Available online at www.sciencedirect.com
ScienceDirect
Physica A 374 (2007) 85–102
PHYSICA A
www.elsevier.com/locate/physa

Order–disorder separation: Geometric revision

Alexander Gorban

Centre for Mathematical Modelling, University of Leicester, University Road, Leicester LE1 7RH, UK

Received 25 July 2005; received in revised form 25 June 2006
Available online 14 August 2006

Abstract

After Boltzmann and Gibbs, the notion of disorder in statistical physics relates to ensembles, not to individual states. This disorder is measured by the logarithm of ensemble volume, the entropy. But recent results about measure concentration effects in analysis and geometry allow us to return from the ensemble-based point of view to a state-based one, at least, partially. In this paper, the order–disorder problem is represented as a problem of relation between distance and measure. The effect of strong order–disorder separation for multiparticle systems is described: the phase space could be divided into two subsets, one of them (set of disordered states) has almost zero diameter, the second one has almost zero measure. The symmetry with respect to permutations of particles is responsible for this type of concentration. Dynamics of systems with strong order–disorder separation has high average acceleration squared, which can be interpreted as evolution through a series of collisions (acceleration-dominated dynamics). The time arrow direction from order to disorder follows from the strong order–disorder separation. But, inverse, for systems in space of symmetric configurations with "sticky boundaries" the way back from disorder to order is typical (Natural selection). Recommendations for mining of molecular dynamics results are also presented.

Keywords: Order; Disorder; Irreversibility; Phase volume; Measure concentration; Entropy

0. Introduction

Is everything clear with the entropy growth? It seems that it is not. A collection of problem statements and approaches was published by Physica A on the eve of the millennium [1–3]. Very recently, V.L. Ginzburg in his Nobel Lecture characterized this problem as one of the greatest challenges for physicists:

> The "great problems" are, first, the increase in entropy, time irreversibility, and the "time arrow" [4].

We usually describe the time arrow as disorder increase, and measure disorder by the (logarithm of) phase volume following the famous Boltzmann epitaph

$$S = k \ln W,$$

E-mail address: ag153@le.ac.uk.

0378-4371/$ - see front matter
doi:10.1016/j.physa.2006.07.034

where W is the volume of an ensemble and S is the entropy of this ensemble. The ensemble-based point of view was expressed recently in the following reasoning [5, p. 329]:

> The well known question of what has more order, a fine castle or a pile of stones, has a profound answer: It depends on which pile you mean. If "piles" are thought as all configurations of stones which are not castles, then there are many more such piles, and so there is less order in such a pile. However, if these are specially and uniquely placed stones (for example, a garden of stones), then there is the same amount of order in such a pile as in a fine castle. *Not a specific configuration is important but an assembly of configurations embraced by one notion.*

It seems to be true, but it is not the whole truth. In this paper, the ensemble-based point of view will be complemented by the state-based one: The notions of order and disorder can describe not only ensembles, but also points.

The following *toy-example* gives us a nice possibility to understand the difference between the state-based and the ensemble-based point of view, and helps us to learn how the measure of order and disorder depends on the human activity and perspective as well as on a state itself. Most of the people are familiar with the situation described in this example.

In book [6], the picture of "order" after intensive play of four children is presented to illustrate the idea: the definition of order depends on a point of view, and the same set of positions and orientations of toys may serve as a representative of rather big ensemble of equivalent disorders ("parents-room"), or as an almost unique configuration that changes sense after small change ("children-room"). Children implicitly use the positions and orientations of all their toys in their play. For parents, these differences are not important. The same room (a state) produces different ensembles, it depends on perspective. The notion "order" distinguishes wide ensemble of the parents-room (big volume, disorder) from narrow ensemble of the children-room (small volume, order), and the entropy measures this difference. This situation should be reflected on deeply before entering into any discussion about order–disorder measurement.

This difference between the parents-room and the children-room can be formalized by the volumes of equivalent configurations. For the parents-room, it seems to be larger, because the parents-equivalence is coarser (they use "other variables" for description of the state of the room). Here, we meet the important operation that replaces a state (a point) by an ensemble. The simplest formal version of this operation is the so-called "fattening": in a metric space with metric $\rho(x, y)$ for any set A and $\varepsilon > 0$ the ε-fattening of A is the set

$$A_\varepsilon = \{x : \rho(x, y) < \varepsilon \text{ for some } y \in A\}. \tag{1}$$

The set A_ε includes all points that belong to A "with accuracy ε."[1] Our first attempt to describe the difference between the parents-room and the children-room is the hypothesis that these ensembles are results of ε-fattening for the same state (a point), but with significantly different ε. The volume of the parents-room-ensemble is much higher than the volume of the children-room-ensemble.

This point of view is not the final one. Later, in this paper, it will be complemented by the permutation analysis: the parents-room has more permutation symmetry than the children-room, and this causes significant difference between their ε-fattening even for the same ε. The symmetrization occurs to be the most important operation for understanding of the difference between thermodynamic order and disorder.

In this paper, the order–disorder problem is represented as a problem of relation between distance and measure. The main focus of our consideration is the effect of *order–disorder separation*: for systems with a large number of particles the available phase space (or configuration space) can be divided into two parts. One part has microscopically small diameter (part D, disorder), another part (part O, order) has microscopically small measure (volume). We call a quantity *microscopically small*, if it tends to 0 when the number of particles tends to ∞. Of course, a proper normalization of the volume and distance is assumed. As a consequence of the order–disorder separation it is worth mentioning the existence of such a microscopically small $\varepsilon > 0$ that for each point x from the part D its ε-fattening $\{x\}_{\varepsilon>0}$ includes almost all volume (the rest of the volume is microscopically small).

We follow the idea of *thin–thick decomposition* (see Gromov book [7, p. 124]). The effect of order–disorder separation is one of the *measure concentration effects*. The geometry of spaces with finite, but very large

[1] The fattening is similar to the Ehrenfest's coarse-graining [8,9].

dimension has some interesting features that simplifies the asymptotic picture in comparison with both the small-dimensional, and the infinite-dimensional pictures. The typical questions refer to various asymptotic relations between the Lebesgue measure and the Euclidean distance. Recently, the effects of this kind have been studied very intensively [7,10,11]. Some links between concentration of measure and works of Boltzmann, Maxwell, Gibbs, and Ehrenfest are presented below (nothing is absolutely new).

For the measure concentration that leads to order–disorder separation, the permutation symmetry between particles (PI—*Permutation Invariance*) is important.

The paper has the following structure. In Section 1, two classical examples of measure concentration are presented: the waist (or Maxwell) concentration of all the volume of multidimensional spheres near equators, and the boundary (or Gibbs) concentration of the volume of multidimensional balls near boundaries (spheres). In Section 2, the Feynman analysis of an example of order increase is collated [12]. The order–disorder separation for the Feynman example is demonstrated in Section 3.

Below, we discuss the statistical idea of order/disorder only. It is based on the analysis of differences between less probable/more probable events for large systems. There exist many other notions of order/disorder, most important of them is the presence/absence of a regular structure. We do not touch them in this paper.

1. The classical measure concentration effects

For large dimension n, the main part of the volume of the unit n-dimensional ball B^n is concentrated in a small neighborhood of its boundary, that is the unit sphere S^{n-1}. This simple, but very seminal fact can be demonstrated, as follows. Let us use the normalized volume $|\bullet|$: $|B^n| = 1$. The correspondent (normalized) surface area of a unit sphere is a constant C_n, $|B^n| = \int_0^1 C_n r^{n-1}\,\mathrm{d}r$, hence, $C_n = n$. The volume of the part of B^n inside the ε-neighborhood of S^{n-1} is

$$V_\varepsilon = 1 - \int_0^{1-\varepsilon} nr^{n-1}\,\mathrm{d}r = 1 - (1-\varepsilon)^n. \tag{2}$$

For small ε and large n (say, $n > 1/\varepsilon$), we obtain the exponential estimate:

$$V_\varepsilon = 1 - (1-\varepsilon)^{(1/\varepsilon)n\varepsilon} \approx 1 - \exp(-n\varepsilon). \tag{3}$$

It implies that for a given ε and $n \to \infty$ the volume $V_\varepsilon \to 1$ as $1 - \exp(-n\varepsilon)$ (exponentially).

The sphere S^{n-1} can be considered as the isoenergetic surface for a very simple energy function, $E = \sum_{i=1}^n x_i^2$, that is, for kinetic energy of n classical particles on a line, or for potential energy of n simplest classical oscillators. Of course, the observed concentration theorem could be proved for more general energy functions. Usually these generalizations are formulated as theorems of ensemble equivalence: for large n the canonical ensemble (ensemble with probability distribution that maximizes the entropy functional for a given average energy value) is equivalent to the microcanonical ensemble (that is equidistribution on the isoenergetic surface with respect to invariant Liouville measure). Khinchin [13] describes the probabilistic theory of ensemble equivalence when energy is a "sum function", this means that the system consists of a large number of non-interacting subsystems. This type of concentration we call the Gibbs concentration. The analysis of ensemble equivalence and non-equivalence is presented in Ref. [14] with relevant references.

The same type of reasoning can be applied to a hemisphere $H^{n-1} = \{x \in S^{n-1} : x_1 \geqslant 0\}$: for large n almost all measure of the hemisphere H^n is concentrated near its boundary $S^{n-2} = \{x \in S^{n-1} : x_1 = 0\}$. Hence, almost all measure of S^{n-1} is concentrated near its $n-2$-dimensional equator. The exponential estimate of the type (3) is also valid. The well-known application of this "waist concentration" is the Maxwell distribution for particle velocity: if the n-particle system in the velocity space is equidistributed on the sphere of radius $R^2 = \sum_{i=1}^n v_i^2 = 3nkT/m$, then, for large n, the velocity of one particle will be distributed due to the Maxwell distribution. The distribution of v_1 has the Gaussian density $\frac{1}{\sqrt{2\pi}\sigma}\exp(-v_1^2/2\sigma^2)$, where $\sigma^2 = kT/m$. In other term, the projection of uniform distribution from the unit sphere S^n onto the first axis has, for large n, the narrow (almost) Gaussian distribution $\frac{1}{\sqrt{2\pi}\sigma}\exp(-v_1^2/2\sigma^2)$, where $\sigma = 1/\sqrt{n}$.

