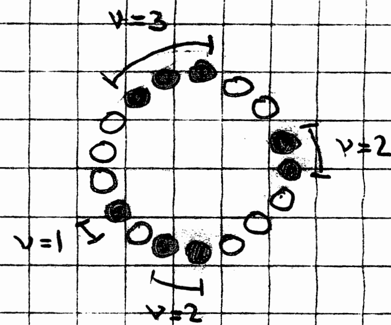
Equilibrium Properties of the Diffusion Problem

(a) The block number

Define the number of contiguous blocks of atoms within a pattern as the block number m (see figure)



Here: $p=17$, $k=8$, $m=4$.

Note: The number of blocks m, and the number of gaps between blocks must be identical.

Since m does not change under rotations of the entire pattern, the block number is an invariant for each primitive pattern.

Under individual jumps of types A–D, the block number behaves according to the following table:

Type A	(motion of a single atom block) :	$\Delta m = 0$
B	(creation of a single atom block) :	$\Delta m = +1$
C	(fusion of a single atom block) :	$\Delta m = -1$
D	(exchange of atom between blocks) :	$\Delta m = 0$

Note that transitions between different patterns are limited by the selection rule $|\Delta m| \leq 1$.

The block number is invariant under reflections.

(b) The distribution of primitives among block numbers

In the following, we determine a formula for the number of primitive patterns $N(p,k,m)$ that share the same block number m.

(i) The generating function method

Each of the m blocks must contain at least one atom, and the total population must equal the number of atoms k in the problem. Now consider the geometric sum:

$$x + x^2 + x^3 + \ldots = \frac{x}{1-x} \tag{1}$$

The sum represents a single block, and the terms $x, x^2, x^3, \ldots$ stand for $1, 2, 3, \ldots$ atoms in this block. As expansion into a sum shows, the product of m of these factors:

$$(x + x^2 + x^3 + \ldots) \cdot \ldots \cdot (x + x^2 + x^3 + \ldots) = \frac{x^m}{(1-x)^m} \tag{2}$$

contains a term $x^{\nu_1} x^{\nu_2} \ldots x^{\nu_m} = x^{\nu_1 + \nu_2 + \ldots + \nu_m}$ for a combination of ν_1 atoms in block 1, ν_2 atoms in block 2, etc. The number of different combinations that contain exactly $\nu_1 + \nu_2 + \ldots + \nu_m = k$ atoms is therefore the coefficient of x^k in the product (2):

$$\frac{x^m}{(1-x)^m} = \sum_{\nu=m}^{\infty} C(\nu, m)\, x^{\nu} \tag{3}$$

where $C(\nu, m)$ is the number of options to sort ν identical atoms into m blocks so that none of the blocks is empty. By expansion into a binomial series,

$$\frac{x^m}{(1-x)^m} = \sum_{\mu=0}^{\infty} \binom{m-1+\mu}{m-1} x^{\mu+m} \tag{4}$$

and therefore, as $k = \mu + m$,

$$C(k, m) = \binom{k-1}{m-1}. \tag{5}$$

There are $\binom{k-1}{m-1}$ possibilities to sort k atoms into m blocks.

In the same vein, the remaining $p-k$ empty sites ("holes") must form m contiguous blocks of at least one hole each. Correspondingly, there are

$$C(p-k,m) = \binom{p-k-1}{m-1} \tag{6}$$

options to perform this task. The product of (5) and (6) then yields the total number of combinations. However, since the blocks and gaps are located on a circle (see figure), the assignment of a block as "first block" is arbitrary; only the order of the block-gap-sequence is fixed. Rotation of the block-gap-sequence thus leads to exactly m members of each primitive equivalence class in this product. (Note that for a prime number of sites, a rotation of a pattern by a nontrivial ($\neq 2k\pi$) angle always yields an equivalent yet distinct pattern.)

It is therefore simple to remove the overcounting by division, and the desired number of primitives that show m blocks reads:

$$\begin{aligned} N(p,k,m) &= \frac{1}{m}\, C(k,m)\, C(p-k,m) \\ &= \frac{1}{m}\binom{k-1}{m-1}\binom{p-k-1}{m-1} \end{aligned} \tag{7}$$

Note that there are at least one and at most $\min(k,p-k)$ blocks of atoms in a pattern with k atoms. Hence, the numbers $N(p,k,m)$ must sum up to the total number of primitives:

$$\sum_{m=1}^{\min(k,p-k)} \frac{1}{m}\binom{k-1}{m-1}\binom{p-k-1}{m-1} = \frac{1}{p}\binom{p}{k} \tag{8}$$

(ii) The combinatorial method

In the following, we present an alternative derivation of $C(k,m)$ (5) that founds on a combinatorial argument instead. We look at the number of ways to put k objects (the atoms) into m non-empty boxes (the contiguous blocks), where the objects are indistinguishable. This numbers can be established in the following procedure:

(α) First, remove m objects from the set of k atoms and put one into each box (to render the blocks non-empty)

(β) Line up the remaining $k-m$ atoms and introduce $m-1$ separators. Example ($k=10$, $m=4$):

$$\bullet\ \bullet\ |\ \bullet\ \bullet\ \bullet\ |\ |\ \bullet \qquad (9)$$

separators

The atoms left to the first separator enter box #1, the atoms between the first and second separator are destined for box #2, etc.; the atoms right to the last separator end up in box #m.

