A gentle introduction to combinatorial stochastic processes V

Enrico Scalas

Department of Mathematics, University of Sussex, UK

Stochastic Models for Complex Systems
10 June - 7 July, 2021

Outline

Binary moves I

So far, we used unary moves and we introduced binary moves in an exercise. Why are binary moves needed? In Physics, interparticle collisions cause a change of category (state) of two particles. A binary collision is represented by two particles in the initial categories (also called *cells* in this case) (i, j) and ending in the final categories (k, l). A binary transition can be represented by $\mathbf{n}_{ij}^{kl}|\mathbf{n}$, where

$$\mathbf{n} = (n_1, \ldots, n_i, \ldots, n_j, \ldots, n_k, \ldots, n_l, \ldots, n_g) \tag{1}$$

and

$$\mathbf{n}_{ij}^{kl} = (n_1, \ldots, n_i - 1, \ldots, n_j - 1, \ldots, n_k + 1, \ldots, n_l + 1, \ldots, n_g). \tag{2}$$

A generalization of the Ehrenfest Brillouin transition probability to binary moves is the following, for i, j, k, l all different:

$$\mathbb{P}(\mathbf{n}_{ij}^{kl}|\mathbf{n}) = P(\mathbf{n}_{ij}|\mathbf{n})P(\mathbf{n}_{ij}^{kl}|\mathbf{n}_{ij}) = 4\frac{n_i}{n}\frac{n_j}{n-1}\frac{(\alpha_k + n_k)(\alpha_l + n_l)}{(\alpha + n - 2)(\alpha + n - 1)} \tag{3}$$

Binary moves II

Neglecting the coefficient 4, the first two terms give the probability to select a particle in the i-th cell, followed by the extraction of a particle in cell j in a hypergeometric process of parameter $\mathbf{n}$, where $\mathbf{n}$ is the initial occupation vector. The third term is an accommodation Pólya process with weights $\alpha_k + n_k$ for the k-th cell when the total weight is $\alpha + n - 2$, and the fourth is an accommodation Pólya process with weights $\alpha_l + n_l$ and total weight given by $\alpha + n - 1$. This mechanism is just the same as for unary moves, considering that each destruction or creation step is conditioned to the current state of the population and that the population diminishes during destructions, and increases during creations. The factor 4 is due to the four paths connecting $\mathbf{n}$ to $\mathbf{n}_{ij}^{kl}$. In fact, denote by $D_1 = i$, $D_2 = j$, $C_1 = k$, $C_2 = l$ the sequence of the two extractions followed by the two accommodations. There are two destruction sequences connecting $\mathbf{n}$ to $\mathbf{n}_{ij}$, that is $D_1 = i, D_2 = j$ and $D_1 = j, D_2 = i$, and there are two creation sequences from $\mathbf{n}_{ij}$ to $\mathbf{n}_{ij}^{kl}$, that is $C_1 = k$, $C_2 = l$ and $C_1 = l$, $C_2 = k$. Due to the exchangeability of destructions $\{D_r\}$ and and creations $\{C_r\}$, the four paths D_1, D_2, C_1, C_2 are equiprobable. The coefficient is not 4 if some index is repeated, but these subtleties can be omitted. The simplest route to the invariant distribution is using the detailed balance conditions, and showing that the invariant distribution of (3) is still the Pólya distribution.

Pólya invariance

There is another way to show that the Pólya distribution is invariant for sequences of creations and destructions. Assume that n elements are described by a dichotomous exchangeable distribution. Then, as known from previous lectures, one can write

$$\mathbb{P}_n(k) = \binom{n}{k} \mathbb{P}(\mathbf{X}(k, n)), \tag{4}$$

where $\mathbf{X}(k, n)$ denotes an individual description with k successes in n trials. Then, a random destruction is performed, leading to

$$\mathbb{P}_{n-1}(k) = \mathbb{P}_n(k+1)\mathbb{P}(k|k+1) + \mathbb{P}_n(k)\mathbb{P}(k|k) = \mathbb{P}_n(k+1)\frac{k+1}{n} + \mathbb{P}_n(k)\frac{n-k}{n}. \tag{5}$$