Due to ensemble equivalence this Maxwell concentration might be demonstrated as concentration of the projection of the equidistribution in the ball B^n on a line. This projection is a probability distribution on the

segment $[-1,1]$ with the density $\sim(\sqrt{1-x^2})^n$. For large n, $(\sqrt{1-x^2})^n \approx \exp(-nx^2/2)$, and the projection density approaches the Gaussian distribution $\sqrt{\frac{n}{2\pi}}\exp(-nx^2/2)$ with the standard deviation $\sigma = 1/\sqrt{n}$.

The waist concentration holds not only for projection on coordinate axis, but for any (non-linear) Lipschitz function $F(x)$ with Lipschitz constant 1 ($|F(x)-F(y)| \leqslant |x-y|$): for large n, the values of such a function on S^n are concentrated in a $\frac{1}{\sqrt{n}}$-small interval around the median value $\overline{F}$ defined by the following statement:

$$\boldsymbol{P}(F(x) \geqslant \overline{F}) \geqslant \tfrac{1}{2} \quad \text{and} \quad \boldsymbol{P}(F(x) \leqslant \overline{F}) \geqslant \tfrac{1}{2}.$$

It is the Levy theorem [15].

The Maxwell distribution was known before statistical mechanics was developed by Gibbs (and almost at the same time and independently by Einstein). The waist concentration, in this sense, was discovered by Maxwell.

Let us mention one important property of the waist concentration: the points on the sphere are distributed uniformly, and are equivalent in any reasonable sense: the measure is concentrated near *every* equator. In one-dimensional projections (both linear and general Lipschitz) this symmetry is destroyed, and there are distinguished points, the median and its $\frac{1}{\sqrt{n}}$-small neighborhood. The complement of this set has small measure, and this set of distinguished points has small diameter; and, of course, a small vicinity of any distinguished point has the same property, it has the "almost full" measure, and the small diameter. It does not matter, if this projection is linear or not, only the Lipschitz property is important. It makes no difference if the projection is not one-dimensional: for any given dimension and for the number of degrees of freedom $n \to \infty$ the result is the same. The final results concerning the waist concentration for different possible relations between n and dimension of projection were obtained by Gromov [11].

We can call the distinguished points as "thermalized" states, or "near-equilibrium" states, but initially, on the multidimensional sphere, all the states are equivalent, and the distinguished points of measure concentration emerge only in a macroscopic projection. In Section 2, we will present the order–disorder separation for microscopic state.

2. Strong order–disorder separation for symmetric microscopic states

All the classical statistical physics is the theory of symmetric ensembles: the density $\rho(x_1, x_2, \ldots, x_n)$ of the full multiparticle probability distribution is symmetric with respect to particles permutations (here x_i is a phase point for the ith particle). In this section, we demonstrate the concentration effect that emerges in the projection of the phase space (or configuration space) of n particles onto the space of permutations orbits. The n-particle space is P^n, where P is an one-particle space. The space of orbits can be presented as the space of n-point subsets in the one-particle space P (in the measure and distance discussion for continuous spaces we can neglect the degenerate case when positions of some particles coincide).

2.1. Feynman's blue and white atoms mixing

Let us start from a simplest example of blue and white atoms mixing analyzed in the book "The Character of Physical Law," by Feynman [12].

> You have atoms of two different kinds (it's ridiculous, but let's call them blue and white) jiggling all the time in thermal motion. If we were to start from the beginning we should have mostly atoms of one kind on one side, and atoms of other kind on the other side. Now these atoms are jiggling around, billions and billions of them, and if we start them with one kind all on one side, and the other kind on the other side, we see that in their perpetual irregular motions they will get mixed up, and that is why the water becomes more or less uniformly blue. …
>
> If you start with a thing that is separated and make irregular changes, it does get more uniform. But if it starts uniform and you make irregular changes, it does not get separated. It *could* get separated. It is not against the law of physics that the molecules bounce around so that they separate. It is just unlikely. It would never happen in a million years. And that is the answer.

This discussion is interesting not only by the clearly explained thing, but by the carefully hidden things also. Let P be the box where the atoms move. The configuration space is P^n, where n is the number of particles. The separated configurations (''with one kind all on one side, and the other kind on the other side'') form an ensemble (a ''drop'') with volume 2^n times smaller than the whole volume of P^n. The concentration effects in the velocity spaces allow us to represent the correspondent ensemble in a phase space as a drop with a constant density inside it also (for example, with equidistribution in a velocity ball). It is convenient for discussion. The volume of this drop is 2^n times smaller than the equilibrium volume (hence, the density is 2^n times larger). This volume is conserved in the mechanical motion. Hence, after some time this ensemble become more mixed, but remains ''oil in water'', that is, a phase space drop with the same volume and density. In the sense of ensembles it is not a ''uniform'' ensemble, and if somebody (the Maxwell demon, for example) carefully inverted all the velocities, this ensemble would return to the initial separated state.

What does Feynman mean: ''starting from homogeneous state we never will get the separation…?'' It is absolutely new ensemble ''uniform states'', it is not a result of the initial ensemble evolution. Starting from the initial separated state we do reach some of the ''uniform states'', but not all such states. The phase volume is different. For ''all uniform states'' it is 2^n larger, where n is the number of particles. How can we get *all* the uniform states (ensemble U) from the states we can reach from our ordered states (ensemble O)?

And here Feynman uses an unexpected new notion, *irregular changes*: ''If you start with a thing that is separated and make irregular changes, it does get more uniform.'' And back: ''if it starts uniform and you make irregular changes, it does not get separated.''

How and who makes these irregular changes and what do they mean? The small portion of irregular changes makes the mixed ''oil in water'' ensemble strictly uniform. Where did this concept come from? We can find a source of this idea in the coarse-graining.

The idea of coarse-graining dates back to Ehrenfest and Ehrenfest-Afanasyeva, and it has been most clearly expressed in their famous paper of 1911 [8]. Ehrenfests considered a partition of the phase space into small cells, and they have suggested to supplement the motions of the phase space ensemble due to the Liouville equation with ''shaking''—averaging of the density of the ensemble over the phase cells. As a result of this process, the convergence to the equilibrium becomes uniform out of the convergence in average. It is the fattening that we mentioned in Introduction. This ''fattening-based'' approach was developed into a general technique of non-equilibrium thermodynamics [9,16]. What is the physical nature of the ε-fattening? First interpretation is noise, any kind of small noise, small perturbations, and ε is the amplitude of this noise. Another interpretation of ε is the possible accuracy of measurement and control.

But, there is a purely mechanical effect: we start from the state with small volume of its ε-fattening and after some time of motion the system typically reaches states with large volume of their ε-fattening.

After some time of mechanical motion a typical state (a point, not an ensemble) becomes ''thick'': permutation symmetrization with microscopically small fattening transforms this point into an uniform ensemble. The explanation of this effect is based on the study of the geometry of a multidimensional simplex that we perform in the Section 2.2.

Let us watch blue particles only, and an one-dimensional box $P = [0, 1]$ (in the direction of separation x). In order to represent the set[2] of n particles as a point in a standard simplex, we introduce *symmetric coordinates* for n-particle systems. Let us enumerate particles in the order of x value: $0 = x_0 \leqslant x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n \leqslant x_{n+1} = 1$. Symmetric coordinates are:

$$s_i = x_i - x_{i-1}, \tag{4}$$

where $i = 1, \ldots, n+1$. Unordered n-particle states form in coordinate s_i a standard simplex Δ_n: $s_i \geqslant 0$, $\sum_i s_i = 1$. The configuration volume transforms into a uniform distribution in this simplex with a constant density $n!$.

[2]Positions of some particles can coincide, and, rigorously, a ''set'' of n particles forms an *unordered tuple*. An unordered tuple of length n of set P is a unordered selection with possible repetitions of set P and is represented by a sorted list of length n. In one-dimensional case, it is convenient to sort positions (numbers) in ascending order.

Of course, it is possible to study the space of permutation orbits as a quotient space endowed by quotient metrics. For the Euclidean metric in the one-particle space, the quotient metrics is

$$d_Q(\{x_1, x_2, \ldots, x_n\}, \{y_1, y_2, \ldots, y_n\}) = \left[\min_\sigma \left\{\sum_{i=1}^{n} \|x_i - y_{\sigma(i)}\|^2\right\}\right]^{1/2}, \tag{5}$$

where minimum is calculated for the set of all n-particle permutations σ. Nevertheless, the use of symmetric coordinates is more transparent. There are several other symmetric representations of n-particle systems: measure representation and functional (distance) representation. They are discussed below.

2.2. Distance–measure relations in large-dimensional simplex

Let us consider an n-dimensional standard simplex Δ_n. The normalized equidistribution in Δ_n has the constant density $n!$. We call the correspondent probability measure the *normalized volume*, and use notation Vol(•): $\mathrm{Vol}(\Delta_n) = 1$. When discussing the probability, we identify the probability of an event $\mathbf{P}\{\bullet\}$ with the volume of a correspondent set Vol(•).

For large n, almost all volume of the simplex Δ_n is concentrated in a small neighborhood of the center of Δ_n, near the point $c = (\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$. The Euclidean radius of this neighborhood R can be chosen of order $\sim n^{-1/2}$. A projection of an n-dimensional Euclidean ball with unit radius on a line is concentrated in an interval of length $\sim n^{-1/2}$. It is the Maxwell (the waist) concentration. Hence, any projection of an n-dimensional standard simplex on a line is concentrated within an interval of length $\sim n^{-1}$. This is true not only for orthogonal projections, but also for any Lipschitz functions with Lipschitz constant 1 (1-Lipschitz functions), as it is for Levy concentration. (See Ref. [7, p. 235].)