Each setting of the $m-1$ separators between the $k-m$ atoms designates a unique distribution of atoms among the blocks.

(γ) The separators occupy $(m-1)$ sites in the scheme (9), while the atoms occupy $k-m$ sites of a total of $(k-m)+(m-1) = (k-1)$ sites. The number of different settings in the arrangement (9) thus equals the number of ways to pick $(m-1)$ numbers in a sequence of length $(k-1)$:

$$C(k,m) = \binom{k-1}{m-1} \qquad (5')$$

as already established in (5).

(c) The block number of a configuration determines its average occupation in equilibrium

Interestingly, the equilibrium distribution of configurations, unlike the kinetics of the problem, depends on the jump rates A, B, C, D merely by the ratio B/C, as the following theorem states:

Theorem. The relative weight $p(\mathfrak{z})$ of a configuration $\mathfrak{z}$ of k atoms on p sites depends only on its block number m and the ratio of jump rates B/C: If two configurations $\mathfrak{z}, \mathfrak{z}'$ possess block numbers m, m', then the ratio of their respective probabilities in equilibrium is given by:

$$\frac{p(\mathfrak{z}')}{p(\mathfrak{z})} = \left(\frac{B}{C}\right)^{m'-m} \tag{10}$$

We perform the proof of this assertion in a number of steps. We first take a look at the rate matrix A which details the temporal evolution of the system. If a jump exists that connects the initial configuration $\mathfrak{z}$ with some final configuration $\mathfrak{z}'$ of the system, then the matrix element $A_{\mathfrak{z}'\mathfrak{z}}$ is given by the rate of the jump, and otherwise vanishes. The diagonal elements $A_{\mathfrak{z}\mathfrak{z}}$ account for the losses, so that:

$$\frac{d}{dt} p(\mathfrak{z}') = \sum_{\mathfrak{z}} A_{\mathfrak{z}'\mathfrak{z}}\, p(\mathfrak{z}) \tag{11}$$

Since the total probability is conserved, $\sum_{\mathfrak{z}'} p(\mathfrak{z}') = 1$, we find that the derivative of the latter expression is zero, or equivalently, that the column sum of A must vanish:

$$\sum_{\mathfrak{z}'} A_{\mathfrak{z}'\mathfrak{z}} = 0 \qquad \text{for all } \mathfrak{z}. \tag{12}$$

The off-diagonal elements of the rate matrix fulfil simple symmetry rules required by reciprocity. We note that $A_{\mathfrak{z}\mathfrak{z}'}$ and $A_{\mathfrak{z}'\mathfrak{z}}$ $(\mathfrak{z} \neq \mathfrak{z}')$ describe the same jump, just with reversed directions. Now the reverse of jumps of type A and D (motion of a single atom block, exchange of an atom between

blocks) is again a jump of the same respective type, while jumps of types B and C (separation of an atom from a block, fusion of an atom to a block) are reverses of each other and exchange their roles.
We summarize the symmetry rules in a table:

$A_{\mathfrak{z}\mathfrak{z}'}$:	A	B	C	D
$A_{\mathfrak{z}'\mathfrak{z}}$:	A	C	B	D

(If $A_{\mathfrak{z}\mathfrak{z}'} = 0$, then also $A_{\mathfrak{z}'\mathfrak{z}} = 0$.)

According to the selection rules listed in Section (a), the type of jump recorded in $A_{\mathfrak{z}'\mathfrak{z}}$ is severely restricted by the block numbers m, m' of the involved configurations $\mathfrak{z}, \mathfrak{z}'$. If $m = m'$, only jumps of type A or D may connect the configurations. For $m' = m+1$, the jump must be of type B, for $m' = m-1$, it must be of type C, and for $|\Delta m| > 1$, no direct transition between $\mathfrak{z}$ and $\mathfrak{z}'$ is possible.
We thus find:

$\Delta m = m' - m$	0	+1	−1	±2, etc.
$A_{\mathfrak{z}'\mathfrak{z}}$	0, A, D	0, B	0, C	0
$A_{\mathfrak{z}\mathfrak{z}'}$	0, A, D	0, C	0, B	0

