, producing a sample with only dimers seems difficult unless one is depositing mass-sorted clusters. In most cases, the deposition of atoms is made by evaporation and condensation. One would suppose that in the case of our concern, one should make a deposit maximizing the number of dimers, but this is not the case, because the monomers present on the surface will certainly diffuse before the dimers. The results obtained by applying the formulas for monomer diffusion show that the ratio of dimers over the total number of clusters is maximized when the amount deposited tends towards 0. The size distribution after monomer diffusion[16] when $\theta \to 0$ is given by the following formula:

$$P_n = \frac{n-1}{n!} \tag{6}$$

Taking as a starting point this distribution we obtain the occupancies listed in the table 3. We notice that the differences are less important but not negligible: Counting on a zone where 1000 particles were initially found, we will find in one case 606 ± 25 clusters and in the other 666 ± 25 clusters.

	initial size distribution	dimer diffusion	dimer explosion
P_0	0	0	0
P_1	0	0	0
P_2	0.5	0.0	0.0
P_3	0.333333	0.2021768	0.2443238
P_4	0.125	0.1516326	0.2268339
P_5	0.033333	0.1213061	0.1239769
P_6	0.006944	0.0673922	0.0498425
P_7	0.0011904	0.0361029	0.0160665
P_8	0.0000144	0.0164087	0.0043355
relative occupancy	1	0.60653	0.66666

Table 3: Comparative changes under diffusion or explosion from the size distribution given by formula 6.

9 Conclusion

While these formulae may be useful to discriminate different dimer disappearance scenarios, this is not their main use. If one considers a set of clusters giving properties to a given sample, optical, magnetic, or catalytic properties, the proper characterization of the proportion of clusters of such or such size is the only way to deduce the individual properties depending on the cluster size. By recording the response of a sample as a function of coverage, if monomer and dimer diffusion is involved, the formulas given in this study added to the formulas given in [16] allow measuring the activities of clusters of size equal to or larger than 3 as a function of their size. Although these formulas succeed in an experimental test, one can hope that they will be confirmed (or invalidated) by dedicated studies, which would allow them to be used without fear as is often done with the Poisson distribution. Lastly, the main object of this study was related to the diffusion of atoms and dimers on a surface in the way it can occur in the systems studied in the field of nanosciences. Nevertheless, these formulas can equally be applied to other domains where atoms or molecules would not be on a surface but in a 3-dimensional matrix. The only importance is to respect the initial hypotheses (Study assumptions).

References

[1] Feynman, R., *There's plenty of room at the bottom.* Engineering and Science, volume 23(1960)(5):22–36. https://resolver.caltech.edu/CaltechES:23.5.0.

[2] Ekardt, W.; Thoai, D.; Frank, F. and Schulze, W., *Static polarizability of metal-clusters and the related blue shift of the dipolar surface-plasmon.* Solid State Commun., volume 46(1983)(7):571–574.

[3] Maier, S.; Brongersma, M. and Atwater, H., *Electromagnetic energy transport along arrays of closely spaced metal rods as an analogue to plasmonic devices.* Appl. Phys. Lett., volume 78(2001)(1):16–18.

[4] Hannour, A.; Bardotti, L.; Prevel, B. et al., *2D arrays of CoPt nanocluster assemblies.* Surf. Sci., volume 594(2005)(1-3):1–11.

[5] Poppa, H., *Model studies in catalysis with UHV-deposited metal particles and clusters.* Vacuum, volume 34(1984)(12):1081–1095.

[6] Gates, B.C.; Flytzani-Stephanopoulos, M.; Dixon, D.A. and Katz, A., *Atomically dispersed supported metal catalysts: perspectives and suggestions for future research.* Catal. Sci. Technol., volume 7(2017)(19):4259–4275.

[7] Will, M.; Atodiresei, N.; Caciuc, V. et al., *A Monolayer of Hexagonal Boron Nitride on Ir(111) as a Template for Cluster Superlattices.* ACS Nano, volume 12(2018)(7):6871–6880.

[8] Dull, F.; Meusel, M.; Spaeth, F. et al., *Growth and stability of Pt nanoclusters from 1 to 50 atoms on h-BN/Rh(111).* Phys. Chem. Chem. Phys., volume 21(2019)(38):21 287–21 295.

[9] Alyabyeva, N.; Ouvrard, A.; Zakaria, A.M. and Bourguignon, B., *Probing Nanoparticle Geometry down to Subnanometer Size: The Benefits of Vibrational Spectroscopy.* J. Phys. Chem. Lett., volume 10(2019)(3):624–629.

[10] Qiao, B.; Wang, A.; Yang, X. et al., *Single-atom catalysis of CO oxidation using Pt-1/FeOx.* Nature Chem., volume 3(2011)(8):634–641.

[11] Liu, L. and Corma, A., *Metal Catalysts for Heterogeneous Catalysis: From Single Atoms to Nanoclusters and Nanoparticles.* Chem. Rev., volume 118(2018)(10):4981–5079.

[12] Sitja, G.; Omar Unac, R. and Henry, C.R., *Kinetic Monte Carlo simulation of the growth of metal clusters on regular array of defects on insulator.* Surf. Sci., volume 604(2010)(3-4):404–408.

[13] N'Diaye, A.T.; Bleikamp, S.; Feibelman, P.J. and Michely, T., *Two-dimensional Ir cluster lattice on a graphene moire on Ir(111).* Phys. Rev. Lett., volume 97(2006)(21).

[14] Liu, X.; Han, Y.; Evans, J.W. et al., *Growth morphology and properties of metals on graphene.* Pro. Surf. Sci., volume 90(2015)(4):397–443.

[15] Poisson, S.D., *Recherches sur la probabilité des jugements en matière criminelle et en matière civile ; précédées des Règles générales du calcul des probabilités.* Bachelier, volume 205(1837):.

[16] Sitja, G., *Statistics modification under coalescence.* arXiv [cond-mat.stat-mech], (2021).

10 Appendix 1 - Calculations for dimer diffusion

10.1 Determining the recurrence formulas

10.1.1 P_0

Let us consider a surface with a very large number N of nucleation centers. The number of empty sites N_0 is such that $N_0 = N \times P_0$. Removing a dimer is equivalent to increase P_0 by $1/N$. Indeed the number of empty sites after the first step is ${N_0}' = N_0 + 1$, $N_0 + 1 = {P_0}' \times N$ which means that ${P_0}' = P_0 + 1/N$. In the rest of the document I will replace $1/N$ by ε to simplify the notation. The second step of the diffusion will consist of dropping randomly this dimer on one of the sites of the surface. Obviously the probability that the dimer "lands" on a site containing n atoms is P_n. Here we are interested in knowing what happens to P_0. At the second stage of the diffusion P_0 is decreased by ${P_0}' \times \varepsilon$, i.e.: ${P_0}' = {}^1P_0 = {P_0}' - {P_0}' \times \varepsilon$ and finally :

$$ {}^1P_0 = (P_0 + \varepsilon)(1 - \varepsilon) $$

Noting ${}^aP_n$ the value of P_n after a elementary diffusions.