In order to demonstrate the main concentration properties of a simplex, let us start with the moment evaluation. The moments give this estimate of concentration radius in simplex, but only power estimates of deviations are achievable in this way. Let us follow Chebyshev's inequality for positive random variable ξ: $\mathbf{P}\{\xi \geqslant a\} \leqslant \mathbf{E}(\xi)/a$, where $\mathbf{E}(\xi)$ is the expectation of ξ (the average).

The distribution density for value s of one coordinate s_i in n-dimensional standard simplex is $p_1(s) = n(1-s)^{n-1}$, the mutual density function for two coordinates, s_1, s_2 is $p_2(s_1, s_2) = n(n-1)(1-s_1-s_2)^{n-2}$, for k coordinates $s_1, s_2, \ldots, s_k$ $(k<n)$ the mutual density is

$$p_k(s_1, s_2, \ldots, s_k) = \frac{n!}{(n-k)!}\left(1 - \sum_{i=1}^{k} s_i\right)^{n-k}.$$

The first moments are: $\mathbf{E}(s) = 1/(n+1) = 1/n + o(1/n)$, $\mathbf{E}(s^2) = 2/[(n+1)(n+2)] = 2/n^2 + o(1/n^2)$,

$$\mathrm{Var}(s) = \mathbf{E}(s^2) - (\mathbf{E}(s))^2 = \frac{n}{(n+1)^2(n+2)} = \frac{1}{n^2} + o\left(\frac{1}{n^2}\right),$$

and for $k<n$,

$$\mathbf{E}(s^k) = \frac{1}{C_n^k} = \frac{(n-k)!k!}{n!} = \frac{k!}{n^k} + o\left(\frac{1}{n^k}\right)$$

(the last equality holds for any given k and $n \to \infty$). For the first mixed moments we get $\mathbf{E}(s_1 s_2) = 1/[(n+1)(n+2)] = 1/n^2 + o(1/n^2)$,

$$\mathrm{Cov}(s_1, s_2) = \mathbf{E}(s_1 s_2) - \mathbf{E}(s_1)\mathbf{E}(s_2) = -\frac{1}{(n+1)^2(n+2)} = -\frac{1}{n^3} + o\left(\frac{1}{n^3}\right),$$

and for the correlation coefficient

$$\mathrm{Cor}(s_1, s_2) = \frac{\mathrm{Cov}(s_1, s_2)}{\sqrt{\mathrm{Var}(s_1)\mathrm{Var}(s_1)}} = -\frac{1}{n}.$$

It is worth mentioning that $\mathrm{Cov}(s_1, s_2)$ has order n^{-3}, $\mathrm{Var}(s)$ has order n^{-2}, hence, correlations between coordinates decrease as n^{-1} (coordinates become independent for large n, and correlation decrease is a symptom of this independence). It is easy to calculate moments of the square of the Euclidean radius $R^2 = \sum_{i=1}^{n+1}(s_i - \mathbf{E}(s_i))^2$, for example,

$$\mathbf{E}(R^2) = (n+1)\mathrm{Var}(s) = \frac{n}{(n+1)(n+2)} = \frac{1}{n} + o\left(\frac{1}{n}\right),$$

and the Chebyshev's inequality gives the simplest estimate:

$$\mathrm{Vol}\{x \in \Delta_n : R^2 > \rho^2\} \leqslant \frac{1}{\rho^2 n}, \tag{6}$$

up to the leading order in n.

With the higher moments of R^2 we can obtain estimates with the higher powers of $1/n$, but already a simple geometrical consideration gives exponential estimates. For any $i = 1, \ldots, n$, the part of Δ_n, where $s_i \geqslant \varepsilon$, has the normalized volume

$$\mathrm{Vol}(s \in \Delta_n : s_i \geqslant \varepsilon) = (1-\varepsilon)^n \approx \exp(-\varepsilon n). \tag{7}$$

Hence, the set $K_\varepsilon \subset \Delta_n$, where $s_i < \varepsilon$ for all $i = 1, \ldots, n+1$, has the normalized volume

$$V_\varepsilon \geqslant (1 - (n+1)(1-\varepsilon)^n) \approx 1 - n\exp(-\varepsilon n). \tag{8}$$

For any point $x = (s_1, \ldots, s_{n+1}) \in K_\varepsilon$ the following inequality holds: $R^2 = \sum_{i=1}^{n} s_i^2 \leqslant \sum_{i=1}^{n} \varepsilon s_i = \varepsilon$. Therefore, the intersection of Δ_n and a Euclidean ball $B_\rho^{(n+1)}$ with the center c includes the set K_ε, if $\varepsilon \leqslant \rho^2$. Hence, for the normalized volume of this intersection, W_ρ, the following inequalities hold:

$$\begin{aligned}\mathrm{Vol}\{x \in \Delta_n : R^2 > \rho^2\} &= W_\rho \geqslant V_{\rho^2} \\ &\geqslant (1 - (n+1)(1-\rho^2)^n) \approx 1 - n\exp(-\rho^2 n).\end{aligned} \tag{9}$$

The estimate (9) implies that for any given positive constant $a < 1$ there exists a positive constant b such that $W_{b\ln n/\sqrt{n}} > a$ for all n. In other words, for any given share a of the simplex volume there exists such a constant $b > 0$ that the Euclidean ball $B^{(n+1)}_{b\ln n/\sqrt{n}}$ with the center c includes this part of the volume for all n. We can guarantee with (9) that the radius of such a ball goes to zero as $\ln n/\sqrt{n}$.

A precise analysis of the concentration effects in L_p balls and in a standard simplex was performed in Refs. [7,17].

The concentration of a simplex measure in a small vicinity of its center can be considered as an effect that is opposite to the Gibbs concentration of volume of a n-dimensional ball B_n in a small vicinity of its boundary, the sphere. On the other hand, it is similar to the waist concentration. And now not only the values of macroscopic projections can be separated into two sets: one with a microscopically small diameter, the other with a microscopically small measure, but also the set of the symmetrized microscopic states. The symmetrization with respect to particles permutations plays the same role as the macroscopic projection. We can say now that this *symmetrization is the main step in the micro–macro transformation*.

2.3. Symmetric coordinates for multidimensional phase space

In order to demonstrate the same effect for one-particle configuration space (or phase space) of non-unit dimension, let us consider a product of m simplices Δ_n for $m \sim n^\alpha$ and some power α. Euclidean diameter of Δ_n^m grows with m as $\sqrt{m} \sim n^{\alpha/2}$, Euclidean diameter of the product of Euclidean balls $(B_\rho^{(n+1)})^m$ is $2\sqrt{m}\rho \sim \rho n^{\alpha/2}$. For the normalized volume of the intersection $\Delta_n^m \cap B^{m(n+1)}_{\sqrt{m}\rho}$ the following estimate holds:

$$\begin{aligned}\mathrm{Vol}(\Delta_n^m \cap B^{m(n+1)}_{\sqrt{m}\rho}) &\geqslant \mathrm{Vol}(\Delta_n^m \cap (B_\rho^{(n+1)})^m) \\ &\approx (1 - n\exp(-\rho\sqrt{n}))^m \sim 1 - n^{1+\alpha}\exp(-\rho\sqrt{n}).\end{aligned} \tag{10}$$

From this estimate it follows that the strong order–disorder separation holds for these Cartesian degrees of simplex also (if $m\sim n^{\alpha}$): almost all volume belongs to an Euclidean ball with the relatively small diameter $R\sim\rho n^{\alpha/2}$. In order to include in this ball any given share of volume we can choose $\rho\sim n^{-1/2}$ with appropriate value of the prefactor. Therefore, the correspondent relation of diameters $R/\mathrm{Diam}(\Delta_n^m)$ goes to zero as $n^{-1/2}$.

Let one-particle space be k-dimensional unit cube Q_k. The space for n-particle system is $(Q_k)^n$. We produce the symmetric map of $(Q_k)^n$ onto product of $n^{1/k}$ dimensional simplices $(\Delta_{n^{1/k}})^{kn^{(n-1)/k}}$ as follows. Let ξ_i, $i=1,\dots,k$, $0\leqslant\xi\leqslant 1$ be coordinates in Q_k. With each coordinate axis we construct a projection of $(Q_k)^n$ onto $(\Delta_{n^{1/k}})^{n^{(n-1)/k}}$. The product of k such projections is the resulting map $(Q_k)^n\to(\Delta_{n^{1/k}})^{kn^{(n-1)/k}}$.

For ξ_k, this projection is the top floor of the "staged tower" of symmetric coordinates. Let us first enumerate particle in the order of ξ_1 value: $0=x_0\leqslant x_1\leqslant x_2\leqslant\cdots\leqslant x_n\leqslant x_{n+1}=1$. First set (the ground floor of the "staged tower") of symmetric coordinates is: $s_i=x_i-x_{i-1}$, where $i=1,\dots,n+1$.