From this table, we infer that:

$$A_{\mathfrak{z}'\mathfrak{z}} = \left(\frac{B}{C}\right)^{\Delta m} A_{\mathfrak{z}\mathfrak{z}'} \tag{13}$$

holds for all off-diagonal matrix elements. After summation over all initial configurations $\mathfrak{z}$, we therefore obtain from (12) and (13):

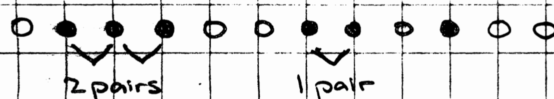
$$\sum_{\mathfrak{z}} A_{\mathfrak{z}'\mathfrak{z}} \left(\frac{B}{C}\right)^{m(\mathfrak{z})} = \sum_{\mathfrak{z}} A_{\mathfrak{z}\mathfrak{z}'} \left(\frac{B}{C}\right)^{m(\mathfrak{z}')}$$
$$= \left(\frac{B}{C}\right)^{m(\mathfrak{z}')} \sum_{\mathfrak{z}} A_{\mathfrak{z}\mathfrak{z}'} = 0 \tag{14}$$

According to (11),

$$p(\mathfrak{z}) = \left(\frac{B}{C}\right)^{m(\mathfrak{z})}$$

is therefore an (unnormalized) equilibrium eigenvector of the system, and the assertion follows at once.

Physical interpretation: Assign a "cohesion energy" E to each neighbouring pair of atoms. The number of pairs in each contiguous block is the number of atoms in the block, less one:



Therefore, the total number of pairs is the number of atoms, less the number of blocks, $k-m$, and the associated energy of the configuration $(k-m)\cdot E$. The equilibrium probability of a configuration $p(\zeta)$ thus only depends on the block number m:

$$\frac{p(\zeta')}{p(\zeta)} = \frac{e^{-(k-m')\beta E}}{e^{-(k-m)\beta E}} = e^{(m'-m)\beta E} \tag{15}$$

Detailed balance now requires $A_{\zeta'\zeta}\, p(\zeta) = A_{\zeta\zeta'}\, p(\zeta')$; using the symmetry property of the elements $A_{\zeta\zeta'}$ established above, we infer for $m' = m+1$ that $A_{\zeta\zeta} = B$, $A_{\zeta\zeta'} = C$ and thus:

$$\frac{p(\zeta')}{p(\zeta)} = \frac{A_{\zeta'\zeta}}{A_{\zeta\zeta'}} = \frac{B}{C} = e^{\beta E} \tag{16}$$

Therefore, $B/C = e^{\beta E}$ follows, and (15) immediately yields the statement (10) of the theorem again.

(d) The partition function

It proves convenient to introduce a quantity akin to the partition function of statistical mechanics by summing over the (unnormalized) equilibrium averages of all primitive atomic configurations:

$$Z = \sum_{\text{primitive } z} \left(\frac{B}{C}\right)^{m(z)} = \sum_{m=1}^{\min(k,p-k)} N(p,k,m) \left(\frac{B}{C}\right)^{m} \tag{17}$$

Here, we used the relative weight of the configurations (10) together with their classification by the block numbers m. (see Section (b).)

For a closed representation of the partition function, we first express $N(p,k,m)$ (7) by Pochhammer symbols:

$$N(p,k,m) = \frac{(k-1)!\,(p-k-1)!}{m!\,(k-m)!\,(m-1)!\,(p-k-m)!}$$

$$= \frac{1}{(m-1)!} \frac{(1-k)(2-k)\cdot\ldots\cdot(m-1-k)\,(1-p+k)(2-p+k)\cdot\ldots\cdot(m-1-p+k)}{2\cdot 3\cdot 4\cdot\ldots\cdot(m-1)\cdot m}$$

$$= \frac{1}{(m-1)!} \frac{(1-k)_{m-1}\,(1-p+k)_{m-1}}{(2)_{m-1}} \tag{18}$$

where $(a)_n = a\cdot(a+1)\cdot(a+2)\cdot\ldots\cdot(a+n-1)$, and $(a)_0 = 1$ by definition. Inserting (18) in (17), and setting $g = m-1$, we find that Z can be recast as a Gaussian hypergeometric function [see e.g. Abramowitz / Stegun, Handbook of Mathematical Functions, Ch. 15]:

$$Z = \sum_{g=0}^{\infty} \frac{1}{g!} \frac{(1-k)_g\,(1-p+k)_g}{(2)_g} \left(\frac{B}{C}\right)^{g+1}$$