Trivially, we can extend this result. After a+1 diffusions :

$$ {}^{a+1}P_0 = ({}^aP_0 + \varepsilon)(1 - \varepsilon) \tag{7} $$

10.1.2 P_1

As mentioned earlier, we expect P_1 to be equal to 0 and remain so, but let's make the calculation as if the monomers were stable. The first step of the diffusion does not affect P_1, but the second step decreases the initial value of P_1 by εP_1. In an immediate and trivial way, we can write that at the end of the first diffusion ${}^1P_1 = P_1(1 - \varepsilon)$. This leads to the following recurrence formula:

$$ {}^{a+1}P_1 = {}^aP_1(1 - \varepsilon) \tag{8} $$

10.1.3 P_2

The case of P_2 is a bit more tricky because the probability of having dimers at step $a + 1$ depends on the probabilities P_0 and P_2 at step a. Obviously, the first step of the diffusion decreases the probability P_2 of ε and we have ${P_2}' = P_2 - \varepsilon$. The second step will increase it by $\varepsilon \times {P_0}'$ in the case that the dimer reaches an empty site, or decrease it by $\varepsilon \times {P_2}'$ if the dimer finishes its walk on a site occupied by another dimer. At the end of the second step of the diffusion, we have:

$$ {}^1P_2 = P_2(1 - \varepsilon) - \varepsilon + \varepsilon P_0 + 2\varepsilon^2 $$

We can safely neglect the quadratic term in ε and write the following recurrence formula:

$$ {}^{a+1}P_2 = {}^aP_2(1 - \varepsilon) - \varepsilon(1 - {}^aP_0) \tag{9} $$

10.1.4 P_3

P_3 is handled almost like P_1. Indeed the first step of the diffusion does not disturb the probability of having a trimer, but the second step will subtract 1 from N_3 in the situation where the dimer arrives on a site containing a trimer. There is another way to modify P_3, indeed if the dimer reaches a site containing a monomer we will add 1 to N_3. The new probability is consequently :

$$^1P_3 = P_3(1-\varepsilon) + P_1\varepsilon$$

A formula that we will generalize to the following recurrence form:

$$^{a+1}P_3 = {}^aP_3(1-\varepsilon) + \varepsilon^a P_1 \tag{10}$$

10.1.5 P_4

The only difference between P_3 and P_4 is that if P_1 was not altered by the first step of the diffusion, P_2 is. P_4 increases by ε if the diffusing dimer meets another dimer. The probability of this event is the number of dimers remaining after the first step of the diffusion, that is $P_2 - \varepsilon$. Finally, we can write:

$$^1P_4 = P_4(1-\varepsilon) + (P_2 - \varepsilon)\varepsilon$$

i.e.:

$$^1P_4 = P_4(1-\varepsilon) + \varepsilon P_2$$

once the quadratic terms removed. A formula that will be extended to the following recurrence formula:

$$^{a+1}P_4 = {}^aP_4(1-\varepsilon) + \varepsilon^a P_2 \tag{11}$$

10.1.6 P_5 and $P_{n\geq 5}$

The evolution of P_n for $n \geq 5$ and P_3 are strictly identical, in the sense that from size 5, the first step of the diffusion will have no impact on P_5: The evolution of P_5 depends on P_5 and P_3 and we can add that for a given size n P_n will only depend on P_n and P_{n-2}. It follows:

$$^1P_n = P_n(1-\varepsilon) + \varepsilon P_{n-2}$$

That is extended to the following recurrence formula:

$$^{a+1}P_n = {}^aP_n(1-\varepsilon) + \varepsilon^a\ P_{n-2} \tag{12}$$

10.2 Iterations

To derive the value of the different probabilities after a given number of single diffusions, one must iterate the recurrence formulas obtained previously.

10.2.1 P_0

Let's see how P_0 evolves. I recall the recurrence formula (7):

$$^{a+1}P_0 = ({}^aP_0 + \varepsilon)(1-\varepsilon)$$

$$\begin{aligned} ^{a+1}P_0 &= {}^aP_0(1-\varepsilon) + \varepsilon(1-\varepsilon) \\ &= {}^aP_0(1-\varepsilon) + \varepsilon - \varepsilon^2 \end{aligned}$$

in which we can remove the quadratic terms in ε :

$$ {}^{a+1}P_0 = {}^{a}P_0 + \varepsilon(1 - {}^{a}P_0) \tag{13} $$

The iterations result in:

$$
\begin{aligned}
{}^{1}P_0 &= P_0(1-\varepsilon) + \varepsilon \\
{}^{2}P_0 &= {}^{1}P_0(1-\varepsilon) + \varepsilon \\
&= (P_0(1-\varepsilon) + \varepsilon)(1-\varepsilon) + \varepsilon \\
&= P_0(1-\varepsilon)^2 + \varepsilon(1-\varepsilon) + \varepsilon \\
{}^{3}P_0 &= {}^{2}P_0(1-\varepsilon) + \varepsilon \\
&= P_0(1-\varepsilon)^3 + \varepsilon(1-\varepsilon)^2 + \varepsilon(1-\varepsilon) + \varepsilon \\
{}^{4}P_0 &= {}^{3}P_0(1-\varepsilon) + \varepsilon \\
&= P_0(1-\varepsilon)^4 + \varepsilon(1-\varepsilon)^3 + \varepsilon(1-\varepsilon)^2 + \varepsilon(1-\varepsilon) + \varepsilon
\end{aligned}
$$

it seems to emerge a fairly simple pattern that would give :

$$ {}^{a}P_0 = P_0(1-\varepsilon)^a + \varepsilon \sum_{i=0}^{a-1} (1-\varepsilon)^i $$

We can easily prove that:

$$ \sum_{i=0}^{a} \alpha^i = \frac{\alpha^{a+1} - 1}{\alpha - 1} $$

hence:

$$ \sum_{i=0}^{a-1} (1-\varepsilon)^i = \frac{(1-\varepsilon)^a - 1}{(1-\varepsilon) - 1} = \frac{1 - (1-\varepsilon)^a}{\varepsilon} $$

and finally:

$$ {}^{a}P_0 = 1 - (1 - P_0)(1-\varepsilon)^a \tag{14} $$

In the limit where ε tends towards 0, we can deduce a continuous formula by noticing that $(1-\varepsilon)^a = e^{-a\varepsilon}$. The continuous version of the formula(14) is then:

$$ P_0(x) = 1 - (1 - P_0)e^{-x} \tag{15} $$

writing $x = a\varepsilon$ which is nothing else than the number of diffusions per site (a/N)

10.2.2 A short break to think

We are interested here in the continuous versions of the evolution of probabilities. It is with them that we can easily do calculations. Obtaining $P_0(x)$ required quite a lot of work, and it would be good to ask ourselves if we could not get rid of the tedious discrete calculations to obtain an equivalent result, but in a simpler way.