Let us divide the particles into $n^{1/k}$ groups G^l, $l=1,\dots,n^{1/k}$ with $n^{(k-1)/k}$ elements in each group in the same order: first $n^{(k-1)/k}$ particles with coordinates $\xi_1=x_1,x_2,\dots,x_{n^{(k-1)/k}}$ belong to the first group, G^1, then follow $n^{(k-1)/k}$ particles from the second group, etc. Let us enumerate particle of each group in the order of ξ_2 value: $0=x_0^l\leqslant x_1^l\leqslant x_2^l\leqslant\cdots\leqslant x_{n^{(k-1)/k}}^l\leqslant x_{n^{(k-1)/k}+1}=1$, where superscript l is the group number.[3]

The first floor of the "staged tower" consists of $n^{1/k}$ sets of symmetric coordinates $s_i^l=x_i^l-x_{i-1}^l$. After that, we can divide each G^l into $n^{1/k}$ groups G^{lm}, $m=1,\dots,n^{1/k}$ with $n^{(k-2)/k}$ elements in each group in the order of ξ_2 value. Let us enumerate particles of each group in the order of ξ_3 value: $0=x_0^{lm}\leqslant x_1^{lm}\leqslant x_2^{lm}\leqslant\cdots\leqslant x_{n^{(k-2)/k}}^{lm}\leqslant x_{n^{(k-2)/k}+1}^{lm}=1$. The second floor consists of $n^{2/k}$ sets of symmetric coordinates $s_i^{lm}=x_i^{lm}-x_{i-1}^{lm}$. Finally, we get k floors (from the ground to $(k-1)$st one). The floor number j ($j=0,\dots,k-1$) consists of $n^{j/k}$ groups of symmetric coordinates with $n^{(k-j)/k}$ coordinates in each group. These coordinates are non-negative, their sums in groups are equal to 1. Therefore, each floor represents a n-dimensional polyhedron that is a product of $n^{j/k}$ standard simplices $\Delta_{n^{(k-j)/k}}$ of dimension $n^{(k-j)/k}$, and the whole tower represents the following product of simplices:

$$\Omega_{k,n}=\prod_{j=0}^{k-1}\Omega_j\quad\text{where }\Omega_j=(\Delta_{n^{(k-j)/k}})^{n^{j/k}}. \tag{11}$$

We are interested in the $(k-1)$st floor that corresponds to ξ_k. It is $\Omega_{k-1}=(\Delta_{n^{1/k}})^{n^{(n-1)/k}}$. Analogous projection for other ξ_k could be obtained by coordinates permutation (cyclic).

We see that for a k-dimensional one-particle space the result is qualitatively the same as for one-dimensional. The only difference is that here the estimates guarantee that the relative Euclidean radius (that is, the relation of the radius to the diameter of the whole space) of the set, where an arbitrary part $a<1$ of measure is concentrated, tends to zero as $1/\sqrt{d}$, instead of $1/\sqrt{n}$. Here, d is the dimension of one simplex from the product, that is, $d=n^{1/k}$ and the relative radius goes to zero as $n^{-1/(2k)}$.

This change of order reflects a simple fact: the typical distance from a particle to the nearest particle in dimension k is $\sim n^{-1/k}$. After summation of d squares of such variables we get the square of radius: $R^2\sim n^{-1/k}$. Then, we take the $n^{(k-1)/k}$th power of d-dimensional simplex and of the ball from this simplex also. The relation of the Euclidean radii does not change after this operation. It remains $\sim n^{-1/(2k)}$.

The same result holds for one-particle space P that is not a cube, but a bi-Lipschitz image of a cube, or can be covered by finite number of such images. We discuss much more general metric-measure (mm) spaces in Section 2.4.

2.4. Other natural distances on symmetrized states

A metric space P with distance $d(x,y)$ and a given measure μ on P is a mm-space [7], if every metric ball is measurable. In this section, we discuss distribution of particles in a mm-space P with a probability measure μ, hence, $\mu(P)=1$.

[3]Of course, it is more rigorous to speak about integer parts of numbers: G^1 consists of IntegerPart($n^{(k-1)/k}$) elements, G^2 consists of IntegerPart($2n^{(k-1)/k}$) − IntegerPart($n^{(k-1)/k}$) elements, and so on, but it adds nothing to the sense, only the notations become cumbersome.

We assume that P is compact[4] and, hence, has a finite diameter. The space of (Radon) measures on P is $C^*(P)$, that is the conjugated space to the space of continuous functions $C(P)$ on P. The action of a measure $v \in C^*(P)$ on a function $f \in C(P)$ is the number $[v,f]$. The action of probability measure μ on f is the expectation: $[\mu,f] = \mathbf{E}(f)$. The probability measures on P are positive and normalized elements of $C^*(P)$. For measures, we use the weak* convergence: $\mu_i \to \mu_0$ if $[\mu_i,f] \to [\mu_0,f]$ for every continuous function f.

An unordered tuple of n points (''particles'') in a *mm*-space P can be represented as a probability measure:

$$\{x_1, x_2, \ldots, x_n\} \mapsto \mu_{x_1,x_2,\ldots,x_n} = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}, \tag{12}$$

where δ_{x_i} is a unit measure concentrated at the point x_i (δ-function). The *law of large numbers* states that $\mu_{x_1,x_2,\ldots,x_n} \to \mu$ for almost all sequences $\{x_1, x_2, \ldots, x_n, \ldots\} \in P^{\infty}$. ''Almost all'' means here: the set of exceptions has zero measure. Let f be an arbitrary bounded continuous function on P. The standard law of large numbers for a random variable f and the probability space P immediately gives the weak* convergence: the sequence of averages $\langle f \rangle_n = \frac{1}{n}\sum_{i=1}^{n} f(x_i)$ converges to the average value of f with respect to the probability measure μ, that is, to the expectation $\mathbf{E}(f)$.

For each domain $W \subset P$ with non-zero measure $\mu(W)$ the probability that all particles are outside W is

$$\mathbf{P}\{x_i \notin W : i = 1, \ldots, n\} = (1 - \mu(W))^n. \tag{13}$$

This estimate is analogous to the estimate of the volume of one wing of the n-dimensional standard simplex (7).

Moreover, *all* balls B_ρ of given radius ρ are non-empty with almost unit probability. Let us take a $\rho/2$ net $\{y_1, \ldots, y_m\}$ in P, where $m = \mathrm{Cap}_{\rho/2}(P)$ is the minimum number of points in $\rho/2$ net in P. Each ball B_ρ in P includes a ball $B_{\rho/2}(y_i)$ of radius $\rho/2$ and center in one of the points $y_1, \ldots, y_m$. Therefore, if there is a ball $B_\rho \subset P$ free of particles, then at least one of the balls $B_{\rho/2}(y_i)$ $(i = 1, \ldots, m)$ is also free of particles

$$\begin{aligned}&\mathbf{P}\{\text{Each ball } B_\rho \subset P \text{ includes a particle}\}\\ &\geqslant 1 - \sum_{i=1}^{\mathrm{Cap}_{\rho/2}(P)} (1 - \mu(B_{\rho/2}(y_i)))^n \geqslant 1 - \mathrm{Cap}_{\rho/2}(P)(1 - \underline{\mu}(\rho/2))^n,\end{aligned} \tag{14}$$

where $\underline{\mu}(\rho) = \inf\{\mu(B_\rho(y)) : y \in P\}$. This estimate is similar to the estimate (8) of the joint volume of the n-dimensional simplex wings. In the final estimate (14) two characteristics of the *mm*-space P are used: the minimum number of points in $\rho/2$ net, $\mathrm{Cap}_{\rho/2}(P)$, and the minimal volume of a ball of radius $\rho/2$, $\underline{\mu}(\rho/2)$. There is also a difference between (8) and (14): the analogue for the ''number of wings'' for the last estimate, $\mathrm{Cap}_{\rho/2}(P)$, does not depend on n.

Natural metrization of the space of probability measures on P in the weak* convergence gives the following metric [7,20]:

$$Lid(v, \eta) = \sup_f |[v - \eta, f]|, \tag{15}$$

where f runs over all 1-Lipschitz functions on P.

Another, functional representation of an unordered tuple of n points (''particles'') in a *mm*-space P is a continuous function

$$\{x_1, x_2, \ldots, x_n\} \mapsto f_{x_1,x_2,\ldots,x_n} : f_{x_1,x_2,\ldots,x_n}(x) = \min_{i=1,\ldots,n} d(x, x_i). \tag{16}$$

This functional representation is an exact analogue for the simplex representation (4): in one-dimensional case, the maximum norm of the function $f_{x_1,x_2,\ldots,x_n}$ is $\frac{1}{2}\max_i s_i$, the average of $|f_{x_1,x_2,\ldots,x_n}(x)|^p$ is $\frac{1}{2^p(p+1)}\sum_i s_i^{p+1}$.

[4]Generalization of most statements to complete, but non-compact *mm*-spaces (for example, to important case of locally compact space) is often possible because the probability measure μ is concentrated on a compact subset of P up to any given accuracy, and after cutting a ''tail'' of distribution μ we can return to compact space. The theory of large deviations and equidistribution in general spaces is presented in Refs. [18,19].

Particularly, the square of the Euclidean (i.e., L_2) norm in a simplex is proportional to L_1 norm of the function $f_{x_1,x_2,\ldots,x_n}$.

For this representation, the estimate (14) has a simple form

$$\begin{aligned}\mathbf{P}\{\|f_{x_1,x_2,\ldots,x_n}\|_{L_\infty}<\rho\} &\geqslant 1-\sum_{i=1}^{m}(1-\mu(B_{\rho/2}(y_i)))^n \\ &\geqslant 1-m(1-\underline{\mu}(\rho/2))^n,\end{aligned} \tag{17}$$

where $\|\bullet\|_{L_\infty}$ is the maximum norm.

In many practically important *mm*-spaces P, the volume of balls is of order ρ^k for some power $k>0$: $\inf \mu(B_\rho)/\rho^k = a>0$. In that case, for sufficiently small ρ and large n

$$(1-\underline{\mu}(\rho/2))^n \approx \exp(-an(\rho/2)^k),$$

and

$$\mathbf{P}\{\|f_{x_1,x_2,\ldots,x_n}\|_{L_\infty}<\rho\} \geqslant 1-\mathrm{Cap}_{\rho/2}(P)\exp(-an(\rho/2)^k), \tag{18}$$

where $\mathrm{Cap}_{\rho/2}(P)$ is the minimum number of points in $\rho/2$ net in P.

For L_1 norm of $f_{x_1,x_2,\ldots,x_n}$ (analogue for the square of the Euclidean norm in a simplex of symmetrical coordinates) we obtain the estimates

$$\|f_{x_1,x_2,\ldots,x_n}\|_{L_1} = \mathbf{E}(f_{x_1,x_2,\ldots,x_n}) \leqslant \max_P |f_{x_1,x_2,\ldots,x_n}(x)|$$

and

$$\begin{aligned}\mathbf{P}\{\|f_{x_1,x_2,\ldots,x_n}\|_{L_1}\leqslant b\} &\geqslant \mathbf{P}\{\|f_{x_1,x_2,\ldots,x_n}\|_{L_\infty}<b\} \\ &\geqslant 1-\mathrm{Cap}_{b/2}(P)\exp(-an(b/2)^k).\end{aligned} \tag{19}$$

And again we observe the simplex-type strong order–disorder separation. In the maximum norm, the whole set of functions that represent n-point tuples has diameter $\mathrm{Diam}(P)$. The measure is concentrated in a ball or radius $\sim n^{-1/k}$ (18) (in the maximum norm also).