$$= \frac{B}{C}\, {}_2F_1\left(1-k,\, 1-p+k,\, 2;\, B/C\right) \tag{19}$$

For the special case $B=C$, all configurations are equally likely, and the partition function simply reflects the number of different primitive atomic arrangements. By using the identity [AS 15.1.20]:

$$Z = {}_2F_1(1-k,\, 1-p+k,\, 2;\, 1) = \frac{\Gamma(2)\,\Gamma(p)}{\Gamma(k+1)\,\Gamma(p-k+1)} \tag{20}$$

we indeed confirm the sum rule (8):

$$Z = \frac{(p-1)!}{k!\,(p-k)!} = \frac{1}{p}\binom{p}{k} \tag{21}$$

(e) Equilibrium probabilities of configurations and patterns

Since every primitive pattern actually represents p congruent configurations $\mathfrak{z}$ that are merely rotated with respect to each other, the normalized equilibrium average for a configuration $\mathfrak{z}$ depends on its block number $m(\mathfrak{z})$ via:

$$p(\mathfrak{z}) = \frac{(B/C)^{m(\mathfrak{z})}}{p \cdot Z} = \frac{1}{p} \frac{(B/C)^{m-1}}{{}_2F_1(1-k, 1-p+k, 2; B/C)} \tag{22}$$

For a pattern that only partly specifies the configuration (i.e., that allows the occurrence of both atoms or holes on some sites), the equilibrium average follows after summation over all configurations that match the pattern. If we denote the number of configurations belonging to a primitive equivalence class $\mathfrak{z}_{prim}$ that match the indefinite pattern η with the "matching number" $\mathcal{M}(\eta, \mathfrak{z})$ (which obviously ranges between 0 and p), then the average occupation of η will be:

$$p(\eta) = \frac{1}{p \cdot Z} \sum_{\text{primitive } \mathfrak{z}} \mathcal{M}(\eta, \mathfrak{z}) \left(\frac{B}{C}\right)^{m(\mathfrak{z})} \tag{23}$$

In several simple cases, the sum may be expressed in closed form, yielding an analytical result for the pattern likelihood $p(\eta)$ (see section below). Otherwise, the matching number must be evaluated by comparison of all possible configurations with the pattern, and the sum performed numerically.

For illustration, we study the average occupation of a single site, i.e., $\eta = \bullet \times \times \ldots \times$, or for short, $\eta = \bullet$. Clearly, for each atom in every primitive configuration, there will be a matching rotated configuration, so the matching number is constant: $\mathcal{M}(\eta, \mathfrak{z}) = k$. Insertion into (23), and comparison with the definition of Z (17) then yields:

$$p(\bullet) = \frac{1}{p \cdot Z} \sum_{\text{prim. } \mathfrak{z}} k \cdot \left(\frac{B}{C}\right)^{m(\mathfrak{z})} = \frac{k \cdot Z}{p \cdot Z} = k/p, \tag{24}$$

the result trivially expected from the equivalence of sites.

(f) Averages that allow expression in closed form

In the following, we examine some important averages that can be cast into closed form as quotients of hypergeometric functions. To this end, we initially write the sum (23) in slightly different form:

$$p(\eta) = \frac{1}{p \cdot Z} \sum_{m=1}^{\min(k,p-k)} \left(\frac{B}{C}\right)^m \sum_{\zeta_m} \mathcal{M}(\eta, \zeta) \qquad (23')$$

where the latter sum is performed over all primitives ζ whose atoms are arranged in m contiguous blocks. The method thus is an extension of the summation technique used to establish the partition function Z in Section (d).

(i) The averages ⟨o●●…●⟩

Clearly, the challenge in (23′) is to enumerate all possible matches between the configurations with m blocks and the comparison pattern ⟨o●●…●⟩. We again use the technique of the generating function introduced in Section (b).

Let us denote the length of the string of atoms in ⟨o●●…●⟩ by r. Then, in any given primitive arrangement ζ_m with m atomic blocks, the matching number $\mathcal{M}(\eta, \zeta)$ will equal the number of blocks of length $\geq r$. We are therefore looking for all combinations of block lengths $\nu_1, \nu_2, \ldots, \nu_m$ where $\nu_1 + \nu_2 + \ldots + \nu_m = k$ and $\nu_1 \geq r$, or $\nu_2 \geq r$. In analogy to the method put forward in Section (b), the number of combinations that have at least r atoms in the first block is the coefficient of x^k in:

$$(x^r + x^{r+1} + \ldots) \cdot (x + x^2 + x^3 + \ldots)^{m-1} = \frac{x^{m+r-1}}{(1-x)^m} \qquad (25)$$