Therefore, in the dichotomous case, if the chain is described by $\mathrm{Polya}((k, n-k); (\alpha_1, \alpha - \alpha_1))$, then, after a sequence of $m \leq n$ hypergeometric destructions, the chain will be described by a $\mathrm{Polya}((k, n-m-k); (\alpha_1, \alpha-\alpha_1))$ and, in the multivariate case, if $\mathbf{m}$ denotes the vector of destructions, one will pass from $\mathrm{Polya}(\mathbf{n}; \boldsymbol{\alpha})$ to $\mathrm{Polya}(\mathbf{n} - \mathbf{m}; \boldsymbol{\alpha})$.
Consider the case in which an element is added to a population of n elements. In the case of creation probabilities following the Pólya scheme, the analogous of equation (5) is given by:

$$\mathbb{P}_{n+1}(k) = \mathbb{P}_n(k-1)\frac{\alpha_1 + k - 1}{\alpha + n} + \mathbb{P}_n(k)\frac{\alpha - \alpha_1 + n - k}{\alpha + n} \tag{6}$$

The reader can directly check that, if $\mathbb{P}_n(k)$ is $\mathrm{Polya}((k, n-k); (\alpha_1, \alpha - \alpha_1))$, then $\mathbb{P}_{n+1}(k)$ is $\mathrm{Polya}((k, n+1-k); (\alpha_1, \alpha - \alpha_1))$. Again this results is valid for any number of steps, and also for the multivariate Pólya distribution. Then applying equation (6) m times in sequence to $\mathrm{Polya}(\mathbf{n} - \mathbf{m}; \boldsymbol{\alpha}))$, one recovers $\mathrm{Polya}(\mathbf{n}; \boldsymbol{\alpha}))$, that is the starting distribution.
Summarizing all the previous discussion: if the population is described by a distribution $\mathrm{Polya}(\mathbf{n}; \boldsymbol{\alpha})$, then after a sequence of $m \leq n$ hypergeometric destructions the shrinked population is described by a distribution which is $\mathrm{Polya}(\mathbf{n} - \mathbf{m}; \boldsymbol{\alpha})$, and after a sequence of m Pólya creations, the restored population is described by the same initial distribution, that is $\mathrm{Polya}(\mathbf{n}; \boldsymbol{\alpha})$.

Collisions in a gas I

Suppose that two particles in a gas collide, and that their final states are different from the initial ones. The transition from the occupation state $\mathbf{n}$ to the occupation state $\mathbf{n}_{ij}^{kl}$ is associated with the events $D_1 = i, D_2 = j$, representing the initial categories of the two selected particles, and $C_1 = k, C_2 = l$, representing the final categories of the colliding particles. Now, further assume that the collision is *elastic* meaning that the total energy of the two particles is the same before and after the collision. In other words, the transition can occur only if

$$\varepsilon(i) + \varepsilon(j) = \varepsilon(l) + \varepsilon(k), \tag{7}$$

where $\varepsilon(i)$ denotes the energy of a particle in the i-th category, the transition can occur, otherwise it is forbidden. In this case, it is useful to call the categories *cells* and divide the g cells into energy levels, grouping cells of equal energy.

Collisions in a gas II

The parameters of the system are n, the number of particles, and the set $\{g_i, \varepsilon_i\}_{i=1,\dots,d}$, where $g = \sum_{i=1}^{d} g_i$ is the decomposition of the g cells according to their energy. If cells are re-ordered with respect to their energy in the following way $(1, \dots, g_1, g_1+1, \dots g_1+g_2, g_1+g_2+1, \dots, \sum_{i=1}^{d} g_i)$, with $g_0 = 0$, the vector

$$N_i = \sum_{j=1+\sum_{k=1}^{i} g_{k-1}}^{\sum_{k=1}^{i} g_k} n_j \tag{8}$$

represents the occupation number of an energy level. The level-occupation vector $\mathbf{N} = (N_1, \dots, N_d)$ is the so-called *macrostate* of the system, whereas the original cell-occupation vector (with cells numbered from 1 to g, as usual) $\mathbf{n} = (n_1, \dots, n_g)$ is the *microstate* of the system. The same procedure can be used for the initial weights, leading to

$$\beta_i = \sum_{j=1+\sum_{k=1}^{i} g_{k-1}}^{\sum_{k=1}^{i} g_k} \alpha_j. \tag{9}$$