There exists a simple connection between measure (12) and functional representations (16) of n-particle systems. Let the radius of a ball with the centre x and volume δ be a continuous function $r_\delta(x)$ for any $\delta>0$. For each continuous function $f(x)$, the L_1 function $f^\delta(x)$ is defined:

$$f^\delta(x) = \begin{cases} \dfrac{1}{\delta}, & \text{if } f(x)<r_\delta(x); \\ 0, & \text{if } f(x)\geqslant r_\delta(x). \end{cases} \tag{20}$$

The distribution $\frac{1}{n}f^\delta_{x_1,x_2,\ldots,x_n}\mu$ approximates the measure $\mu_{x_1,x_2,\ldots,x_n}$ (12) when $\delta\to 0$.

2.5. Statistics of local structures

In this paper, we discuss the statistics of large sets of particles with permutation invariance. For spaces of k-particle sets, two embeddings are considered: into space of measures (12) and into spaces of functions (16). These embeddings are useful for theoretical purposes, but for practical needs embeddings into finite-dimension Euclidean space are necessary, as well as systems of internal coordinate charts on spaces of k-particle sets with one labeled point. In this subsection, we discuss the embedding and coordinate choice and give a non-technical introduction into statistical analysis on non-Euclidean metric spaces.

Molecular dynamics gives us many examples of particle configurations. We never had such detailed information before, and the question is how to process it with maximally useful output. The classical approach of statistical physics is based on k-particle distribution functions for small k. It is not sufficient, for example, for the following problems.

Let the configuration of n particles be given: we know all the positions of molecules. For each particle (point x) and any $k<n$ we define a k-particle local configuration, or *germ* (k-germ) of the configuration at x, that is the set of k particles nearest to x represented in the reference system with origin x. The set of k-germs for all

possible central particles form a cloud of points in the space of k-germs. Are there clusters or clots in this cloud? Is the distribution of k-germs in physical space (R^3) homogeneous? If it is heterogeneous, then how can we find boundaries between locally homogeneous clusters?

For systems in isotropic conditions, instead of k-germs it is necessary to consider orbits of k-germs under the action of rotation group. Molecules with internal structure can also be considered without principal problems (but with some technical complications).

The problem of local heterogeneities in water is most attractive [21–23]. But, even for hard spheres systems the cluster boundaries localization is not trivial.

It is not obvious, how many particles in local configuration should we take into account: where the heterogeneities are hidden. It is necessary to study statistics of k-germs for different k and evaluate the informativity of transition from k to $k+1$.

The classical *statistical geometry* gives some tools for quantitative analysis of configuration structure. Important sources of ideas and methods for the local configuration analysis are the theory of random packing [24], the molecular geometry of liquids [25] and the theory of liquid–glass transition [26]. The main tool of statistical geometry that is in wide use for molecular dynamics data mining is the analysis of the Voronoi polyhedra and the Delaunay simplices statistics [27–29]. The Voronoi polyhedron is the domain around a particle, such that all points of this domain are closer to this particle than to any other. A group of four particles, whose Voronoi polyhedra meet at one vertex, forms another basic object of statistical geometry, the Delaunay simplex. Statistics of the Voronoi polyhedra and the Delaunay simplices gives us information about local order in the *nearest vicinity* of particles: the Voronoi polyhedron describes the coordination of the nearest atomic environment while the Delaunay simplex describes the shape of the cavities between the nearest atoms.

In order to extend this vicinity to an arbitrary number of neighbors and coordination spheres, we need the statistics of k-germs. We propose systematic study of statistics of k-germs: (non-linear) principal component analysis (PCA), cluster analysis, and analysis of the fields of obtained statistical characteristics in the physical space–time. Many old questions could be revisited in this way, especially the problems of local heterogeneity.

The crucial question is the choice of a space where the statistical analysis will be performed.

For statistical computations, the embedding of the space of germs into Euclidean space is convenient. For any sequence of functions in R^m, $\mathscr{F} = \{f_1, \ldots, f_M\}$ let us define

$$\mathscr{F}(x_1, \ldots, x_k) = \left\{ \frac{1}{k} \sum_{i=1}^{k} f_j(x_i) \right\}_{j=1}^{M}. \tag{21}$$

These coordinates serve for computation of distance ρ between germs (just a standard Euclidean distance in these coordinates can be chosen; the second choice is the locally Euclidean Riemannian metric, the geodesic distance).

For systems with rotational symmetry, it is necessary to study statistics of rotational orbits of germs. The space of functions spanned by $\{f_1, \ldots, f_M\}$ should be rotationally invariant and represented as a sum of irreducible subspaces. In this case, the coordinate tuple for a germ (21) is a direct sum of irreducible tensors, and it is easy to write the complete system of rotational invariants and to define invariant distance on the space of germs.

For this purpose, it is convenient to choose $\{f_1, \ldots, f_M\}$ as eigenfunctions of a Schrödinger operator with central force, sorted by eigenvalues and momentum (isotropic oscillator eigenfunctions, for example). These eigenfunctions are spherical harmonics multiplied on radial functions $f_j(\rho)$ that decay when $\rho \to \infty$, hence, a far particle has less influence on the distance between germs than the nearest one.

Statistics in Euclidean spaces with coordinates (21) is not statistics of germs: the average of germs for this Euclidean statistics is already not a germ. Let us consider the space of germs as a non-Euclidean metric space with metric ρ.

Following Frechet [30], we can define an average point $\langle z \rangle$ for a finite subset $\{z_1, \ldots, z_q\}$ of a metric space $\mathscr{K}$ as a minimizer of average squared distance

$$\langle z \rangle = \operatorname*{argmin}_{z \in \mathscr{K}} \left\{ \frac{1}{q} \sum_j \rho^2(z, z_j) \right\}. \tag{22}$$

On the basis of this approach, statistics on Riemannian spaces is developed, from simple averaging to moments calculation and definition of normal distribution [31]. For shape statistics, the method of principal geodesic analysis is proposed, that is a generalization of PCA to the manifold setting [32].

We can interpret the Frechet averaging (22) as minimization of elastic energy of springs that connect data points with an average point. The statistical analysis on metric spaces may be represented as minimization of "elastic energy" [33–35]. This energetic metaphor works successfully for model reduction problems, cluster analysis and analysis of data with complex topology. Let us give a sketch of this approach following Ref. [36]. For simplicity, we consider a metric space embedded into Euclidean space with Euclidean distance between points.

Let G be a simple undirected graph with set of vertices Y and set of edges E. For $k \geqslant 2$, a k-star in G is a subgraph with $k+1$ vertices $y_{0,1,\dots,k} \in Y$ and k edges $\{(y_0, y_i) \mid i = 1, \dots, k\} \subset E$. Suppose for each $k \geqslant 2$, a family S_k of k-stars in G has been selected. We call a graph G with selected families of k-stars S_k an *elastic graph* if, for all $E^{(i)} \in E$ and $S_k^{(j)} \in S_k$, the correspondent elasticity moduli $\lambda_i > 0$ and $\mu_{kj} > 0$ are defined. Let $E^{(i)}(0), E^{(i)}(1)$ be vertices of an edge $E^{(i)}$ and $S_k^{(j)}(0), \dots, S_k^{(j)}(k)$ be vertices of a k-star $S_k^{(j)}$ (among them, $S_k^{(j)}(0)$ is a central vertex). For any map $\phi : Y \to R^m$ the *energy of the graph* is defined as

$$U^{\phi}(G) := \sum_{E^{(i)}} \lambda_i \|\phi(E^{(i)}(0)) - \phi(E^{(i)}(1))\|^2 + \sum_{S_k^{(j)}} \mu_{kj} \left\| \sum_{i=1}^{k} \phi(S_k^{(j)}(i)) - k\phi(S_k^{(j)}(0)) \right\|^2. \tag{23}$$

Very recently, a simple but important fact was noticed [37]: every system of elastic finite elements could be represented by a system of springs, if we allow some springs to have negative elasticity coefficients. The energy of a k-star s_k in R^m with y_0 in the centre and k endpoints $y_{1,\dots,k}$ is $u_{s_k} = \mu_{s_k}(\sum_{i=1}^{k} y_i - k y_0)^2$, or, in the spring representation, $u_{s_k} = k\mu_{s_k} \sum_{i=1}^{k} (y_i - y_0)^2 - \mu_{s_k} \sum_{i>j} (y_i - y_j)^2$. Here, we have k positive springs with coefficients $k\mu_{s_k}$ and $k(k-1)/1$ negative springs with coefficients $-\mu_{s_k}$.

For a given map $\phi : Y \to R^m$, we divide the dataset D into subsets K^y, $y \in Y$. The set K^y contains the data points for which the node $\phi(y)$ is the closest one in $\phi(Y)$. The *energy of approximation* is

$$U_A^{\phi}(G, D) := \sum_{y \in Y} \sum_{x \in K^y} w(x) \|x - \phi(y)\|^2, \tag{24}$$

where $w(x) \geqslant 0$ are the point weights.

The simple and very popular algorithm for minimization of the energy $U^{\phi} = U_A^{\phi}(G, D) + U^{\phi}(G)$ is the splitting algorithm, in the spirit of the classical k-means clustering: for a given system of sets $\{K^y \mid y \in Y\}$ we minimize U^{ϕ}, then for a given ϕ we find new $\{K^y\}$, and so on; stop when no change occurs. This is the constrained minimization: the nodes move along the k-germs space embedded into Euclidean space, while the distance in this example is Euclidean one. This algorithm gives a local minimum, and the global minimization problem arises. There are many methods for improving the situation, but without guarantee of the global minimization.