Clearly, the number of combinations that have at least r atoms in block $2, 3, \ldots, m$ is exactly the same, so the generating function for the desired number of r-blocks in all configurations with block number m is:

$$\frac{m\, x^{m+r-1}}{(1-x)^m} = m \sum_{\mu=0}^{\infty} \binom{m-1+\mu}{m-1} x^{m+r-1+\mu} = \sum_{\nu=m+r-1}^{\infty} C_r(\nu, m)\, x^\nu \qquad (26)$$

Hence, we obtain for this number $C_r(k,m)$:

$$C_r(k,m) = m\binom{k-r}{m-1} \tag{27}$$

Since we are free to choose any combination of lengths for the m gaps between the atomic blocks (as each gap contains at least one empty site), each arrangement of blocks incidentally represents $C(p-k,m)$ atomic configurations (6), so the total number of matches in configurations with m blocks is $C_r(k,m)\,C(p-k,m)$. However, the assignment of a block as the "first" block etc. in (25) is artificial (only the sequence of blocks is fixed). To remove the duplicate states generated by rotation of the block pattern, we simply divide by the number of blocks m. We then obtain for the sum in (23'):

$$\sum_{z_m} \mathcal{M}(\eta,z) = \frac{1}{m} C_r(k,m)\,C(p-k,m)$$

$$= \binom{k-r}{m-1}\binom{p-k-1}{m-1} \tag{28}$$

This coefficient can be formulated in terms of Pochhammer symbols $(a)_n$ again (cf. (18), (19)):

$$\sum_{z_m} \mathcal{M}(\eta,z) = \frac{1}{(m-1)!}\,\frac{(r-k)_{m-1}\,(1-p+k)_{m-1}}{(1)_{m-1}} \tag{29}$$

which allows for summation as a hypergeometric series:

$$\sum_{m=1}^{\min(k,p-k)} \left(\frac{B}{C}\right)^m \sum_{z_m} \mathcal{M}(\eta,z) = \frac{B}{C}\,{}_2F_1\!\left(r-k,\,1-p+k,\,1;\,\frac{B}{C}\right) \tag{30}$$

Equation (23), together with (19), then yields $(r=1,2,\dots,k)$:

$$\langle \circ\bullet\bullet\dots\bullet \rangle = \frac{1}{p}\,\frac{{}_2F_1(r-k,\,1-p+k,\,1;\,B/C)}{{}_2F_1(1-k,\,1-p+k,\,2;\,B/C)} \tag{31}$$

By particle-hole symmetry, also:

$$\langle \bullet\circ\circ\dots\circ \rangle = \frac{1}{p}\,\frac{{}_2F_1(1-k,\,r-p+k,\,1;\,B/C)}{{}_2F_1(1-k,\,1-p+k,\,2;\,B/C)} \tag{32}$$

In the case $B=C$ (all configurations carry equal weight), the hypergeometric series reduces to a simple product [AS 15.1.20]:

$$ {}_2F_1(a,b,c;1) = \frac{\Gamma(c)\,\Gamma(c-a-b)}{\Gamma(c-a)\,\Gamma(c-b)} \tag{33} $$

The average (31) reduces then to:

$$ \langle \circ\bullet\bullet\cdots\bullet\rangle = \frac{1}{p}\,\frac{k!\,(p-k)!\,(p-r-1)!}{(p-1)!\,(p-k-1)!\,(k-r)!} = \frac{\binom{p-r-1}{k-r}}{\binom{p}{k}} \tag{34} $$

Indeed, the number of all configurations is $\binom{p}{k}$, and the number of arrangements that start with $\circ\bullet\bullet\cdots\bullet$ are all possible configurations of the $k-r$ remaining atoms on the $p-r-1$ sites not considered so far, i.e., $\binom{p-r-1}{k-r}$ combinations, as stated by (34).

In the limit $B/C \to 0$, all atoms clump together – only the single primitive configuration with $m=0$ (7) is occupied, and only one of the p rotationally equivalent states will match the pattern. Hence, we find:

$$ \lim_{B/C\to 0} \langle \circ\bullet\bullet\cdots\bullet\rangle = \frac{1}{p} \tag{35} $$

Since ${}_2F_1(a,b,c;0) = 1$, this is confirmed by (31).

(ii) The averages $\langle \bullet\bullet\cdots\bullet\rangle$

Next, we examine the equilibrium average to find a contiguous block of atoms of length r. We again use the method of generating functions (Section (6)). First, consider a block of length $\nu \geq r$:

$$ \circ\circ\bullet\bullet\bullet\bullet\bullet\circ\circ\ldots $$

As illustrated, there are 3 possibilities to match a 3-block into a sequence of 5 atoms. In general, there are $\nu-r+1$ ways to fit a r-block into a block of length ν.