Collisions in a gas III

If the system starts from a given microstate **n** and energy is conserved, with unary moves, it cannot leave the initial macrostate **N**. On the contrary, binary moves can mimic elastic collisions, and the transition probability $\mathbb{P}(\mathbf{n}_{ij}^{kl}|\mathbf{n})$ still has the form of equation (3), with a slight modification due to the existence of forbidden transitions. In fact, the destruction term is given by the usual term $2(n_i/n)(n_j/(n-1))$ for $i \neq j$, or by $(n_i/n)(n_i-1)/(n-1)$ for $j = i$; if creation concerns cells not involved in the previous destructions, the creation term is proportional to $2(\alpha_k + n_k)(\alpha_l + n_l)$ for $l \neq k$, or to $(\alpha_k + n_k)(\alpha_k + n_k + 1)$ for $l = k$. These cumbersome notes are useful in order to write a program simulating the dynamics: at any stage of the move, one must use the current occupation number, possibly modified by previous partial moves. In order to normalize the creation probability, one must consider all final states reachable from $\mathbf{n}_{ij}$ adding two particles whose total energy is $\varepsilon(l) + \varepsilon(k) = \varepsilon(i) + \varepsilon(j)$.

The effect of constraints

Defining the term

$$Q(\mathbf{n}_{ij}^{kl}|\mathbf{n}_{ij}) = (2-\delta_{lk})(\alpha_k + n'_k)(\alpha_l + n'_l + \delta_{lk}), \tag{10}$$

where n'_k and n'_l are the occupation numbers of $\mathbf{n}_{ij}$, then one has

$$\mathbb{P}(\mathbf{n}_{ij}^{kl}|\mathbf{n}_{ij}) = \frac{Q(\mathbf{n}_{ij}^{kl}|\mathbf{n}_{ij})}{\sum_{k',l'} Q(\mathbf{n}_{ij}^{k'l'}|\mathbf{n}_{ij})} \tag{11}$$

and, finally, the transition probability for the binary move is

$$\mathbb{P}(\mathbf{n}_{ij}^{kl}|\mathbf{n}) = \mathbb{P}(\mathbf{n}_{ij}|\mathbf{n})\mathbb{P}(\mathbf{n}_{ij}^{kl}|\mathbf{n}_{ij}) =$$

$$= (2-\delta_{ij})\frac{n_i}{n}\frac{n_j}{n-1}\frac{Q(\mathbf{n}_{ij}^{kl}|\mathbf{n}_{ij})}{D}\delta[\varepsilon(l)+\varepsilon(k)-(\varepsilon(i)+\varepsilon(j))] \tag{12}$$

where

$$D = \sum_{k',l'} Q(\mathbf{n}_{ij}^{k'l'}|\mathbf{n}_{ij}), \tag{13}$$

and the sum is on the set

$$A(\mathbf{n},i,j) = \{k',l' : \ \varepsilon(l')+\varepsilon(k') = \varepsilon(i)+\varepsilon(j)\}. \tag{14}$$

Useful remarks

Not all the "first neighbours" of the initial microstate $\mathbf{n}$ can be reached in a single step, as the motion is confined to the so-called *constant energy surface*. Nevertheless, it is remarkable that, if $\mathbb{P}(\mathbf{n}_{ij}^{kl}|\mathbf{n}) > 0$, then also the inverse transition is possible, that is $\mathbb{P}(\mathbf{n}|\mathbf{n}_{ij}^{kl}) > 0$. This is called *microscopic reversibility*. If one considers all the states reachable from the initial state $\mathbf{Y}_0 = \mathbf{n}(0)$ by means of repeated applications of the transition probability (12), for all these states $\mathbf{Y}_t = \mathbf{n}(t)$, one finds that $\sum_{j=1}^{g} n_j(t)\varepsilon(j) = \sum_{j=1}^{g} n_j(0)\varepsilon(j) = E(\mathbf{n}(0))$ is a constant of motion. One can assume that all the states compatible with the constraint, i.e. the *energy surface* or *energy shell* $\{\mathbf{n} : \sum_{j=1}^{g} n_j\varepsilon(j) = E(\mathbf{n}(0))\}$, can be reached by means of binary collisions (if this is not the case, in principle, nothing prevents to introduce m-ary collisions). In any case, the energy surface is a subset of the unconstrained state space S_g^n, which strongly depends on the values $\{g_i, \varepsilon_i\}_{i=1,2,\ldots,d}$. Therefore, the transition matrix (12) rules a multidimensional random walk on the energy surface which is an irreducible and aperiodic Markov chain.