The next problem is the elastic graph construction. Here, we should find a compromise between simplicity of graph topology, simplicity of geometrical form for a given topology, and accuracy of approximation. Geometrical complexity is measured by the graph energy $U^{\phi}(G)$, and the error of approximation is measured by the energy of approximation $U_A^{\phi}(G, D)$. Both are included in the energy U^{ϕ}. Topological complexity will be represented by means of elementary transformations: it is the length of the energetically optimal chain of elementary transformation from a given set applied to initial simple graph.

Graph grammars [38,39] provide a well-developed formalism for the description of elementary transformations. An elastic graph grammar is presented as a set of production (or substitution) rules. Each rule has a form $A \to B$, where A and B are elastic graphs. When this rule is applied to an elastic graph, a copy of A is removed from the graph together with all its incident edges and is replaced with a copy of B with edges that connect B to graph. For a full description of this language, we need the notion of a *labeled graph*. Labels are necessary to provide the proper connection between B and the graph.

A link in the energetically optimal transformation chain is constructing by finding a transformation application that gives the largest energy descent (after an optimization step), then the next link, and so on, until we achieve the desirable accuracy of approximation, or the limit number of transformations (some other termination criteria are also possible).

As a simple (but already rather powerful) example, we use a system of two transformations: "add a node to a node" and "bisect an edge." These transformations act on a class of *primitive elastic graphs*: all non-terminal nodes with k edges are centers of elastic k-stars, which form all the k-stars of the graph. For a primitive elastic graph, the number of stars is equal to the number of non-terminal nodes—the graph topology prescribes the elastic structure.

The transformation "*add a node*" can be applied to any vertex y of G: add a new node z and a new edge (y,z). The transformation "*bisect an edge*" is applicable to any pair of graph vertices y, y' connected by an edge (y,y'): delete edge (y,y'), add a vertex z and two edges, (y,z) and (z,y'). The transformation of elastic structure (change in the star list) is induced by the change of topology, because the elastic graph is primitive. This two-transformation grammar with energy minimization builds *principal trees* (and principal curves, as a particular case) for datasets.

For applications, it is useful to associate with these principal trees one-dimensional continuums. Such a continuum consists of node images $\phi(y)$ and of pieces of lines that connect images of linked nodes.

The *first task* of k-germs statistical analysis is dimension reduction. The method of choice here is PCA. Its linear version is now classics and textbook material [40], and non-linear PCA is developed recently [33,35,41,42]. The methods of elastic manifolds and graphs [35] does not require Euclidean space of data. The *second task* is cluster analysis. The described method of elastic graphs is a tool for non-linear PCA, and for cluster analysis, both.

The *third task* that is specific for statistical physics is the analysis of k-germs distribution in physical space–time. After that, we can discuss structural non-uniformity, quasi-chemical representation of kinetics [5], and many other topics. Of course, additional topological information about various bonds between particles could be added to this metric description.

Internal coordinates on the space of germs are necessary for gradient optimization of energy. Topologically, the space of k-germs near a point x can be defined as the space of permutation orbits. Let us enumerate k particles $x_1, x_2, \dots, x_k$ nearest to the point x in order of their distance to x, $\rho_i = \|x_i - x\|$: $\rho_1 < \rho_2 < \dots < \rho_k$. If all particles are in generic positions, then any two distances are distinct. This ordered representation $\{x_1, \dots, x_k\}$ has discontinuity points when some ρ_i coincide.

The following internal coordinates on the space of rotational orbits of germs give generically a representation of these orbits with discontinuity points when some ρ_i coincide.

Let us enumerate k particles $x_1, x_2, \dots, x_k$ nearest to the point x in order of their distance to x, $\rho_i = \|x_i - x\|$: $\rho_1 < \rho_2 < \dots < \rho_k$. We assume that all particles are in generic positions, hence, any two distances are distinct and three particles could not belong to one straight line. The distances from x_i to x, x_1, x_2 will be the main coordinates of the k-germ. These are $3k-3$ numbers: $\{\rho_i\}_{i=1,\dots,k}$, $\{\rho'_j\}_{j=2,\dots,k}$, $\{\rho''_l\}_{l=3,\dots,k}$, where $\rho'_j = \|x_j - x_1\|$, $\rho''_l = \|x_l - x_2\|$. An additional set of coordinates consists of $k-2$ signs, $\sigma_i = \pm 1$, $i = 3, \dots, k$. The triangle $\{x, x_1, x_2\}$ belongs to a plane Γ. This plane divides the space into two half-spaces, L_+ and L_-. We define the signs subscripts $L_{+,-}$ by triangle orientation $\{x, x_1, x_2\}$ according to the standard "screw rule" (or the "right-hand rule"). The sign $\sigma_i = +1$, if $x_i \in L_+$, $\sigma_i = -1$, if $x_i \in L_-$. Generically, there are no particles on Γ. The whole set of coordinates consists of $3k-3$ real numbers and $k-2$ signs. If we, in addition to rotation symmetry, assume the reflection symmetry, then there are only $k-3$ signs: $\sigma'_j = \sigma_3\sigma_j$, $j = 4, \dots, k$. The worst violations of the continuity condition for the proposed coordinates are jumps of basis triangle x, x_1, x_2 near some of the configurations. For example, for a body-centered cubic lattice there are three non-equivalent choices of particles x_1, x_2 nearest to the central particle x (in this symmetric case, $\rho_1 = \rho_2$): along the cube edge, along a face diagonal, and along a main diagonal of the cube. Therefore, in the vicinity of this symmetric configuration, the distance $\rho'_2 = \|x_2 - x_1\|$ cannot be a continuous function of k-germ ($k \geqslant 4$).

Statistical theory of shapes of finite sets in R^3 was launched in 1970s (see a survey [43]). Statistical analysis of configuration germs is an interdisciplinary area between statistics of shapes and statistical physics.

2.6. Dynamics in systems with strong order–disorder separation

In previous subsections, we discussed relations between measure and distance in high-dimensional systems with permutational symmetry. But the main property of the measure under consideration is its *invariance* with respect to mechanical motion. In this subsection, we consider dynamics in phase spaces with concentration.

Without such a return to dynamics the consideration of order–disorder relations in statistical mechanics is incomplete, and we can loose some important effects.

The strong order–disorder separation causes very special peculiarities of dynamical systems with conservation of measure. Let P_n be the n-particle phase space for a system with strong order–disorder separation. Assume that for given $\delta>0$, the radius $\rho(\delta,n)$ of a $(1-\delta)$-concentration ball $B^{n\,\mathrm{conc}}_{\rho(\delta,n)}$ with the $1-\delta$ goes to 0 at $n\to\infty$ and the diameter of P_n is bounded: $\alpha\geqslant \mathrm{Diam}\, P_n\geqslant\beta>0$.[5] Phase flow transformations form a one-dimensional semigroup of injective maps $T_t: P_n\to P_n$, T_t $(t>0)$ is a shift over time t. For any $t>0$, the map T_t keeps the most part of the $1-\delta$-concentration ball $B^{n\,\mathrm{conc}}_{\rho(\delta,n)}$ in it:

$$\mathbf{P}(T_t(B^{n\,\mathrm{conc}}_{\rho(\delta,n)})\cap B^{n\,\mathrm{conc}}_{\rho(\delta,n)})\geqslant 1-2\delta, \tag{25}$$

because the measure of complement of $B^{n\,\mathrm{conc}}_{\rho(\delta,n)}$ in P_n is less than δ.

For time averages of bounded differentiable vector-functions on $[0,\infty[$ (with bounded derivatives) an elementary identity holds:

$$\langle \dot{f}^2(t)\rangle = -\langle (f(t),\ddot{f}(t))\rangle, \tag{26}$$

if all averages exist, hence,

$$\langle \dot{f}^2(t)\rangle \leqslant \langle f^2(t)\rangle^{1/2}\langle \ddot{f}^2(t)\rangle^{1/2}, \tag{27}$$

and

$$\frac{\langle \ddot{f}^2(t)\rangle}{\langle \dot{f}^2(t)\rangle}\geqslant\frac{\langle \dot{f}^2(t)\rangle}{\langle f^2(t)\rangle}. \tag{28}$$

Let us choose the origin in the center of $B^{n\,\mathrm{conc}}_{\rho(\delta,n)}$. In this case, under standard assumptions,

$$\langle a^2\rangle\geqslant\frac{\langle v^2\rangle^2}{\rho^2(\delta,n)}, \tag{29}$$

where $\langle a^2\rangle$ is the average square of acceleration and $\langle v^2\rangle$ is the average square of velocity.

It means that in systems with concentration for given average square of velocity $\langle v^2\rangle$ the average square of acceleration tends to ∞ with the number of particles, even if the velocity (n-particle) remains normalized. Just to imagine the orders let us assume: $\langle v^2\rangle\sim\frac{1}{2}nkT$, $\rho^2(\delta,n)\sim n^{-1}$. In this case, $\langle a^2\rangle\geqslant \mathrm{const}\times n^3$.

Dynamics of particles with elastic collisions on an interval is equivalent to billiards in a multidimensional simplex (see elsewhere, for example, Ref. [44]). Even if the particles are transparent (if there are no physical collisions at all), the symmetric representation of the system (by a point in the simplex) evolves with velocity jumps. These jumps take place every time when the particles change their order on the line.

For a functional representation of moving particles (16) (for any dimension of the one-particle space) the time derivative of $f_{x_1,x_2,\ldots,x_n}(x)$ has a discontinuity when the nearest to x particle changes its number.

In all these cases the motion in symmetric coordinates is only piecewise differentiable, and average square of acceleration does not exist at all (is infinite).

For example, the described *acceleration–dominated dynamics* makes no differences between real physical interaction and jumps of velocities caused by geometry of permutation symmetry. For motion of particles on a line, the particles can be transparent and do not interact at all. In this case one particle will come through the other, but any change of their order on a line causes in symmetric representation jump of velocities. On the other hand, particles can interact, collide, and do not change their order on a line. The result will be the same. For instantaneous elastic collisions the difference does not exist, but for softer potentials the picture of acceleration dominance also holds. The system without interaction is a billiard in the standard n-dimensional simplex. Interaction changes (smoothes) collisions and bends trajectories between them.