As usual, the remaining $m-1$ blocks of atoms have no restriction with regard to their respective lengths, as do the m blocks of holes, except that the total atom and hole numbers must add up to k and $p-k$. Thus, the appropriately weighed generating function is [cf. (2)]:

$$m(x^r + 2x^{r+1} + 3x^{r+2} + \dots)(x + x^2 + x^3 + \dots)^{m-1}(y + y^2 + y^3 + \dots)^m \qquad (36)$$

where the variables x and y represent atoms and holes, respectively. Now, we have:

$$\begin{aligned} x^r + 2x^{r+1} + 3x^{r+2} + \dots &= x^r(1 + 2x + 3x^2 + \dots) \\ &= x^r \frac{\partial}{\partial x}(x + x^2 + x^3 + \dots) \\ &= x^r \frac{\partial}{\partial x}\left(\frac{1}{1-x}\right) = \frac{x^r}{(1-x)^2} \end{aligned} \qquad (37)$$

and therefore the generating function (36) becomes:

$$m \frac{x^r}{(1-x)^2}\left[\frac{x}{1-x}\right]^{m-1}\left[\frac{y}{1-y}\right]^m = m \frac{x^{r+m-1}}{(1-x)^{m+1}} \frac{y^m}{(1-y)^m} \qquad (38)$$

By comparison with (4), the number of matches within the subset of primitive configurations with m blocks (23′) is given by the coefficient of $x^k y^{p-k}$ in the power series of (38):

$$\sum_{\mathfrak{Z}_m} \mathcal{M}(\eta, \mathfrak{z}) = \binom{k-r+1}{m}\binom{p-k-1}{m-1} \qquad (39)$$

(This result is already corrected for the m-fold rotational degeneracy discussed in Section (b) and above.) Again, (39) is favorably expressed in terms of Pochhammer symbols (see eq. (18)):

$$\sum_{\mathfrak{Z}_m} \mathcal{M}(\eta, \mathfrak{z}) = -\frac{(r-k-1)_m (1+k-p)_{m-1}}{(m-1)!\,(1)_m}$$

$$= \frac{k-r+1}{(m-1)!} \frac{(r-k)_{m-1}(1+k-p)_{m-1}}{(2)_{m-1}} \qquad (40)$$

The sum (23') thus yields, similarly to (30):

$$\langle \bullet\bullet\ldots\bullet \rangle = \frac{k-r+1}{p} \frac{{}_2F_1(r-k,1+k-p,2;B/C)}{{}_2F_1(1-k,1+k-p,2;B/C)} \tag{41}$$

In particular, $\langle\bullet\rangle = \frac{k}{p}$, as argued in (24). For $B=C$, the result simplifies as in (33), (34):

$$\langle \bullet\bullet\ldots\bullet \rangle = \frac{k-r+1}{p} \frac{k!\,(p-k)!\,(p-r)!}{(k-r+1)!\,(p-k)!\,(p-1)!} = \frac{\binom{p-r}{k-r}}{\binom{p}{k}} \tag{42}$$

This is expected from combinatorics.

(iii) <u>The averages $\langle \circ\bullet\bullet\ldots\bullet\circ \rangle$</u>

Similarly, we study the probability to find an isolated block of atoms of length r. Since the length of one block is fixed, but the lengths of all other atomic blocks and the gaps in between them are not constrained, the generating function here reads [cf. (2), (25), (36)]:

$$m \cdot x^r \cdot \left[\frac{x}{1-x}\right]^{m-1} \left[\frac{y}{1-y}\right]^m \tag{43}$$

and the coefficient of $x^k y^{p-k}$ in its power series is the desired number of matches within the m-block subset. After eliminating the m-fold degeneracy, one finds using (4):

$$\sum_{3m} M(\eta,\mathfrak{z}) = \binom{k-r-1}{m-2}\binom{p-k-1}{m-1}^{*} \tag{44}$$

$$= -\frac{1}{(m-2)!} \frac{(1+k-p)_{m-1}(1+r-k)_{m-2}}{(1)_{m-1}}$$

$$= \frac{1}{(m-2)!}(p-k-1)\frac{(2+k-p)_{m-2}(1+r-k)_{m-2}}{(2)_{m-2}} \tag{45}$$

The summation in (23) again yields a hypergeometric series (19):

$$\langle \circ\bullet\bullet\ldots\bullet\bullet\circ \rangle = \frac{p-k-1}{p}\,\frac{B}{C}\,\frac{{}_2F_1(1+r-k,2+k-p,2;B/C)}{{}_2F_1(1-k,1+k-p,2;B/C)} \tag{46}$$

---

* This formula is valid only for $m \geq 2$ or $k \neq r$. If $k=r$ and $m=1$, the sum yields unity, as comparison with (43) shows.