For this, let us return to the formula (13) and see what it has to say. This formula is recalled here:

$$ {}^{a+1}P_0 = {}^{a}P_0 + \varepsilon(1 - {}^{a}P_0) $$

Each new diffusion increases a by 1. Obviously, the number of elementary diffusions to reach a given state will be proportional to the number N of sites considered, and we can introduce here the number x of diffusions per site which is of course a/N. We have $x = a/N$, and making an additional diffusion increases x by a small amount δx. Then we have: $x + \delta x = (a+1)/N$. It follows that $\delta x = \varepsilon$. We can add that $P_0(x)$ is also slightly modified by this additional diffusion. The formula (**??**) can thus be rewritten as follows:

$$P_0(x) + \delta P_0 = P_0(x) + (1 - P_0(x))\delta x$$

Obviously in the limit where δx tends to 0 we have :

$$P_0(x) + \mathrm{d}P_0 = P_0(x) + (1 - P_0(x))\mathrm{d}x \tag{16}$$

The equation (16) allows us to find a differential equation for which the solution should be what we are searching for:

$$\frac{\mathrm{d}P_0}{\mathrm{d}x}(x) = 1 - P_0(x)$$

i.e.:

$$P_0'(x) = 1 - P_0(x) \tag{17}$$

Solving this equation is straightforward at this point, since we already know the solution that was computed earlier by the iteration method (15). We are going nevertheless to check that this is the correct solution:

$$\begin{aligned} P_0(x) &= 1 - (1 - P_0)e^{-x} \\ P_0'(x) &= (1 - P_0)e^{-x} \end{aligned}$$

$$\begin{aligned} 1 - P_0(x) &= 1 - [1 - (1 - P_0)e^{-x}] \\ &= (1 - P_0)e^{-x} = P_0'(x) \end{aligned}$$

which satisfies the differential equation (17).

10.2.3 P_1

From the recurrence formula (8) for P_1 let's try to determine the differential equation that governs $P_1(x)$. We have:

$$\begin{aligned} {}^{a+1}P_1 &= {}^{a}P_1(1 - \varepsilon) \\ P_1(x) + \mathrm{d}P_1(x) &= P_1(x) - P_1(x)\mathrm{d}x \\ \frac{\mathrm{d}P_1(x)}{\mathrm{d}x} &= -P_1(x) \end{aligned}$$

That can simply be written:

$$P_1'(x) = -P_1(x) \tag{18}$$

The result is immediate:

$$P_1(x) = K_1 e^{-x}$$

for $x = 0$; $P_1(0) = K_1 = P_1$; and finally:

$$P_1(x) = P_1 e^{-x} \tag{19}$$

10.2.4 P_2

P_2 is a bit more complicated because it depends not only on P_2 but also on P_0. The recurrence formula (9) is arranged as follows:

$$\begin{aligned} {}^{a+1}P_2 &= {}^{a}P_2 - {}^{a}P_2\varepsilon + {}^{a}P_0\varepsilon - \varepsilon \\ &= {}^{a}P_2 + ({}^{a}P_0 - {}^{a}P_2 - 1)\varepsilon \end{aligned}$$

The resulting differential equation is :

$$P_2'(x) = -P_2(x) + P_0(x) - 1$$

And since we know the solution for $P_0(x)$ (15), we finally obtain the differential equation that we will attempt to solve:

$$P_2'(x) = -P_2(x) - (1 - P_0)e^{-x} \tag{20}$$

We will assume that $P_2(x) = f(x)e^{-x}$. We then have:

$$\begin{aligned} P_2'(x) &= f'(x)e^{-x} - f(x)e^{-x} \\ &= (f'(x) - f(x))e^{-x} \end{aligned}$$

but (20) leads to:

$$\begin{aligned} P_2'(x) &= -P_2(x) - (1 - P_0)e^{-x} \\ &= -f(x)e^{-x} - (1 - P_0)e^{-x} \\ &= [-f(x) - (1 - P_0)]e^{-x} \end{aligned}$$

If the solution has the form that we assume, we must have:

$$-f(x) - (1 - P_0) = f'(x) - f(x)$$

that is

$$f'(x) = -(1 - P_0)$$

it follows trivially :

$$f(x) = -(1 - P_0)x + K_2$$

i.e.

$$P_2(x) = [K_2 - (1 - P_0)x]e^{-x}$$

As $P_2(0) = P_2$, we finally find:

$$P_2(x) = [P_2 - (1 - P_0)x]e^{-x} \tag{21}$$

10.2.5 P_3

P_3 is not more complicated than P_2. Indeed, from the recurrence formula (10) we derive the following differential equation:

$$P_3'(x) = -P_3(x) + P_1(x) \tag{22}$$

We will once again assume that the solution has the following form $f(x)e^{-x}$, and it comes :

$$P_3'(x) = (f'(x) - f(x))e^{-x} = (-f(x) + P_1)e^{-x}$$

with the differential equation to solve:

$$f'(x) = P_1$$

the solution, once the limit for $x = 0$ is included, is $f(x) = P_3 + xP_1$, from which we conclude:

$$P_3(x) = [P_3 + P_1 x]e^{-x} \tag{23}$$

10.2.6 P_4

The differential equation deduced from the formula (11) giving the evolution of P_4 is :

$$P_4'(x) = -P_4(x) + P_2(x)$$

Again using the same approach, $P_4(x) = f(x)e^{-x}$, we get :

$$f'(x) = P_2 - (1 - P_0)x$$

and finally:

$$P_4(x) = [P_4 + P_2 x - \frac{1}{2}(1 - P_0)x^2]e^{-x} \tag{24}$$

10.2.7 P_n

The point here is to understand why the generic solution $P_n(x) = f(x)e^{-x}$ works to find the solution to the differential equation from the formulas giving the evolutions of the probabilities. the evolution of P_n (12) gives us the following differential equation :

