CUBO A Mathematical Journal
Vol. 9, Nº 2, (15 - 36). August 2007.

Algorithmic Complexity and Statistical Mechanics[1]

Vladimir V'yugin
Institute for Information Transmission Problems, Russian Academy of Sciences, Bol'shoi Karetnyi per. 19, Moscow GSP-4, 127994, Russia.
vyugin@iitp.ru

Victor Maslov
Moscow State University, Physical Faculty, Vorobyevy Gory, Moscow 119899 Russia.
v.p.maslov@mail.ru

ABSTRACT

We apply the algorithmic complexity theory to statistical mechanics; in particular, we consider the maximum entropy principle and the entropy concentration theorem for non-ordered data in a non-probabilistic setting. The main goal of this paper is to deduce asymptotic relations for the frequencies of energy levels in a non-ordered collection $\omega^N = [\omega_1, \dots, \omega_N]$ from the assumption of maximality of the Kolmogorov complexity $\mathrm{K}(\omega^N)$ given a constraint $\sum_{i=1}^{N} f(\omega_i) = NE$, where E is a number and f is a numerical function; $f(\omega_i)$ is an energy level.

We also consider a combinatorial model of the securities market and give some applications of the entropy concentration theorem to finance.

RESUMEN

Aplicamos la teoría de complexidad algorítmica para mecánica estadística, en particular, consideramos el principio de entropía máxima y el teorema de concentración entrópica para datos no ordenados en un contexto no probabilístico. El

[1]This work was partially supported by Russian foundation for fundamental research: 06-01-00122. A part of the paper was presented in the conference paper [13].

primer objetivo de este artículo es deducir relaciones asintóticas para las frecuencias de niveles de energía en una colección no ordenada $\omega^N = [\omega_1, \dots, \omega_N]$ con la suposición de maximilidad de la complejidad de Kolmogorov $K(\omega^N)$ dado una coacción $\sum_{i=1}^{N} f(\omega_i) = NE$, donde E es un número y f es una función numérica; $f(\omega_i)$ es un nivel de energía.

También consideramos un modelo combinatorial de mercados de seguridad y damos aplicaciones del teorema de concentración entrópica a financiar.

Key words and phrases: *Algorithmic complexity; Algorithmic information theory; Statistical mechanics; Maximum entropy principle; Jaynes' entropy concentration theorem; Distribution of investments*

Math. Subj. Class.: *68Q30; 82B05; 91B28; 91B50*

1 Introduction

Generally, main notions and results of statistical mechanics are presented in the probabilistic framework. In this paper we pose some ideas and theorems on this subject in a non-probabilistic form for non-ordered data. We use algorithmic complexity theory [12] as the main tool to obtain corresponding results. We obtain the algorithmic versions of the Jaynes' *maximum entropy principle* and of the *entropy concentration theorem*.

Jaynes' maximum entropy principle is well-known as the principle of inductive inference and probabilistic forecasting; it is used in many applications for the construction of optimal probability distributions when some a priori constrains for the mathematical expectation and other moments are given (Jaynes [5], Cover and Thomas [2], Section 11). Extreme relations between the cost of the information transmission and the capacity of the channel were considered in [16] (Chapter 3). This principle originate from statistical physics; it is used for computation the numerical characteristics of ideal gases in the equilibrium state (Landau and Lifshitz [11]).

Let $f(a)$ be a function taking numerical values at all letters of an alphabet $B = \{a_1, \dots a_M\}$. For each collection of letters $\omega_1, \dots, \omega_N$ from the alphabet B we consider the sum $\sum_{i=1}^{N} f(\omega_i)$. The value $f(\omega_i)$ can have various physical or economic meanings. It may describe the cost of the element ω_i of a message or a loss under the occurrence of the event ω_i. In thermodynamics, $f(\omega_i)$ is the energy of a particle or volume element in the state ω_i.

In contrast to [2], [5], [16], we consider non-ordered collections or bags; to be more concise, we consider a variant of the well known in statistical physics Bose - Einstein model [11], [15] for a system of N indistinguishable particles of n types. This type of data is also typical for finance, where non-ordered and indistinguishable collections of items (like shares of stocks) are considered.

In this work, we do not assume the existence of any probabilistic mechanism generating elements of a collection $\omega_1, \dots, \omega_N$. Instead of this, we consider combinatorial models for describing possible collections of outcomes $\omega_1, \dots, \omega_N$ typical for statistical mechanics; more precise, we assume that the collection of outcomes under consideration are "chaotic" or "generic" elements in some simple sets. The notions of chaotic and simple objects are introduced using the algorithmic complexity (algorithmic entropy) introduced by Kolmogorov in [6].