Alternatively, this result may also be obtained from the earlier result (31) in an entirely different fashion. Note that

$$\langle \circ \underbrace{\bullet\bullet\cdots\bullet}_{r} \rangle = \langle \circ \underbrace{\bullet\bullet\cdots\bullet\bullet}_{r+1} \rangle + \langle \circ \underbrace{\bullet\bullet\cdots\bullet}_{r}\circ \rangle \qquad (47)$$

(The site immediately right to the sequence on the left-hand side will host either an atom or a hole). Since two of the three averages are known, $\langle \circ\bullet\bullet\cdots\bullet\circ\rangle$ follows by subtraction:

$$\langle \circ \underbrace{\bullet\cdots\bullet}_{r}\circ \rangle = \frac{{}_2F_1(r-k, 1-p+k, 1; \frac{B}{C}) - {}_2F_1(r+1-k, 1-p+k, 1; \frac{B}{C})}{p\, {}_2F_1(1-k, 1-p+k, 2; B/C)} \qquad (48)$$

The equivalence of (46) and (48) is established using the contiguity relations between hypergeometric functions (see AS, Ch. 15.2). We first use [AS 15.2.17]:

$$(c-a-1)\, {}_2F_1(a,b,c;z) + a\, {}_2F_1(a+1,b,c;z) =$$
$$= (c-1)\, {}_2F_1(a,b,c-1,z) \qquad (49)$$

For our case $c=1$, the right-hand side must be replaced by its limiting value [AS 15.1.2]:

$$\lim_{c\to 1} \frac{1}{\Gamma(c-1)}\, {}_2F_1(a,b,c-1;z) =$$
$$= \lim_{c\to 1} (c-1)\, {}_2F_1(a,b,c-1;z) = abz\, {}_2F_1(a+1,b+1,2;z) \qquad (50)$$

When $c=1$, (49) therefore becomes $(a \neq 0)$:

$${}_2F_1(a,b,1;z) - {}_2F_1(a+1,b,1;z) = -bz\, {}_2F_1(a+1,b+1,2;z) \qquad (51)$$

Applying this relation to (48) yields the original formula (46) at once.

(iv) <u>Averages of compact patterns</u>

The foregoing ideas can be extended to embrace all compact patterns, i.e., any specified contiguous sequence of ν atoms and μ empty sites serving as a comparison pattern η.

Examples include:

⟨●●●○○●○●●○○⟩, ⟨●○○●●○●●⟩, etc.

Generally, these sequences are made up from a left and right open end sequence of sites of one kind (either atoms or holes), and several enclosed blocks of atoms and holes. To match the comparison pattern, the atomic arrangement must possess a sequence of atom and hole blocks that exactly fits the pattern of enclosed blocks in η, while the blocks bordering this interior sequence must be at least as long as the respective end sequences in the comparison pattern. Thus, compact patterns combine elements of case (i) (the open-ended patterns ⟨○●●…●⟩) and (iii) (the enclosed patterns ⟨○●●…●⟩).

We may discern two different classes among the compact patterns:

(α) Patterns whose end sequences are of different types (one end atoms, the other end holes)

(β) Patterns whose end sequences are of the same type (both atoms or both holes)

The patterns examined in (i) and (iii) are thus the simplest representatives of both classes. For the sake of simplicity, here we assume that both end sequences in type (β) are of the atom type.

Let us denote the number of enclosed atom blocks in the pattern by σ, and the number of enclosed hole blocks by τ. Alternation between atoms and holes then requires that:

$$\sigma = \tau \quad [\text{type } (\alpha)], \qquad \sigma = \tau - 1 \quad [\text{type } (\beta)] \tag{52}$$

For the moment, we designate the lengths of every participating block in the comparison pattern (enclosed and open-end blocks alike) by ν_a (atoms) and μ_a (holes), but it will turn out that only the total numbers ν of atoms and μ of holes enters the result. In the now familiar fashion, we set up the generating function [cf. (25), (43)]. We first examine patterns of type (α), and assume that the pattern starts with an atomic sequence:

$$m\,(x^{\nu_0}+x^{\nu_0+1}+\dots)\,y^{\mu_1}x^{\nu_1}y^{\mu_2}x^{\mu_2}\dots x^{\mu_\sigma}(y^{\mu_\sigma}+y^{\mu_\sigma+1}+\dots)\times$$
$$\times\left(\frac{x}{1-x}\right)^{m-\sigma-1}\left(\frac{y}{1-y}\right)^{m-\sigma-1} \tag{53}$$