Invariant distribution I

In order to discuss the inverse transition, one can start from

$$\mathbf{n}_{ij}^{kl} = \mathbf{n} + \mathbf{k} + \mathbf{l} - \mathbf{i} - \mathbf{j}, \tag{15}$$

where $\mathbf{k}$, $\mathbf{l}$, $\mathbf{i}$, and $\mathbf{j}$ indicate vectors which are zero everywhere except for the positions k, l, i, and j, respectively, where they are equal to 1. Now, consider the path $D_1 = l, D_2 = k$, leading to $\mathbf{n}_{ij} = \mathbf{n} - \mathbf{i} - \mathbf{j}$, and the path $C_1 = j, C_2 = i$, leading to $\mathbf{n}$. Then, one has

$$\mathbb{P}(\mathbf{n}_{ij}|\mathbf{n}_{ij}^{kl}) = (2 - \delta_{lk})\frac{n_k}{n}\frac{n_l}{n-1}, \tag{16}$$

whereas

$$\mathbb{P}(\mathbf{n}_{ij}^{ij}|\mathbf{n}_{ij}) = \frac{Q(\mathbf{n}_{ij}^{ij}|\mathbf{n}_{ij})}{D}, \tag{17}$$

where $\mathbf{n}_{ij}^{ij} = \mathbf{n}$ and the denominator D coincides with the one in equation (12). In words, any path connecting $\mathbf{n}$ to $\mathbf{n}_{ij}^{kl}$ passes through the intermediate state $\mathbf{n}_{ij}$ in both directions. In both cases the destructions end in $\mathbf{n}_{ij}$, which is the starting state for creations, all constrained to belong to the set $A(\mathbf{n}, i, j)$. Note that if one sets $\mathbf{n}_{ij}^{kl} = \mathbf{n}'$, then $\mathbf{n}_{ij} = \mathbf{n}'_{kl}$, and $A(\mathbf{n}, i, j) = A'(\mathbf{n}', k, l)$, as $\varepsilon(l) + \varepsilon(k) = \varepsilon(i) + \varepsilon(j)$.

Invariant distribution II

Given that the multipliclity factors are equal, one gets from equation (12) and from the previous discussion

$$\frac{\mathbb{P}(\mathbf{n}_{ij}^{kl}|\mathbf{n})}{\mathbb{P}(\mathbf{n}|\mathbf{n}_{ij}^{kl})} = \frac{n_i n_j (\alpha_k + n_k)(\alpha_l + n_l)}{(n_k + 1)(n_l + 1)(\alpha_i + n_i - 1)(\alpha_j + n_j - 1)}, \tag{18}$$

for the case $i \neq j \neq k \neq l$. Indeed, there exists a probability distribution which satisfies the detailed balance conditions, such that

$$\frac{\pi(\mathbf{n}_{ij}^{kl})}{\pi(\mathbf{n})} = \frac{\mathbb{P}(\mathbf{n}_{ij}^{kl}|\mathbf{n})}{\mathbb{P}(\mathbf{n}|\mathbf{n}_{ij}^{kl})} \tag{19}$$

where the right hand side is given by equation (18), and it is the following:

$$\pi(\mathbf{n}) \propto \prod_{j=1}^{g} \frac{\alpha_j^{[n_j]}}{n_j!}. \tag{20}$$

It is the usual generalized Pólya distribution now restricted to (or conditional on) all the states belonging to the energy surface. It is tedious but straightforward to check that this works in all cases.