The entropy concentration theorem is considered as some justification of the maximum entropy principle [5]. The main goal of Sections 2, 3 and 4 is to present results closely connected with this theorem; we deduce asymptotic relations for the frequencies of the energy levels in an non-ordered collection $\omega^N = [\omega_1, \dots, \omega_N]$ from the assumption of maximality of the Kolmogorov complexity $\mathrm{K}(\omega^N)$ given the constraint $\sum\limits_{i=1}^{N} f(\omega_i) = EN$.

In Section 5 we present some applications of the entropy concentration theorem; we consider a simple combinatorial model of the securities market and the problem of optimal distribution of investments among different securities (stocks, bonds, etc.). We use the notions of algorithmic complexity and entropy of the market as basic notions of our model. We show that the value of a sufficiently complex portfolio of securities changes in the same way as the mean value of the rate of return of all market securities.

2 Preliminaries

We refer readers for details of the theory of Kolmogorov complexity and algorithmic randomness to [12]. In this section we briefly introduce some definitions used in the following.

The *Kolmogorov complexity* is defined for arbitrary constructive (finite) objects. A set of all words over a finite alphabet is a typical example of the set of constructive objects. For any set of constructive objects we can effectively identify its elements and finite binary sequences. The definition of Kolmogorov complexity is based on the theory of algorithms. Algorithms define computable functions transforming constructive objects. Let $B(p, y)$ be an arbitrary computable function of two arguments, where p is a finite binary word, and y is a word in some alphabet. We consider the function $B(p, y)$ as a method of decoding of constructive objects, where p is a code of an object under a condition y. We suppose also that the method of decoding is prefix-free: if $B(p, y)$ and $B(p', y)$ are defined then $p \not\subset p'$ and $p' \not\subset p$, where $\subset$ is the relation of words extension. The measure of complexity (with respect to B) of a constructive object x given a constructive object y is defined

$$K_B(x|y) = \min\{l(p) \mid B(p, y) = x\},$$

where $l(p)$ is the length of the binary word p (we set $\min \emptyset = \infty$). A decoding method $B(p, y)$ is called optimal if for any other method of decoding $B'(p, y)$ the inequality $K_B(x|y) \leq K_{B'}(x|y) + O(1)$ holds, where the constant $O(1)$ does not depend on x and y (but does depend on the function B') [2] .. An optimal method of decoding exists [12]. Any two optimal decoding methods determine measures of complexity differing by a constant. We fix one such optimal decoding method, denote the corresponding measure of complexity by $\mathrm{K}(x|y)$, and call it the (conditional) Kolmogorov complexity of x with respect to y. The unconditional complexity of the word x is defined as $\mathrm{K}(x) = \mathrm{K}(x|\Lambda)$, where Λ is the empty sequence.

We will use the following relations which hold for the prefix complexity of positive integer numbers n (see the book [12]); where any positive integer number is identified with its binary representation. For any $\epsilon > 0$

$$\mathrm{K}(n) \leq \log n + (1+\epsilon)\log\log n + O(1). \tag{1}$$

Besides,

$$\mathrm{K}(n) \geq \log n + \log\log n \tag{2}$$

for infinitely many n.

To specify an element x of a finite set D, it suffices to know the set D, say given as a list of its elements, and the index of x in this list; the binary sequence representing this index has length [3] $\leq \lceil \log |D| \rceil$. This encoding is prefix-free. Then we have

$$\mathrm{K}(x|D) \leq \log |D| + O(1).$$

Moreover, for any $c > 0$, the number of all $x \in D$ such that

$$\mathrm{K}(x|D) < \log |D| - c, \tag{3}$$

is at most $2^{-c}|D|$; i.e., most elements of the set D are of conditional Kolmogorov complexity close to its maximal value. Kolmogorov [8], [9] defined the notion of the *deficiency of algorithmic randomness* of an element x of a finite set D of constructing objects

$$d(x|D) = \log |D| - \mathrm{K}(x|D).$$

Denote by $\mathrm{Rand}_m(D) = \{x \in D : d(x|D) \leq m\}$ the set of all *m-random (chaotic)* elements of D. It holds $|\mathrm{Rand}_m(D)| \geq (1 - 2^{-m})|D|$.