$$P_n'(x) = -P_n(x) + P_{n-2}(x)$$

Taking $P_n(x) = f_n(x)e^{-x}$, we will systematically obtain $P_n'(x) = (f_n'(x) - f_n(x))e^{-x}$, and if by chance, P_{n-2} is such that $P_{n-2} = f_{n-2}(x)e^{-x}$, we will always have to solve:

$$(f_n'(x) - f_n(x))e^{-x} = -f_n(x)e^{-x} + f_{n-2}(x)e^{-x}$$

which will always be simplified to:

$$f_n'(x) = f_{n-2}(x) \tag{25}$$

which will have as a solution:

$$f_n(x) = P_n + \int_0^x f_{n-2}(y)\mathrm{d}y \tag{26}$$

In order to make $P_{n-2}(x)$ of the form $f_{n-2}(x)e^{-x}$, we need $P_{n-4}(x)$ to be of the form $f_{n-4}(x)e^{-x}$. And in order to make $P_{n-4}(x)$ of the form $f_{n-4}(x)e^{-x}$, we need $P_{n-6}(x)$ to be of the form $f_{n-6}(x)e^{-x}$.... this until we find a value of n for which we know it is true. This is precisely the case since for $n = 1$ and $n = 2$ we have: $P_1(x) = P_1e^{-x}$ and $P_2(x) = [P_2 - (1 - P_0)x]e^{-x}$.

From what has just been said, we can very simply give the successive $f_n(x)$ by integrating from one step to the next $f_{n-1}(x)$:

$$\begin{aligned}
f_2(x) &= P_2 - x(1 - P_0) \\
f_4(x) &= P_4 + xP_2 - \frac{x^2}{2}(1 - P_0) \\
f_6(x) &= P_6 + xP_4 + \frac{x^2}{2}P_2 - \frac{x^3}{6}(1 - P_0) \\
f_8(x) &= P_8 + xP_6 + \frac{x^2}{2}P_4 + \frac{x^3}{6}P_2 - \frac{x^4}{4!}(1 - P_0) \\
f_{10}(x) &= P_{10} + xP_8 + \frac{x^2}{2}P_6 + \frac{x^3}{6}P_4 + \frac{x^4}{4!}P_2 - \frac{x^5}{5!}(1 - P_0) \\
f_{12}(x) &= P_{12} + xP_{10} + \frac{x^2}{2}P_8 + \frac{x^3}{6}P_6 + \frac{x^4}{4!}P_4 + \frac{x^5}{5!}P_2 - \frac{x^6}{6!}(1 - P_0) \\
f_n(x) &= P_n + xP_{n-2} + \frac{1}{2}x^2P_{n-4} + \frac{1}{6}x^3P_{n-6} + \frac{1}{24}x^4P_{n-8} + ... \\
&\quad + \frac{1}{(n/2-1)!}x^{n/2-1}P_2 - \frac{1}{(n/2)!}x^{n/2}(1 - P_0)
\end{aligned}$$

and the probability for n even can be summarized by:

$$\begin{aligned}
P_{n\neq 0}(x) &= e^{-x}\left[\frac{1}{(n/2)!}x^{n/2}(P_0 - 1) + \sum_{k=0}^{n/2-1}\frac{1}{k!}x^kP_{n-2k}\right] \\
P_0(x) &= 1 - (1 - P_0)e^{-x}
\end{aligned} \tag{27}$$

For n odd, successive integrations of $f_n(x)$ lead to :

$$\begin{aligned}
f_1(x) &= P_1 \\
f_3(x) &= P_3 + xP_1 \\
f_5(x) &= P_5 + xP_3 + \frac{1}{2}x^2P_1 \\
f_7(x) &= P_7 + xP_5 + \frac{1}{2}x^2P_3 + \frac{1}{6}x^3P_1 \\
f_9(x) &= P_9 + xP_7 + \frac{1}{2}x^2P_5 + \frac{1}{6}x^3P_3 + \frac{1}{24}x^4P_1 \\
f_{11}(x) &= P_{11} + xP_9 + \frac{1}{2}x^2P_7 + \frac{1}{6}x^3P_5 + \frac{1}{4!}x^4P_3 + \frac{1}{5!}x^5P_1 \\
f_n(x) &= P_n + xP_{n-2} + \frac{1}{2}x^2P_{n-4} + \frac{1}{6}x^3P_{n-6} + \frac{1}{24}x^4P_{n-8} + ... \\
&\quad + \frac{1}{[(n-1)/2-1]!}x^{(n-1)/2-1}P_3 + \frac{1}{[(n-1)/2]!}x^{(n-1)/2}P_1
\end{aligned}$$

and the probability for n odd can be summarized by:

$$P_n(x) = e^{-x}\left[\sum_{k=0}^{(n-1)/2}\frac{1}{k!}x^kP_{n-2k}\right] \tag{28}$$

10.3 Diffusion termination

Of course, as we are dealing here with the diffusion of dimers, the process will stop when the number of dimers is equal to 0. This is summarized by the condition $P_2(x) = 0$. The value of x_a for which $P_2(x_a) = 0$ is easily calculated:

$$x_a = \frac{P_2}{1 - P_0} \tag{29}$$

10.4 Diffusion summary.

To conclude, the size distribution after dimer diffusion is summarized in the following formulas:

$$\underline{P_0} = 1 - (1 - P_0)e^{-x_a}$$

$$n \text{ pair } : \underline{P_{n\neq 0}} = e^{-x_a}\left[\frac{1}{(n/2)!}x_a^{n/2}(P_0 - 1) + \sum_{k=0}^{n/2-1}\frac{1}{k!}x_a^k P_{n-2k}\right]$$

$$n \text{ impair } : \underline{P_n} = e^{-x_a}\left[\sum_{k=0}^{(n-1)/2}\frac{1}{k!}x_a^k P_{n-2k}\right] \tag{30}$$

$$with$$

$$x_a = \frac{P_2}{1 - P_0}$$

11 Appendix 2 - Calculations for dimer explosion

11.1 Determination of increments, of differential equations, and of probabilities

11.1.1 P_0

The first step of the diffusion increases P_0 by ε. The following steps do not influence P_0, since once deposited, the atom will diffuse very quickly to an already occupied site. So we get :

$$\begin{aligned} P_0' &= P_0 + \varepsilon \\ P_0'' &= P_0' \\ P_0''' &= P_0'' \end{aligned}$$

that is:

$$^1P_0 = P_0 + \varepsilon$$

The differential equation is:

$$P_0'(x) = 1 \tag{31}$$

$P_0(x)$ is very easy to find in this case:

$$P_0(x) = P_0 + x \tag{32}$$

11.1.2 P_1

Here, monomers are definitely prohibited, and no matter what happens :

$$P_1(x) = 0 \tag{33}$$

11.1.3 P_2

P_2 is affected by each of the three steps of the diffusion: Step 1 removes ε from P_2 the two following steps add ε according to the probability of the released monomers to fall on a dimer. One must be extremely careful here because the probability that an atom lands on a dimer is not equal to the probability of having a dimer (P_2): Indeed, in the case we are interested in here, putting an atom on an empty site does not make sense, it can only stabilize while waiting for the next diffusion on a site already occupied. After the first step of the diffusion, the number of occupied sites is $(1 - P_0') = 1 - P_0 - \varepsilon$. The sequence of the three diffusion steps thus modifies the probability of having a dimer as follows:

$$\begin{aligned} P_2' &= P_2 - \varepsilon \\ P_2'' &= P_2' - \frac{P_2'\varepsilon}{1 - P_0 - \varepsilon} \\ P_2''' &= P_2'' - \frac{P_2''\varepsilon}{1 - P_0 - \varepsilon} \end{aligned}$$

Once expanded and negligible terms dropped we obtain :

$$^1P_2 = P_2 - \frac{2P_2}{1 - P_0}\varepsilon - \varepsilon$$

Resulting in the following differential equation:

$$P_2'(x) = -\frac{2P_2(x)}{1 - P_0(x)} - 1$$

which becomes when replacing $P_0(x)$ given by the formula (32) :

$$P_2'(x) = -\frac{2P_2(x)}{1 - P_0 - x} - 1 \tag{34}$$

To simplify the notation, we will assume here and for the following $1 - P_0 = A$. We can rewrite the previous equation:

$$P_2'(x) = -\frac{2P_2(x)}{A - x} - 1 \tag{35}$$

This differential equation is not easy to solve. One can for example iterate the individual diffusions to give clues. An other method, more direct, is to note that since we have the term $-\frac{2P_2(x)}{A-x}$ in the equation, a solution that could work would be of the form:

$$P_2(x) = f_2(x)(A - x)^2$$

so:

$$\begin{aligned} P_2'(x) &= (A - x)^2 f_2'(x) - 2(A - x)f_2(x) \\ &= (A - x)^2 f_2'(x) - \frac{2P_2(x)}{A - x} \end{aligned}$$

The equation (35) results in :

$$(A - x)^2 f_2'(x) = -1$$

that is to say :

$$f_2'(x) = -(A - x)^{-2}$$

leading to

$$f_2(x) = -(A - x)^{-1} + K_2$$

hence

$$P_2(x) = -(A - x) + K_2(A - x)^2$$

We must have $P_2(0) = P_2$, which determines the constant K_2 :

$$\begin{gathered} K_2 A^2 - A = P_2 \\ \Longrightarrow \\ K_2 = \frac{A + P_2}{A^2} \end{gathered} \tag{36}$$

finally

$$P_2(x) = \frac{A + P_2}{A^2}(A - x)^2 - (A - x) \tag{37}$$

And substituting A with $1 - P_0$:

$$P_2(x) = \frac{1 - P_0 + P_2}{(1 - P_0)^2}(1 - P_0 - x)^2 - (1 - P_0 - x) \tag{38}$$

The verification of this solution can be found in paragraph Validation of the formula (38) for $P_2(x)$ in the supplementary information.

We can now know when the diffusion of dimers will end, obviously, there are two solutions to annul $P_2(x)$:

$$x_1 = A$$

and

$$x_2 = A(1 - \frac{A}{A + P_2})$$

The first solution seems not very realistic since it depends only on P_0, and we expect it, at least, to depend on P_2. Note that $A \geq 0$ and $P_2 \geq 0$ leading to $[1 - A/(A + P2)] \leq 1$. x_2 is then necessarily smaller than x_1 since it is the result of x_1 multiplied by a factor smaller than 1. x_2 will be the first value to be reached during the explosion process, and we can forget about the solution x_1 which is obviously not a physical solution. x_2 has the expected behaviour: it starts from 0 when $P_2 = 0$ and grows by increasing P_2. This is the solution we shall use to determine the end of the "diffusion" of the dimers.

$$x_a = (1 - P_0)(1 - \frac{1 - P_0}{1 - P_0 + P_2}) \tag{39}$$

11.1.4 P_3

The 3 steps of diffusion alter the probability of having a trimer in the following way:

$$\begin{aligned} P_3' &= P_3 \\ P_3'' &= P_3' - \frac{P_3'\varepsilon}{1 - P_0 - \varepsilon} + \frac{P_2'\varepsilon}{1 - P_0 - \varepsilon} \\ P_3''' &= P_3'' - \frac{P_3''\varepsilon}{1 - P_0 - \varepsilon} + \frac{P_2''\varepsilon}{1 - P_0 - \varepsilon} \end{aligned}$$

Once all the negligible terms removed, we obtain the following differential equation:

$$P_3(x)' = \frac{2}{1 - P_0(x)}\left(P_2(x) - P_3(x)\right)$$

substituting $P_0(x)$ and $P_2(x)$ with the previously determined expressions (32) and (38), and still substituting $1 - P_0$ with A we get :

$$P_3'(x) = \frac{2}{A - x}\left[\frac{A + P_2}{A^2}(A - x)^2 - (A - x)\right] - 2\frac{P_3(x)}{A - x}$$

or,once reduced:

$$P_3'(x) = 2\left[\frac{A+P_2}{A^2}(A-x)-1\right] - 2\frac{P_3(x)}{A-x} \tag{40}$$

We note that $P_3'(x)$ depends on $2P_3(x)/(A-x)$ which leads us, as for $P_2(x)$, to a solution of the form: $P_3(x) = f_3(x)(A-x)^2$.