Deriving quantum statistics

If $\forall j$ one has $\alpha_j = 1$ (the Bose-Einstein case), then one finds $\pi(\mathbf{n}_{ij}^{kl}) = \pi(\mathbf{n})$ and one gets the uniform distribution *restricted* to the energy shell; if $\forall j$ $\alpha_j = -1$ (the Fermi-Dirac case), then one has $\pi(\mathbf{n}_{ij}^{kl}) = \pi(\mathbf{n})$ for all $\mathbf{n}$: $n_i \in \{0, 1\}$; finally, the case in which $\forall j$, one has $\alpha_j \to \infty$ (the Maxwell-Boltzmann case) yields $\pi(\mathbf{n}) \propto \left(\prod_{j=1}^{g} n_j!\right)^{-1}$, which is uniform on all the allowed individual descriptions. In other words, what was previously written in the abstract unconstrained case, here is valid, but *within the energy shell*. Further note that these results were obtained without knowing a detailed description of the energy shell.

From cells to levels

In all the three cases decribed above, all the parameters α_j have the same value (all the cells are on a par, if allowed by the energy constraint); therefore, it is useful to introduce the new parameter $c = 1/\alpha_j$. The creation factor $\alpha_j + n_j$ becomes proportional to $1 + cn_j = c[(1/c) + n_j]$, and the three statistics relevant in Physics are obtained for $c = 0, \pm 1$. For the Markov chain of macrostates $\mathbf{N}$, the creation factor becomes $g_i + cN_i = c[(g_i/c) + N_i]$. From equation (20), the invariant weight (to be normalized) of a macrostate is given by

$$W(\mathbf{N}) = \prod_{i=1}^{d} \frac{c^{N_i} (g_i/c)^{[N_i]}}{N_i!}, \tag{21}$$

which become the usual weight for macrostates in statistical mechanics:

$$W_{\mathrm{BE}}(\mathbf{N}) = \prod_{i=1}^{d} \frac{g_i^{[N_i]}}{N_i!} = \prod_{i=1}^{d} \binom{g_i + N_i - 1}{N_i}, \tag{22}$$

for $c = 1$ (Bose-Einstein case),

$$W_{\mathrm{FD}}(\mathbf{N}) = \prod_{i=1}^{d} \frac{g_{i[N_i]}}{N_i!} = \prod_{i=1}^{d} \binom{g_i}{N_i}, \tag{23}$$

for $c = -1$ (Fermi-Dirac case), and

$$W_{\mathrm{MB}}(\mathbf{N}) = \prod_{i=1}^{d} \frac{{g_i}^{N_i}}{N_i!}, \tag{24}$$

for $c = 0$ (Maxwell-Boltzmann case).

Boltzmann's approximate method I

In the unconstrained case, it is possible to get rid of the multivariate equilibrium distribution by marginalizing on a single category, and then obtain expected values for the multivariate distribution; on the contrary, in the constrained case, the initial symmetry of categories is broken by the energy function, and the marginal chain for a single category is very complicated. The only way to extract a univariate distribution from equation (21) was envisaged by Boltzmann for the first time: it is the method of the most probable macrostate. The reason to maximize the probability of the macrostate given in equation (21) instead of the microstate probability (20) is due to the assumed uniform distributions on the underlying microstates. Moreover, in Physics, macrostates are obtained by merging many energy levels, in order to obtain large values of N_i and g_i so that Stirling's approximation for factorials can be used. However, it is possible to use a better procedure than Stirling's approximation as one knows that equation (21) is exact.

Boltzmann's approximate method II

The maximum of $\ln W(\mathbf{N})$ with the constraints $\sum_{i=1}^{d} N_i = N$ and $\sum_{i=1}^{d} N_i \varepsilon_i = E$ is obtained by introducing Lagrange multipliers β, ν and by requiring that $\Delta(\ln W(\mathbf{N}) - \beta E + \nu N) = 0$. The variation is discrete. Considering the i-th level, assume that N_i becomes $N_i \pm 1$, so that $\Delta E_i = \pm\varepsilon_i$ and $\Delta N_i = \pm 1$. Now define $W(N_i)$ as

$$W(N_i) = \frac{c^{N_i}(g_i/c)^{[N_i]}}{N_i!}, \tag{25}$$

considering that $\ln W(\mathbf{N}) = \sum_{i=1}^{d} \ln W(N_i)$ one can calculate the variation of the weight for a level if N_i becomes $N_i \pm 1$ using (21):

$$\Delta^{+} \ln W(N_i) = \ln \frac{W(N_i+1)}{W(N_i)} = \ln \frac{c(g_i/c + N_i)}{N_i + 1} = \ln \frac{g_i + cN_i}{N_i + 1} \tag{26}$$

and

$$\Delta^{-} \ln W(N_i) = \ln \frac{W(N_i-1)}{W(N_i)} = \ln \frac{N_i}{g_i + c(N_i - 1)} = -\ln \frac{g_i + c(N_i-1)}{N_i} \tag{27}$$

Note that the term $N_i + 1$ in equation (26) allows an accommodation in an empty cell, whereas the term N_i in equation (27) forbids a destruction in an empty cell. These terms are important when N_i is small.