[5]We consider compact spaces to avoid trivial technical complications that are needed for locally compact spaces.

3. Sticky faces and natural selection

In this section we discuss general dynamical systems, not necessarily Hamiltonian ones, or systems with conservation of volume. Let a multidimensional simplex be positively invariant with respect to dynamics: if a motion start in this simplex at some time t_0, then it belongs to the simplex at any moment $t > t_0$. For such a dynamical system, we can guess that the motion spends most of the time in a small vicinity of the simplex centre. It is a very natural expectation because of the concentration of the simplex volume near its centre, and some theorems in the form "for a typical dynamical system with positively invariant multidimensional simplex a typical motion spends most of the time in a small vicinity of the simplex centre" could be proved for the appropriate definition of typicalness. But, there exists an important opposite type of dynamic behavior. Let us assume that the faces of the simplex are also positively invariant. In this case, the typical picture of dynamic behavior changes drastically: motions tend to a small vicinity of the small-dimensional skeleton of the simplex.

Let us first explain the sense of such "sticky faces." The standard simplex Δ_n has natural interpretation as a space of n-dimensional probability distributions $p_1, \ldots, p_n$ defined on n states. A dynamic system with positively invariant Δ_n is a kinetic equation. The faces of Δ_n are positively invariant, if the ith state could not be produced from the jth one for $i \neq j$, and only the birth–death rate of ith state depends on the whole distribution $p_1, \ldots, p_n$. It is the general form of *inheritance* property, and such dynamical systems are standard objects for study in mathematical biology after Volterra [45], Lottka, and Gause [46]; review of some modern works could be found in Ref. [47].

The concentration of motions for $t \to \infty$ in a small vicinity of the small-dimensional skeleton of the simplex is exactly the phenomenon of *natural selection* [48,49]. Many physical applications of this phenomenon are known [50–52].

It is easy to demonstrate this phenomenon on the example of n particles moving on an interval $[0, 1]$. The effect of sticky faces implies here that if the position of the ith particle is 0 or 1, then it does not move. Following natural hypothesis of smoothness, we can extract a multiplier x_i near 0 and $(1 - x_i)$ near 1 from the velocity of the particle at the position x_i. It means that in new coordinates $y_i = \ln x_i - \ln(1 - x_i)$, we can expect more or less uniform distribution of particles. But, it is the equidistribution on the whole line. It is impossible in the classical sense, but if we take it seriously, we come to a finite-additive distribution (or to an approximation with equidistributions on a sequence of extended intervals). In any case, the expected number of particles at a given distance from the interval ends (or, in y_i coordinates, at a given bounded interval) should be small in comparison with the total number of particles: almost all particles are concentrated near interval ends.

The whole effect of sticky faces in a simplex means that if some coordinates s_i are zero, then their time derivatives are also zero. For particles moving on $[0, 1]$ it implies that they stick to each other, and for such a system we can observe particle agglutination, in addition to particle concentration near the interval ends.

In order to achieve exact estimation and theorems, it is useful to start with an infinite number of particles. A variant of such a theory for continuous families of particles is developed in Ref. [48], see English version in Ref. [53]. The main result remains the same: for the dynamics on a simplex with sticky boundaries, almost all motions tend to a small vicinity of the small-dimensional skeleton of the simplex. Estimates of the skeleton dimension and asymptotic expansion for motions near this skeleton are also obtained.

4. Discussion

For a large number of particles, the available phase space (or configuration space) could be divided into two parts. One part has microscopically small diameter (part D, disorder), another part (part O, order) has microscopically small measure (volume). This is the strong *order–disorder separation*.

Permutation invariance is crucial for the strong order–disorder separation. For example, the volume of a high-dimensional cube is concentrated near its boundary. After symmetrization the cube transforms into a simplex, and the volume of the simplex is concentrated near its center. Order is in the long, but thin wings of the simplex, while disorder is in the small, but thick vicinity of its center. This effect allows many generalizations for spaces of permutation orbits.

All individual configurations of n distinguishable particles in R^m are equivalent: the measures of their ε-vicinities coincide, and are equal just to a volume of nm-dimensional ball of radius ε. The permutation symmetry enforced us to replace any single configuration of n particles x by the set $\mathbf{S}_n x$ of $n!$ configurations that are generated from x by particles permutations. These finite ensembles are already not equivalent. For a given bounded domain $P \subset R^m$ (a box) and large n, there exists such a configuration x_0 of n particles ("almost equidistribution") in P that ε-fattening of $\mathbf{S}x_0$, $\{\mathbf{S}_n x_0\}_\varepsilon$, has "almost all" volume of the configuration space P^n:

$$\mathrm{Vol}(\{\mathbf{S}_n x_0\}_\varepsilon)/\mathrm{Vol}(P^m) > 1 - \delta$$

for $\varepsilon \sim n^{-1/m}$, and a given small number δ. If such a configuration exists, then, obviously, all points x from $\{\mathbf{S}_n x_0\}_\varepsilon$ have the same property (with twice increased ε):

$$\mathrm{Vol}(\{\mathbf{S}_n x\}_{2\varepsilon})/\mathrm{Vol}(P^m) > 1 - \delta.$$

This finite ensemble $\mathbf{S}_n x$ is not an ε-net in P^n, moreover, the rest, $P^n \backslash \{\mathbf{S}_n x\}_\varepsilon$, has macroscopic (non-small) diameter and macroscopic Hausdorff distance from $\mathbf{S}_n x$. This is the essence of the strong order–disorder separation: disorder has microscopic diameter (but macroscopic, almost all measure), order has microscopic measure (but macroscopic diameter).

The disordered states from the set D are macroscopically indistinguishable, because the distance between them is microscopically small. We can use a notion of "observability" from the control theory, and say that the difference between these states is macroscopically inobservable.

In our definition of order and disorder we use the state-based approach: a state may be ordered or disordered. (Of course, in the definition of order we use ε-fattening (1), hence, an ensemble is present too, but the notion of order relates to states.) The state-based point of view in foundation of statistical physics becomes more popular very recently [54,55].

The time arrow that leads from order to disorder has the following interpretation: if a motion starts from an ordered state, then, after some time, the state becomes disordered, and we can be almost sure that it will remain disordered during time T with microscopically small inverse T^{-1} (probability of fluctuation from disorder to order could be estimated on the basis of Eq. (25)). Neither chaotic dynamics, nor dynamical stirring have any relation to this behavior. Even ergodicity is not especially important: if ergodic components are multiparticle, then the same order–disorder separation is expected on them. But, without strong order–disorder separation it is impossible even to formulate such a statement: if a motion starts from an ordered state, then, after some time, the state becomes disordered, and …

Dynamics of systems with strong order–disorder separation has a very special property: in symmetric representation the average square of acceleration is very high. It can be interpreted as evolution through a series of collisions even for non-interacting particles. It is a hint to a possible solution of an essential open problem, the *problem of indivisible events*. For a macroscopically small time, a small microscopic subsystems can go through "its whole life", from the beginning to the limit state (or, more accurate, to the limit behavior which may be not only a state, but also a type of motion, etc.). The evolution of the microscopic subsystems in a macroscopically small time Δt should be described as an "*ensemble of indivisible events*". An excellent hint is given by the Boltzmann equation with its indivisible collisions, another good hint gives the chemical kinetics with indivisible events of elementary reactions. Now, we understand that the solution could be found in the high acceleration for systems with strong order–disorder separation (Section 2.6), but do not know yet even a form of a proper answer.

The effect of strong order–disorder separation and time arrow direction from order to disorder turn to inverse, if we assume invariance of the boundary (sticky boundaries). If we consider these dynamical systems as kinetic equations, then the effect of sticky boundaries can be presented as inheritance: if some species (or genes—for our choice) are not present in the system now, then they will not appear in the future. In this case, the evolution from disorder to order has a special name: Natural selection. Many applications of this effect are known in physics: from mode selection in lasers to wave turbulence. It is as general as order–disorder separation, and appears together with any sort of inheritance.

The role of the permutation symmetry in statistical physics was discussed many times, from different points of view: as a basic axiom [56,57], as a practical question related to entropy definition and measurement [58,59]; even an ontological status of this assumption was discussed quite thoroughly [60]. In this paper, in addition to

this discussion, we demonstrate the importance of permutation invariance for order–disorder separation and for direction of time arrow from order to disorder.

The idea of measure concentration already affects even applied computer science [61]. The history of physical applications starts more than 100 years ago, and now the measure concentration is one of the central ideas of statistical physics, we should only recognize this properly.

And what about the children- and parents-rooms? Of course, the children-room has no permutation symmetry: any toy has it own sense, and a permutation destroys the sense of the configuration. But in parents-room there is perfect permutation symmetry: the toys there are just some things that should be returned in the toy-box. Hence, the children room is in order, but the parents room is in full disorder. Moreover, the children-room has no order–disorder separation, because each configuration has its own sense: the disorder is impossible in the children-room!

Acknowledgments

I am very grateful to M. Gromov and H.C. Öttinger for inspiring discussions, and to R. Davidchack with A. Zinovyev whose first computer experiments encourage me to publish here recommendations about local statistical analysis of particle configuration.