$$\begin{aligned} P_3'(x) &= (A-x)^2 f_3'(x) - 2(A-x)f_3(x) \\ &= (A-x)^2 f_3'(x) - \frac{2P_3(x)}{A-x} \end{aligned}$$

thus:

$$\begin{aligned} (A-x)^2 f_3'(x) &= 2\left[\frac{A+P_2}{A^2}(A-x)-1\right] \\ f_3'(x) &= 2\frac{A+P_2}{A^2}\frac{1}{A-x} - \frac{2}{(A-x)^2} \\ f_3(x) &= -2\frac{A+P_2}{A^2}ln(A-x) - \frac{2}{(A-x)} + K_3 \end{aligned}$$

leading to:

$$\begin{aligned} P_3(x) &= f_3(x)(A-x)^2 \\ &= (A-x)^2\left[-2\frac{A+P_2}{A^2}ln(A-x) - \frac{2}{(A-x)} + K_3\right] \\ &= -2\frac{A+P_2}{A^2}ln(A-x)(A-x)^2 - 2(A-x) + K_3(A-x)^2 \end{aligned}$$

as $P_3(0) = P_3$

$$P_3 = -2\frac{A+P_2}{A^2}ln(A)A^2 - 2A + K_3A^2 K_3 = \frac{P_3+2A}{A^2} + 2\frac{(A+P_2)}{A^2}ln(A)$$

$$K_3 = \frac{P_3+2A}{A^2} + 2\frac{(A+P_2)}{A^2}ln(A) \tag{41}$$

finally:

$$\begin{aligned} P_3(x) &= -2\frac{A+P_2}{A^2}ln(A-x)(A-x)^2 - 2(A-x) + \frac{P_3+2A}{A^2}(A-x)^2 + 2\frac{(A+P_2)}{A^2}ln(A)(A-x)^2 \\ &= -2\frac{A+P_2}{A^2}ln(A-x)(A-x)^2 + \frac{P_3+2A}{A^2}(A-x)^2 + 2\frac{(A+P_2)}{A^2}ln(A)(A-x)^2 - 2(A-x) \\ &= \left[-2\frac{A+P_2}{A^2}ln(A-x) + \frac{P_3+2A}{A^2} + 2\frac{(A+P_2)}{A^2}ln(A)\right](A-x)^2 - 2(A-x) \\ &= \left[\frac{P_3+2A}{A^2} + 2\frac{(A+P_2)}{A^2}ln(A) - 2\frac{A+P_2}{A^2}ln(A-x)\right](A-x)^2 - 2(A-x) \\ &= \left[\frac{P_3+2A}{A^2} + 2\frac{A+P_2}{A^2}(\ ln(A) - ln(A-x)\ )\right](A-x)^2 - 2(A-x) \end{aligned}$$

that is :

$$P_3(x) = \left[\frac{P_3 + 2A}{A^2} + 2\frac{A + P_2}{A^2}(\ ln(A) - ln(A - x)\ )\right](A - x)^2 - 2(A - x) \tag{42}$$

$$with\ A = 1 - P_0$$

One can be convinced of the validity of this formula looking at the section : Verification of the formula (42) for $P_3(x)$ in the supplementary materials

Before going further we can rearrange this solution in terms of $ln(A - x)$; $(A - x)$; and $(A - x)^2$:

$$P_3(x) = -2\frac{A + P_2}{A^2}ln(A - x)(A - x)^2 - 2(A - x) + \frac{P_3 + 2A}{A^2}(A - x)^2 + 2\frac{(A + P_2)}{A^2}ln(A)(A - x)^2$$

$$= \left[\frac{P_3 + 2A}{A^2} + 2\frac{(A + P_2)}{A^2}ln(A)\right](A - x)^2 - 2\frac{A + P_2}{A^2}ln(A - x)(A - x)^2 - 2(A - x)$$

Recalling the values of K_2 (36) and K_3 (41) we obtain:

$$P_3(x) = K_3(A - x)^2 - 2K_2ln(A - x)(A - x)^2 - 2(A - x) \tag{43}$$

11.1.5 P_4 and $P_{n\geq 3}$

The 3 steps of the diffusion change the probability of having an cluster of 4 atoms in the following way:

$$\begin{aligned} P_4' &= P_4 \\ P_4'' &= P_4' - \frac{P_4'\varepsilon}{1 - P_0'} + \frac{P_3'\varepsilon}{1 - P_0'} \\ P_4''' &= P_4'' - \frac{P_4''\varepsilon}{1 - P_0'} + \frac{P_3''\varepsilon}{1 - P_0'} \end{aligned}$$

Once all the negligible terms are cleared, we obtain the following differential equation:

$$P_4(x)' = \frac{2}{1 - P_0(x)}\left[P_3(x) - P_4(x)\right]$$

We will notice first that, starting from P_3, the differential equations are all of the same form:

$$P_n(x)' = \frac{2P_{n-1}(x)}{A - x} - \frac{2P_n(x)}{A - x} \tag{44}$$

We will also remark that assuming that $P_n(x) = f_n(x)(A - x)^2$ the term $-\frac{2P_n(x)}{A-x}$ appears in $P_n'(x)$:

$$\begin{aligned} P_n(x)' &= \left[f_n(x)(A-x)^2\right]' = f_n'(x)(A-x)^2 - 2f_n(x)(A-x) \\ &= f_n'(x)(A-x)^2 - \frac{2f_n(x)(A-x)^2}{A-x} \\ &= f_n'(x)(A-x)^2 - \frac{2P_n(x)}{A-x} \end{aligned} \tag{45}$$

From (44) and (45) it follows :

$$f_n'(x)(A-x)^2 = \frac{2P_{n-1}(x)}{A-x}$$

that is :

$$f_n'(x)(A-x)^2 = \frac{2f_{n-1}(x)(A-x)^2}{A-x}$$

which yields to the following relationship for $n \geq 3$:

$$f_n'(x) = \frac{2f_{n-1}(x)}{A-x} \tag{46}$$

Starting from $f_2(x)$, multiplying by $\frac{2}{(A-x)}$ and integrating, will give us the solutions we are searching for, and this from the size 2 for which the generic differential equation (44) starts to apply. We will make it simpler by noting that :

$$([ln(A-x)]^n)' = -n\frac{[ln(A-x)]^{n-1}}{A-x}$$

and then:

$$\int \frac{[ln(A-x)]^n}{A-x} = -\frac{1}{n+1}[ln(A-x)]^{n+1}$$

and also, in a less stressful way:

$$\int \frac{1}{(A-x)^2} = \frac{1}{A-x}$$

Thus, the solution sequence begins as follows:

$$f_2(x) = K_2 - \frac{1}{A-x}$$

$$2\frac{f_2(x)}{A-x} = \frac{2K_2}{A-x} - \frac{2}{(A-x)^2}$$
$$f_3(x) = K_3 + \int \frac{2K_2}{A-x} - \int \frac{2}{(A-x)^2}$$
$$= K_3 - 2K_2 ln(A-x) - \frac{2}{A-x}$$

$$2\frac{f_3(x)}{A-x} = \frac{2K_3}{A-x} - \frac{4K_2 ln(A-x)}{A-x} - \frac{4}{(A-x)^2}$$
$$f_4(x) = K_4 + \int \frac{2K_3}{A-x} - \int \frac{4K_2 ln(A-x)}{A-x} - \int \frac{4}{(A-x)^2}$$
$$= K_4 - 2K_3 ln(A-x) + \frac{4K_2[ln(A-x)]^2}{2!} - \frac{4}{A-x}$$

$$2\frac{f_4(x)}{A-x} = \frac{2K_4}{A-x} - \frac{4K_3 ln(A-x)}{A-x} + \frac{8[K_2 ln(A-x)]^2}{2!(A-x)} - \frac{8}{(A-x)^2}$$
$$f_5(x) = K_5 + \int \frac{2K_4}{A-x} - \int \frac{4K_3 ln(A-x)}{A-x} + \int \frac{8K_2[ln(A-x)]^2}{2!(A-x)} - \int \frac{8}{(A-x)^2}$$
$$= K_5 - 2K_4 ln(A-x) + \frac{4K_3[ln(A-x)]^2}{2!} - \frac{8K_2[ln(A-x)]^3}{3!} - \frac{8}{A-x}$$