(The last two factors allow for an arbitrary arrangement of the remaining atoms and holes into $(m-\sigma-1)$ contiguous blocks each.) In compact form, the numbers of matching configurations is therefore given by the coefficient of $x^k y^{p-k}$ in the expansion of:

$$x^\nu y^\mu \frac{(xy)^{m-\sigma-1}}{[(1-x)(1-y)]^{m-\sigma}} =$$

$$= \sum_{a=0}^{\infty}\binom{m-\sigma-1+a}{m-\sigma-1}x^{m-\sigma-1+\nu+a}\sum_{b=0}^{\infty}\binom{m-\sigma-1+b}{m-\sigma-1}y^{m-\sigma-1+\mu+b} \tag{54}$$

(Here, we already performed the customary division by m.) Thus, the number of matches in the m-block subset is (23′):

$$\sum_{z_m} \mathcal{M}(\eta, z) = \binom{k-\nu}{m-\sigma-1}\binom{p-k-\mu}{m-\sigma-1}$$
$$= \frac{1}{(m-\sigma-1)!}\;\frac{(\nu-k)_{m-\sigma-1}\,(\mu+k-p)_{m-\sigma-1}}{(1)_{m-\sigma-1}} \tag{55}$$

Hence, the sum in (23′) is a (terminating) hypergeometric series, and the average can be expressed as:

$$\langle \text{type } (\alpha)\rangle = \frac{1}{p}\left(\frac{B}{C}\right)^{\sigma}\frac{{}_2F_1(\nu-k,\mu+k-p,1;\,B/C)}{{}_2F_1(1-k,1+k-p,2;\,B/C)} \tag{56}$$

The special case (31) emerges from (56) if we set $\nu=r$, $\mu=1$, and $\sigma=0$.

In the case of atom-terminated compact patterns [case (β)], the generating function (53) is correspondingly modified:

$$m(x^{\nu_0}+x^{\nu_0+1}+\dots)\,y^{\mu_1}x^{\mu_1}y^{\mu_2}\cdot\ldots\cdot y^{\mu_{\sigma+1}}(x^{\nu_{\sigma+1}}+x^{\nu_{\sigma+1}+1}+\dots)\times$$
$$\times\left(\frac{x}{1-x}\right)^{m-\sigma-2}\left(\frac{y}{1-y}\right)^{m-\sigma-1} \qquad (57)$$

After elimination of duplicate states introduced by rotation of blocks, the generating function becomes:

$$x^{\nu-2}y^{\mu}\left(\frac{x}{1-x}\right)^{m-\sigma}\left(\frac{y}{1-y}\right)^{m-\sigma-1} \qquad (58)$$

and the sum over all matching states within the m-block subset is the coefficient of $x^k y^{p-k}$ in the series expansion of (58) [cf. (4), (40), (54)]:

$$\sum_{\mathfrak{z}_m} M(\eta,\mathfrak{z}) = \binom{k+1-\nu}{m-\sigma-1}\binom{p-k-1-\mu}{m-\sigma-2}$$
$$= -\frac{1}{(m-\sigma-2)!}\;\frac{(\nu-k-1)_{m-\sigma-1}\,(k+1+\mu-p)_{m-\sigma-2}}{(1)_{m-\sigma-1}}$$
$$= \frac{k+1-\nu}{(m-\sigma-2)!}\;\frac{(\nu-k)_{m-\sigma-2}\,(k+1+\mu-p)_{m-\sigma-2}}{(2)_{m-\sigma-2}} \qquad (59)$$

In the latter form, the sum over averages (23') can be performed in closed form and yields another hypergeometrical function:

$$\langle \text{type } (\beta)\rangle = \frac{k+1-\nu}{p}\left(\frac{B}{C}\right)^{\sigma+1}\frac{{}_2F_1(\nu-k,\,1+k+\mu-p,\,2;\,B/C)}{{}_2F_1(1-k,\,1+k-p,\,2;\,B/C)} \qquad (60)$$

This result even embraces the block averages $\langle \bullet\bullet\ldots\bullet\rangle$ considered in Section (ii): Their equilibrium occupation (41) follows from (60) if we formally set $\sigma=-1$.

The analogous result for the averages of hole-terminated patterns $\langle \circ\ldots\circ\rangle$ is available from (60) by replacing $k\leftrightarrow p-k$, $\nu\leftrightarrow\mu$, and $\sigma\leftrightarrow\tau=\sigma-1$:

$$\langle \text{type } (\beta)\rangle = \frac{p-k+1-\mu}{p}\left(\frac{B}{C}\right)^{\sigma}\frac{{}_2F_1(1+\nu-k,\,\mu+k-p,\,2;\,B/C)}{{}_2F_1(1-k,\,1+k-p,\,2;\,B/C)} \qquad (61)$$

The averages $\langle \circ\bullet\bullet\ldots\bullet\circ\rangle$ (46) [Section (iii)] follow with $\nu=r$, $\mu=2$, and $\sigma=1$.