Boltzmann's approximate method III

But all this procedure is intended for macrostates where $N_i \gg 1$. In this case $\Delta^+ \ln W(N_i)$ and $\Delta^- \ln W(N_i)$ become equal in absolute value, N_i can be treated as a continuous variable, and one can write

$$\frac{d}{dN_i} \ln W(N_i) = \ln \frac{g_i + cN_i}{N_i} = \beta\varepsilon_i - \nu, \tag{28}$$

and deduce a simple analitic formula for the solution. If there exists a macrostate $\mathbf{N}^*$ such that

$$\ln \frac{(g_i + cN_i^*)}{N_i^*} = \beta\varepsilon_i - \nu, \tag{29}$$

that is

$$N_i^* = \frac{g_i}{\exp[\beta\varepsilon_i - \nu] - c}, \tag{30}$$

then it is the most probable state, which for $c = 1, -1, 0$ represents the Bose-Einstein, the Fermi-Dirac and the Maxwell-Boltzmann case, respectively.

Lagrange multipliers and statistical equilibrium

The two new parameters β, ν must be chosen so that $\sum_{i=1}^{d} N_i^* = N$ and $\sum_{i=1}^{d} N_i^* \varepsilon_i = E$. Returning to the exact formulae (26) and (27), considering the transition $\mathbb{P}(\mathbf{N}_{ij}^{kl}|\mathbf{N})$, the variation of $\ln W(\mathbf{N})$ is just $\ln W(\mathbf{N}_{ij}^{kl}) - \ln W(\mathbf{N})$, which contains four terms, if all levels are different, leading to

$$\ln W(\mathbf{N}_{ij}^{kl}) - \ln W(\mathbf{N}) =$$

$$= \Delta^+ \ln W(N_k) + \Delta^+ \ln W(N_l) + \Delta^- \ln W(N_i) + \Delta^- \ln W(N_j) =$$

$$= \ln \frac{N_i}{g_i + c(N_i - 1)} \frac{N_j}{g_j + c(N_j - 1)} \frac{g_k + cN_k}{N_k + 1} \frac{g_l + cN_l}{N_l + 1} =$$

$$= \ln \frac{\mathbb{P}(\mathbf{N}_{ij}^{kl}|\mathbf{N})}{\mathbb{P}(\mathbf{N}|\mathbf{N}_{ij}^{kl})} = \ln \frac{\pi(\mathbf{N}_{ij}^{kl})}{\pi(\mathbf{N})} \tag{31}$$

where the last equalities are a direct consequence of equation (19) applied to macrostates. In the limit of large occupation numbers, if the starting $\mathbf{N}$ is $\mathbf{N}^*$, and applying equation (28), one finds that

$$\Delta \ln W(\mathbf{N}^*) = \beta(\varepsilon_k + \varepsilon_l - \varepsilon_i - \varepsilon_j) = 0 \tag{32}$$

if $\mathbf{N}_{ij}^{kl}$ has the same energy as $\mathbf{N}^*$. This means that the probability is "flat" around $\mathbf{N}^*$, as all its first neighbours nearly have the same probability. In words, starting from $\mathbf{N}^*$, and considering any transition which satisfies the energy constraint (that is such that $\varepsilon_k + \varepsilon_l = \varepsilon_i + \varepsilon_j$), the final vector is still in the region of maximum probability. For macroscopic physical systems, where n is roughly 10^{23}, the assumption of statistical mechanics is that,the region of motion is overhelmingly concentrated around $\mathbf{N}^*$, which then summarizes all the underlying dynamics and, at least, all equilibrium properties.