Let $\mathcal{X}$ be a finite set of constructive objects and $\Xi(\alpha)$ be a computable function from $\mathcal{X}$ to some set $\mathcal{Y}$ of constructive objects. We refer to $\Xi(\alpha)$ as to a *sufficient statistics* (this notion is studied in [4], [12]). We will identify the value $\Xi(\alpha)$ of sufficient statistics and its whole prototype $\Xi^{-1}(\Xi(\alpha))$. So, we identify the sufficient statistics Ξ and the corresponding *partition* of the set $\mathcal{X}$. We refer to $\Xi(\alpha)$ as to a *macrostate* generated by a *microstate* α.

[2] The expression $f(x_1, \ldots, x_n) \leq g(x_1, \ldots, x_n) + O(1)$ means that there exists a constant c such that the inequality $f(x_1, \ldots, x_n) \leq g(x_1, \ldots, x_n) + c$ holds for all $x_1, \ldots, x_n$. The expression $f(x_1, \ldots, x_n) = g(x_1, \ldots, x_n) + O(1)$ means that $f(x_1, \ldots, x_n) \leq g(x_1, \ldots, x_n) + O(1)$ and $g(x_1, \ldots, x_n) \leq f(x_1, \ldots, x_n) + O(1)$. In the following the expression $F(N) = O(G(N))$ means that a constant c exists (not depending on N) such that $|F(N)| \leq cG(N)$ for all N.

[3] In the following $\log r$ denotes the logarithm of r on the base 2, $\lceil r \rceil$ is the minimal integer number $\geq r$, $|D|$ is the cardinality of the set D.

We will use the following Levin - Gács theorem which is valid for the prefix Kolmogorov complexity:

$$\mathrm{K}(x,y) = \mathrm{K}(y) + \mathrm{K}(x|y, \mathrm{K}(y)) + O(1), \tag{4}$$

where x and y are arbitrary constructive objects [12].

Since $\mathrm{K}(\alpha, \Xi(\alpha)) = \mathrm{K}(\alpha) + O(1)$, we obtain from (4) a natural representation of the complexity of a microstate through its conditional complexity with respect to its macrostate and the complexity of the macrostate itself

$$\mathrm{K}(\alpha) = \mathrm{K}(\alpha|\Xi(\alpha), \mathrm{K}(\Xi(\alpha))) + \mathrm{K}(\Xi(\alpha)) + O(1). \tag{5}$$

The deficiency of randomness of a microstate α with respect to a sufficient statistics $\Xi(\alpha)$ (or to the corresponding partition Ξ) is defined

$$d_{\Xi}(\alpha) = \log|\Xi(\alpha)| - \mathrm{K}(\alpha|\Xi(\alpha), \mathrm{K}(\Xi(\alpha))). \tag{6}$$

By definition for any $\alpha \in \mathcal{X}$ it holds $\mathrm{K}(\alpha|\Xi(\alpha), \mathrm{K}(\Xi(\alpha))) \leq \log|\Xi(\alpha)| + O(1)$. We have $d_{\Xi}(\alpha) \geq -c$ for all $\alpha \in \mathcal{X}$, where $c \geq 0$ is a constant. Moreover, for any $m \geq 0$ the number of all $\alpha \in \mathcal{X}$ such that $d_{\Xi}(\alpha) > m$ is not greater than $2^{-m}|\mathcal{X}|$.

By (6) the following representation of the complexity of an element $\alpha \in \mathcal{X}$ is valid

$$\mathrm{K}(\alpha) = \log|\Xi(\alpha)| + \mathrm{K}(\Xi(\alpha)) - d_{\Xi}(\alpha) + O(1). \tag{7}$$

Let $B = \{a_1, \dots, a_M\}$ be some finite alphabet, $f(a)$ be a function on B, and let $\lambda_1, \dots, \lambda_n$ be all its different values (energy levels); we suppose that these values are computable real numbers. We also suppose that $3 \leq n \leq M$. Define $G_j = \{a : f(a) = \lambda_j\}$, $g_j = |G_j|$, $j = 1, \dots n$. It holds $\sum_{j=1}^{n} g_j = M$.