We notice that from one step to the other the term $[ln(A-x)]^n]$ turns into $\frac{-2}{(n+1)}[ln(A-x)]^{n+1}$, and that the term $\frac{1}{A-x}$ turns into $\frac{2}{A-x}$. This allows us to obtain the following general law:

$$f_{(n\geq 3)}(x) = K_n + \left[\sum_{i=1}^{n-2} K_{n-i}\frac{(-2)^i}{i!}[\ ln(A-x)\ ]^i\right] - \frac{2^{(n-2)}}{A-x} \tag{47}$$

Knowing that $P_n = P_n(0) = f_n(0)A^2$ we deduce the value of the constants :

$$K_{(n\geq 3)} = \frac{P_n + 2^{n-2}A}{A^2} - \sum_{i=1}^{n-2} K_{n-i}\frac{(-2)^i}{i!}[\ ln(A)\ ]^i \tag{48}$$

The correctness of the general formula (47) is proved in the section Validation of the formula (47) for $P_{n\geq 3}(x)$ of supplementary materials.

11.2 Explosion recap

In the end, we will summarize in these few following lines the statistics after the "explosion" of all dimers:

$$
\begin{aligned}
\underline{P}_0 &= P_0 + x_a \\
\underline{P}_1 &= 0 \\
\underline{P}_2 &= 0 \\
\underline{P}_{n\geq 3} &= \left[K_n + \sum_{i=1}^{n-2} K_{n-i} \frac{(-2)^i}{i!} [\ ln(A - x_a)\ ]^i \right] (A - x_a)^2 - 2^{(n-2)}(A - x_a)
\end{aligned}
$$

with

$$
\begin{aligned}
A &= 1 - P_0\ ; \\
K_{(n\geq 3)} &= \frac{P_n + 2^{n-2}A}{A^2} - \sum_{i=1}^{n-2} K_{n-i} \frac{(-2)^i}{i!} [\ ln(A)\ ]^i\ ; \\
K_2 &= \frac{A + P_2}{A^2}\ ;
\end{aligned}
\tag{49}
$$

and

$$
x_a = A - \frac{A^2}{A + P_2}
$$

Where the values of $P_n(x_a)$ are written $\underline{P}_n$.

12 Appendix 3 - Calculations for dimer evaporation

12.1 Modification of the probabilities

In the same manner as we did in the two previous sections, we will determine the variations of the probabilities according to ε and derive laws that apply to all sizes. Always noting « ' » the situation after the first step and « " » the situation after the second step. The situation is the same as for "diffusion" in the sense that, as all sites are eligible to receive a monomer, there will not be the decrease in the possibilities that introduced the logarithms in the case of the "explosion" of dimers.
The transformations of the probabilities can be summarized as follows:

$$\begin{array}{lllll}
P_0 & \longrightarrow & P_0' = P_0 & \longrightarrow & P_0'' = P_0 - \varepsilon P_0 \\
P_1 & \longrightarrow & P_1' = P_1 + \varepsilon & \longrightarrow & P_1'' = P_1 + \varepsilon - \varepsilon(P_1 + \varepsilon) + \varepsilon P_0 \\
P_2 & \longrightarrow & P_2' = P_2 - \varepsilon & \longrightarrow & P_2'' = P_2 - \varepsilon - \varepsilon(P_2 - \varepsilon) + \varepsilon(P_1 + \varepsilon) \\
P_3 & \longrightarrow & P_3' = P_3 & \longrightarrow & P_3'' = P_3 - \varepsilon P_3 + \varepsilon(P_2 - \varepsilon) \\
P_4 & \longrightarrow & P_4' = P_4 & \longrightarrow & P_4'' = P_4 - \varepsilon P_4 + \varepsilon P_3 \\
P_n & \longrightarrow & P_n' = P_n & \longrightarrow & P_n'' = P_n - \varepsilon P_n + \varepsilon P_{n-1}
\end{array} \tag{50}$$

Discarding the negligible terms we have:

$$\begin{aligned}
P_0'' &= P_0 - \varepsilon P_0 \\
P_1'' &= P_1 + \varepsilon - \varepsilon P_1 + \varepsilon P_0 \\
P_2'' &= P_2 - \varepsilon - \varepsilon P_2 + \varepsilon P_1 \\
P_3'' &= P_3 - \varepsilon P_3 + \varepsilon P_2 \\
P_n'' &= P_n - \varepsilon P_n + \varepsilon P_{n-1}
\end{aligned} \tag{51}$$

12.2 Differential equation and their solutions

From the formulas (52) one deduces the following differential equations:

$$\begin{aligned}
P_0'(x) &= -P_0(x) \\
P_1'(x) &= 1 + P_0(x) - P_1(x) \\
P_2'(x) &= -1 + P_1(x) - P_2(x) \\
P_3'(x) &= P_2(x) - P_3(x) \\
P_n'(x) &= P_{n-1}(x) - P_n(x)
\end{aligned} \tag{52}$$

Here we are exactly in the same case as in the section Dimer Diffusion. We have the derivative $P_n'(x)$ which involves $-P_n(x)$ and the natural solution is $P_n(x) = f_n(x)e^{-x}$. In the same way as in the section Dimer Diffusion, we will have:

$$P_n'(x) = f_n'(x)e^{-x} - P_n(x)$$

These differential equations are easily solved:

12.2.1 P_0

$$P_0'(x) = f_0'(x)e^{-x} - P_0(x) = -P_0(x) \Leftrightarrow f_0'(x)e^{-x} = 0$$

The obvious solution is $f_0(x) = P_0$ and then we have:

$$P_0(x) = P_0 e^{-x} \tag{53}$$

12.2.2 P_1

We have to solve :

$$P_1'(x) = 1 + P_0(x) - P_1(x)$$

assuming that $P_1(x) = f_1(x)e^{-x}$, The differential equation For P_1 can be summarized as:

$$f_1'(x)e^{-x} = 1 + P_0(x)$$

that is not difficult to solve:

$$\begin{aligned} f_1'(x)e^{-x} &= 1 + P_0 e^{-x} \\ f_1'(x) &= e^x + P_0 \end{aligned}$$

leading trivially to:

$$f_1(x) = K_1 + e^x + xP_0$$

and

$$P_1(x) = K_1 e^{-x} + 1 + xP_0 e^{-x}$$

Of course, K_1 is such as $P_1(0) = P_1$

$$\begin{aligned} P_1 &= K_1 + 1 \\ K_1 &= P_1 - 1 \end{aligned}$$

finally:

$$P_1(x) = 1 - [1 - P_1 - xP_0]e^{-x} \tag{54}$$

12.2.3 P_2

For P_2, we have:

$$P_2'(x) = -1 + P_1(x) - P_2(x)$$

$$\begin{aligned} f_2'(x)e^{-x} &= -1 + P_1(x) \\ f_2'(x) &= -e^x + \left\{1 - [1 - P_1 - xP_0]e^{-x}\right\} e^x \\ &= -e^x + e^x - [1 - P_1 - xP_0] \\ &= P_1 + xP_0 - 1 \end{aligned}$$

In a trivial way :

$$f_2(x) = K_2 + xP_1 + \frac{1}{2}x^2 P_0 - x$$

and simply $K_2 = P_2$, which results in:

$$P_2(x) = \left[P_2 + xP_1 + \frac{1}{2}x^2P_0 - x\right]e^{-x} \tag{55}$$

12.2.4 $P_{n\geq 3}$

From P_3, the differential equations are all the same: $(P_n'(x) = P_{n-1}(x) - P_n(x))$. Moreover, $P_2(x)$ has the form of a simple polynomial in powers of x multiplied by a factor e^{-x}. The solution: $f_n'(x)e^{-x} = \mathcal{P}_{n-1}(x)e^{-x}$ will always lead to $f_n'(x) = \mathcal{P}_{n-1}(x)$ which require $f_n(x)$ to be a polynomial as well: $f_n(x) = \mathcal{P}_n(x)$. The successive solutions of $f_n(x)$ are therefore polynomials whose constant will necessarily be equal to P_n since all the other terms are factors of a power of x and that $e^0 = 1$.

$$\begin{aligned}
P_2(x) &= \left[P_2 \quad + \quad xP_1 \quad + \quad \frac{1}{2}x^2P_0 \quad - \quad x \quad\right]e^{-x} \\
P_3(x) &= \left[P_3 \quad + \quad xP_2 \quad + \quad \frac{1}{2}x^2P_1 \quad + \quad \frac{1}{3!}x^3P_0 \quad - \quad \frac{1}{2}x^2\right]e^{-x} \\
P_4(x) &= \left[P_4 \quad + \quad xP_3 \quad + \quad \frac{1}{2}x^2P_2 \quad + \quad \frac{1}{3!}x^3P_1 \quad + \quad \frac{1}{4!}x^4P_0 \quad - \quad \frac{1}{3!}x^3\right]e^{-x}
\end{aligned} \tag{56}$$

And the appropriate generalization for $P_n(x)$ is:

$$P_{n\geq 2}(x) = \left[\sum_{i=0}^{n} P_{n-i}\frac{x^i}{i!} \quad - \quad \frac{x^{(n-1)}}{(n-1)!}\right]e^{-x} \tag{57}$$

12.3 halting evaporation

Dimer evaporation stops naturally when all dimers have finally disappeared. This condition is achieved when $P_2(x) = 0$.

$$P_2(x) = 0 \Leftrightarrow P_2 + xP_1 + \frac{1}{2}x^2P_0 - x = 0 \tag{58}$$

There are two solutions to annul $P_2(x)$

$$x = \frac{(1-P_1) \pm \sqrt{(1-P_1)^2 - 2P_0P_2}}{P_0}$$

We will notice two things. First, $1 - P_1 \geq 0$ and $P_0P_2 \geq 0$, and if $(1-P_1)^2 - 2P_0P_2 \geq 0$, then $\sqrt{(1-P_1)^2 - 2P_0P_2} \leq (1-P_1)$. From these remarks it follows that:

$$\frac{(1-P_1) - \sqrt{(1-P_1)^2 - 2P_0P_2}}{P_0} \leq \frac{(1-P_1) + \sqrt{(1-P_1)^2 - 2P_0P_2}}{P_0}$$

The solution we are looking for is therefore the first one encountered, i.e.:

$$x_a = \frac{(1-P_1) - \sqrt{(1-P_1)^2 - 2P_0P_2}}{P_0} \tag{59}$$

Note that if P_0 tends towards 0, then we can either do a Taylor expansion of $\sqrt{(1-P_1)^2 - 2P_0P_2}$:

$$\lim_{P_0\to 0}\sqrt{(1-P_1)^2 - 2P_0P_2} = (1-P_1) - \frac{2P_0P_2}{2\sqrt{(1-P_1)^2}} = (1-P_1) - \frac{P_0P_2}{(1-P_1)}$$

leading to:

$$\lim_{P_0 \to 0} x_a = \frac{(1-P_1) - (1-P_1) + \frac{P_0 P_2}{(1-P_1)}}{P_0} = \frac{P_2}{(1-P_1)}$$

Simpler, we can take the condition 58 and set $P_0 = 0$:

$$P_2 + x_a P_1 - x_a = 0$$

Leading by chance to the same formula:

$$x_a = \frac{P_2}{1-P_1}$$

A priori, if the probability of having dimers is such that we use these formulas, the probability of having empty sites will certainly not be equal to 0.

12.4 Summary of the evaporation

The probabilities after evaporation of the dimers are summarized by the following formulas:

$$\begin{aligned}
\underline{P}_0 &= P_0 e^{-x} \\
\underline{P}_1 &= 1 - [1 - P_1 - x_a P_0] e^{-x_a} \\
\underline{P}_{n\geq 2} &= \left[\sum_{i=0}^{n} P_{n-i} \frac{x_a^i}{i!} \quad - \quad \frac{x_a^{(n-1)}}{(n-1)!} \right] e^{-x_a}
\end{aligned} \tag{60}$$

with

$$\begin{aligned}
x_a &= \frac{(1-P_1) - \sqrt{(1-P_1)^2 - 2P_0 P_2}}{P_0} \; if \; P_0 \neq 0 \\
x_a &= \frac{P_2}{1-P_1} \; if \; P_0 = 0
\end{aligned}$$