References

[1] J.L. Lebowitz, Microscopic origins of irreversible macroscopic behavior, Physica A 263 (1999) 516–527.
[2] E.H. Lieb, Some problems in statistical mechanics that I would like to see solved, Physica A 263 (1999) 491–499.
[3] I. Prigogine, Laws of nature, probability and time symmetry breaking, Physica A 263 (1999) 528–539.
[4] V.L. Ginzburg, Nobel lecture: on superconductivity and superfluidity (what I have and have not managed to do) as well as on the "physical minimum" at the beginning of the XXI century, Rev. Mod. Phys. 76 (2004) 981–998.
[5] A.N. Gorban, I.V. Karlin, Invariant Manifolds for Physical and Chemical Kinetics, Lecture Notes in Physics, vol. 660, Springer, Berlin, Heidelberg, 2005.
[6] H.C. Öttinger, Beyond Equilibrium Thermodynamics, Wiley, NJ, 2005.
[7] M. Gromov, Metric Structures for Riemannian and non-Riemannian Spaces. With appendices by M. Katz, P. Pansu, S. Semmes. Translated from the French by Sean Michael Bates. Progress in Mathematics, vol. 152, Birkhauser Boston, Inc., Boston, MA, 1999.
[8] P. Ehrenfest, T. Ehrenfest-Afanasyeva, in: Mechanics Enziklopädie der Mathematischen Wissenschaften, vol. 4, Leipzig, 1911 (Reprinted in: Ehrenfest, P., Collected Scientific Papers, North-Holland, Amsterdam, 1959, pp. 213–300).
[9] A.N. Gorban, I.V. Karlin, H.C. Öttinger, L.L. Tatarinova, Ehrenfest's argument extended to a formalism of nonequilibrium thermodynamics, Phys. Rev. E 63 (2001) 066124.
[10] V. Milman, Topics in asymptotic geometric analysis, Geom. Funct. Anal. Special Volume—GAFA2000 (2000) 792–815.
[11] M. Gromov, Isoperimetry of waists and concentration of maps, Geom. Funct. Anal. 13 (2003) 178–215.
[12] R. Feynman, The Character of Physical Law, Lecture No. 5, Cox and Wyman, London, 1965.
[13] A.I. Khinchin, Mathematical Foundations of Statistical Mechanics, Dover Publications, 1960.
[14] H. Touchette, R.S. Ellis, B. Turkington, An introduction to the thermodynamic and macrostate levels of nonequivalent ensembles, Physica A 340 (1–3) (2004) 138–146.
[15] P. Levy, Problèmes concretes d'analyse fonctionelle, Gauthier-Villars, Paris, 1951.
[16] A.N. Gorban, I.V. Karlin, P. Ilg, H.C. Öttinger, Corrections and enhancements of quasi-equilibrium states, J. Non-Newtonian Fluid Mech. 96 (2001) 203–219.
[17] G. Schechtman, M. Schuckenschlager, Another Remark on the Volume of the Intersection of two L_p^n Balls, Lecture Notes in Mathematics, vol. 1469, Springer, New York, 1991, pp. 174–178.
[18] R. Ellis, Entropy, Large Deviation, and Statistical Mechanics, Springer, New York, 1985.
[19] R.S. Ellis, An overview of the theory of large deviations and applications to statistical mechanics, Scandinavian Actuarial J. 1 (1995) 97–142.
[20] A.N. Shiryaev, Probability, Springer, New York, 1996.
[21] J.R. Errington, P.G. Debenedetti, Relationship between structural order and the anomalies of liquid water, Nature 409 (2001) 318–321.
[22] M. Canpolat, F.W. Starr, M.R. Sadr-Lahijany, A. Scala, O. Mishima, S. Havlin, H.E. Stanley, Local structural heterogeneities in liquid water under pressure, Chem. Phys. Lett. 294 (1998) 9–12.
[23] H.E. Stanley, S.V. Buldyrev, N. Giovambattista, Static heterogeneities in liquid water, Physica A 342 (2004) 40–47.
[24] J.L. Finney, Random packings and structure of simple liquids. 1. Geometry of random close packing, Proc. R. Soc. London Ser. A—Math. Phys. Sci. 319 (1539) (1970) 479–493.

[25] J.L. Finney, Random packings and structure of simple liquids. II. The molecular geometry of simple liquids, Proc. R. Soc. London Ser. A—Math. Phys. Sci. 319 (1539) (1970) 493–495.
[26] M.H. Cohen, G.S. Grest, Liquid-glass transition, a free-volume approach, Phys. Rev. B 20 (3) (1979) 1077–1098.
[27] W. Brostow, M. Chybicki, R. Laskowski, J. Rybicki, Voronoi polyhedra and Delaunay simplexes in the structural analysis of molecular-dynamics-simulated materials, Phys. Rev. B 57 (21) (1998) 13448–13458.
[28] P. Jund, D. Caprion, R. Jullien, Local investigation of the glass transition: molecular dynamics and Voronoi tessellation, Europhys. Lett. 37 (8) (1997) 547–552.
[29] I.I. Vaisman, M.L. Berkowitz, Statistical geometry analysis of water structure, J. Am. Chem. Soc. 114 (20) (1992) 7889–7896.
[30] M. Frechet, Les éléments aléatoires de nature quelconque dans un espace distancié, Ann. Inst. H. Poincaré 10 (1948) 215–310.
[31] X. Pennes, Probabilities and statistics on Riemannian manifolds: a geometric approach, INRIA, Rapport de recherche 5093, Sophia Antipolis, 2004.
[32] P.T. Fletcher, C. Lu, S.M. Pizer, S. Joshi, Principal geodesic analysis for the study of nonlinear statistics of shape, IEEE Trans. Med. Imaging 23 (2004) 995–1005.
[33] A.N. Gorban, A.A. Rossiev, Neural network iterative method of principal curves for data with gaps, J. Comput. Syst. Sci. Int. 38 (5) (1999) 825–831.
[34] A. Zinovyev, Visualization of Multidimensional Data, Krasnoyarsk State University Press Public., 2000, 168pp.
[35] A. Gorban, A. Zinovyev, Elastic principal graphs and manifolds and their practical applications, Computing 75 (2005) 359–379 E-print arXiv cond-mat/0405648.
[36] A.N. Gorban, N.R. Sumner, A.Y. Zinovyev, Topological grammars for data approximation, Appl. Math. Lett. (2006), in press, doi:10.1016/j.aml.2006.04.022. E-print arXiv cs.NE/0603090.
[37] A. Gusev, Finite element mapping for spring network representations of the mechanics of solids, Phys. Rev. Lett. 93 (2) (2004) 034302.
[38] M. Nagl, Formal languages of labelled graphs, Computing 16 (1976) 113–137.
[39] M. Löwe, Algebraic approach to single-pushout graph transformation, Theor. Comp. Sci. 109 (1993) 181–224.
[40] I.T. Jolliffe, Principal Component Analysis, Springer, New York, 1986.
[41] T. Hastie, W. Stuetzle, Principal curves, J. Am. Stat. Assoc. 84 (406) (1989) 502–516.
[42] B. Kégl, A. Krzyzak, T. Linder, K. Zeger, Learning and design of principal curves, IEEE Trans. Pattern Anal. Mach. Intell. 22 (2) (2000) 281–297.
[43] D. Kendall, A survey of the statistical theory of shape (with comments and rejoinder), Stat. Sci. 4 (1989) 87–120.
[44] I.P. Kornfeld, S.V. Fomin, Ya.G. Sinai, Ergodic Theory, Springer, New York, 1982 (Engl. transl).
[45] V. Volterra, Lecons sur la Théorie Mathematique de la Lutte Pour la vie, Gauthier-Villars, Paris, 1931.
[46] G.F. Gause, The Struggle for Existence, Williams and Wilkins, Baltimore, 1934 Online: http://www.ggause.com/Contgau.htm.
[47] O. Diekmann, A beginner's guide to adaptive dynamics, in: Mathematical Modelling of Population Dynamics, Banach Center Publications, vol. 63, Institute of Mathematics Polish Academy of Sciences, Warszawa, 2004, pp. 47–86.
[48] A.N. Gorban, Equilibrium Encircling, Equations of Chemical Kinetics and Their Thermodynamic Analysis, Nauka, Novosibirsk, 1984.
[49] M. Gromov, A dynamical model for synchronisation and for inheritance in micro-evolution: a survey of papers of A. Gorban, The talk given in the IHES Seminar, ''Initiation to functional genomics: biological, mathematical and algorithmical aspects'', Institut Henri Poincaré, November 16, 2000.
[50] V.E. Zakharov, V.S. L'vov, S.S. Starobinets, Turbulence of spin-waves beyond threshold of their parametric-excitation, Uspekhi Fizicheskikh Nauk 114 (4) (1974) 609–654
English translation Sov. Phys. - Usp. 17(6) (1975) 896–919.
[51] V.E. Zakharov, V.S. L'vov, G.E. Falkovich, Kolmogorov spectra of turbulence, in: Wave Turbulence, vol. 1, Springer, Berlin, 1992.
[52] V.S. L'vov, Wave Turbulence under Parametric Excitation: Applications to Magnets, Springer, Berlin, Heidelberg, 1994.
[53] A.N. Gorban, Systems with inheritance: dynamics of distributions with conservation of support, natural selection and finite-dimensional asymptotics, E-print arXiv cond-mat/0405451.
[54] A.N. Gorban, Order–disorder separation: geometric revision, E-print arXiv cond-mat/0507644.
[55] S. Goldstein, J.L. Lebowitz, R. Tumulka, N. Zanghi, Canonical typicality, Phys. Rev. Lett. 96 (2006) 050403 E-print arXiv cond-mat/0511091.
[56] D. Costantini, Symmetry and indistinguishability of classical particles, Phys. Lett. A 123 (1987) 433–436.
[57] A.S. Sant'Anna, A.M.S. Santos, Quasi-set-theoretical foundations of statistical mechanics: a research program, Found. Phys. 30 (1) (2000) 101–120.
[58] L.A. Blumenfeld, A.Y. Grosberg, Gibbs paradox and the notion of construction in thermodynamics and statistic physics, Biofizika 40 (3) (1995) 660–667.
[59] J.F. Nagle, Regarding the entropy of distinguishable particles, J. Stat. Phys. 117 (5–6) (2004) 1047–1062.
[60] B.L. Gordon, Maxwell-Boltzmann statistics and the metaphysics of modality, Synthese 133 (3) (2002) 393–417.
[61] V. Pestov, On the geometry of similarity search: dimensionality curse and concentration of measure, Inform. Process. Lett. 73 (1–2) (2000) 47–51.