Statistical equilibrium for small systems

Surprisingly, a formula similar to equation (30) holds true also for small values of n. In fact, considering the *expected* flux

$$\phi_{ij}^{kl} = \sum_{\mathbf{N}} \mathbb{P}(\mathbf{N}_{ij}^{kl}|\mathbf{N})\mathbb{P}(\mathbf{N}), \tag{33}$$

which is the probability that a transition $i, j \to k, l$ occurs whatever the initial state $\mathbf{N}$ may be, and the reversed flux

$$\phi_{kl}^{ij} = \sum_{\mathbf{N}} \mathbb{P}(\mathbf{N}_{kl}^{ij}|\mathbf{N})\mathbb{P}(\mathbf{N}), \tag{34}$$

then assuming that the two fluxes are equal, one has that $\phi_{ij}^{kl}/\phi_{kl}^{ij} = 1$. Now, one can write

$$\phi_{ij}^{kl}/\phi_{kl}^{ij} = \frac{\mathbb{E}[N_i N_j (g_k + cN_k)(g_l + cN_l)]}{\mathbb{E}[N_k N_l (g_i + cN_i)(g_j + cN_j)]}. \tag{35}$$

If one can approximate the expectation of the product with the product of expectations, then from $\phi_{ij}^{kl}/\phi_{kl}^{ij} = 1$ one gets

$$\ln \frac{\mathbb{E}[N_i]}{g_i + c\mathbb{E}[N_i]} \frac{\mathbb{E}[N_j]}{g_j + c\mathbb{E}[N_j]} \frac{g_k + c\mathbb{E}[N_k]}{\mathbb{E}[N_k]} \frac{g_l + c\mathbb{E}[N_l]}{\mathbb{E}[N_l]} = 0; \tag{36}$$

in other words, if for any energy level $\ln[\mathbb{E}[N_i]/(g_i + c\mathbb{E}[N_i])] = \nu - \beta\varepsilon_i$, then for any conservative collision the two fluxes equalize. Then equation (30) can be re-written also for expected values

$$\mathbb{E}[N_i] = \frac{g_i}{\exp[\beta\varepsilon_i - \nu] - c}, \tag{37}$$

Final remarks I

Although both equations (37) and (30) are approximate, in some applications (30) could be meaningless, whereas equation (37) is always meaningful. For instance, consider a system with $g_i = 1$ for all the values of i. This is strange in Physics, where g_i in an increasing function of i, but it can be the case in Economics or in other fields. Then equation (30) could still apply, because it was only for $N_i \gg 1$, while the usual Stirling's approximation also needs $g_i \gg 1$. In this particular case, macrostates and microstates do coincide. Further assume that $c = 1$. One knows from the distribution for occupation vectors in the quantum Bose-Einstein case that the equilibrium distribution is flat, that is all occupation states are equiprobable. Hence there is no $\mathbf{N}^*$ at all! On the contrary, and this is can be shown by computer simulation, $\mathbb{E}[N_i]$ can be derived from the time average of occupation numbers for the i-th level as time goes by, and an agreement with equation (37) is easily obtained. Note that if there exists $\mathbf{N}^*$ such that $\pi(\mathbf{N}^*) \approx 1$, then $\mathbb{E}[N_i] = \sum N_i \pi(\mathbf{N}) \approx N_i^*$, but $\mathbb{E}[N_i]$ also exists when there is no predominant macrostate $\mathbf{N}^*$.

Final remarks II

Finally, in order to justify the assumption leading from equation (35) to equation (36), note that if g random variables satisfy the constraint $\sum_1^g Y_i = n$, then $\mathbb{V}(\sum_1^g Y_i) = \sum_1^g \mathbb{V}(Y_i) + \sum_{j \neq i} \mathbb{C}(Y_i, Y_j) = 0$. If all the variables Y_i are equidistributed, one finds that $g(g-1)\mathbb{C}(Y_i, Y_j) = -g\mathbb{V}(Y_i)$ and the Bravais-Pearson correlation becomes

$$\rho(Y_i, Y_j) = \frac{\mathbb{C}(Y_i, Y_j)}{\sqrt{\mathbb{V}(Y_i)\mathbb{V}(Y_j)}} = -\frac{1}{g-1}. \tag{38}$$

Given that in any concrete case, the number g of levels involved in the dynamics is far larger than n, the correlation among the occupation numbers is negligible, and the corresponding approximation works fairly well.