Let us consider the set $\mathcal{B}^N$ of all non-ordered collections (bags or multi-sets) $\omega^N = [\omega_1, \dots, \omega_N]$ of size N; this set can be obtained from the set B^N by factorization with respect to the group of all permutations of the elements $\omega_1, \dots, \omega_N$. Any multiset ω^N can be identified with the constructive object - M-tuple

$$\overline{n} = (n_1, \dots, n_M), \tag{8}$$

where $n_i = |\{j : \omega_j = a_i\}|$ is the multiplicity of the letter a_i in ω^N, $i = 1, \dots, M$. The size of this multi-set ω^N is equal to the sum of all multiplicities $N = \sum_{i=1}^{M} n_i$. Therefore, the notion of complexity $\mathrm{K}(\omega^N) = \mathrm{K}(\overline{n})$ of any non-ordered collection ω^N is well defined.

Let Π_N^n be a set (simplex) of all n-tuples of nonnegative integer numbers $\overline{N} = (N_1, \dots, N_n)$ such that $\sum_{i=1}^{n} N_i = N$. A sufficient statistics $\Xi(\omega^N)$ on $\mathcal{B}^N$ with the range in Π_N^n is defined as follows. Put $\Xi(\omega^N) = \overline{N}$, where

$$N_i = N_i(\omega^N) = |\{j : 1 \leq j \leq N, f(\omega_j) = \lambda_i\}| \tag{9}$$

for $i = 1, \dots n$. This means that the element $\Xi_{\overline{N}}$ defined by $\overline{N} = (N_1, \dots, N_n)$ of the corresponding partition of the set $\mathcal{B}^N$ consists of all non-ordered collections

$\omega^N = [\omega_1, \ldots, \omega_N]$ satisfying (9). In other words, this element consists of M-tuples $\overline{n} = (n_1, \ldots, n_M)$ such that

$$\sum_{a_j \in G_i} n_j = N_i$$

for $i = 1, \ldots, n$. By definition

$$\Xi_{\overline{N}} = \Pi^{g_1}_{N_1} \otimes \cdots \otimes \Pi^{g_n}_{N_n}.$$

Therefore, we have

$$|\Xi_{\overline{N}}| = \binom{g_1 + N_1 - 1}{g_1 - 1} \cdots \binom{g_n + N_n - 1}{g_n - 1}. \tag{10}$$

The number $\rho = N/M$ is called *density*. Let $p_i = g_i/M$, $i = 1, \ldots, n$. We will consider asymptotic relations for $N \to \infty$, $M \to \infty$ such that $N/M = \rho > 0$. We suppose that the numbers ρ and $p_1, \ldots, p_n$ are positive constants.

Let $\overline{n} \in \Xi_{\overline{N}}$, $\nu_i = \nu_i(\overline{n}) = N_i/N$, $i = 1, \ldots, n$. By Stirling formula and by (6), (7) we obtain

$$K(\overline{n}) = NH(\overline{\nu}) -$$

$$-\sum_{i=1}^{n} \frac{1}{2} \log \frac{N_i(g_i - 1)}{N_i + g_i - 1} + K(\Xi(\overline{n})) - d_{\Xi}(\overline{n}) + O(1), \tag{11}$$

where the leading (linear by N) member of this representation is defined by

$$H(\overline{\nu}) = \tag{12}$$

$$= \sum_{i=1}^{n} ((\nu_i + p_i\rho^{-1}) \log(\nu_i + p_i\rho^{-1}) - \nu_i \log \nu_i - p_i\rho^{-1} \log(p_i\rho^{-1}))$$

and is called the *Bose entropy* of the frequency distribution $\overline{\nu} = (\nu_1, \ldots, \nu_n)$. The second subtracted term of (11) is approximated by $\frac{1}{2} n \log N + O(1)$ for $N \to \infty$, where the term $O(1)$ depends on p_i, $i = 1, \ldots, n$, and on ρ.

Let the mean value E of the function f be given (we suppose that E is a computable real number), and let $\alpha(N)$ be a computable nondecreasing function such that $\alpha(N) = o(N)$. We suppose that $\alpha(N)$ is unbounded; note that, the results of this paper also hold for the case, where $\alpha(N)$ is a sufficiently large constant. Denote by $C_N(\overline{\lambda})$ the set of all microstates $\overline{n}$ of size N satisfying

$$|\sum_{i=1}^{n} \lambda_i N_i - EN| \leq \alpha(N), \tag{13}$$

where $\overline{\lambda} = (\lambda_1, \ldots, \lambda_n)$ and the numbers N_i, $i = 1, \ldots, n$, are defined by $\overline{n}$ (in other words, $\overline{n} \in \Xi_{\overline{N}}$).

Let us suppose that the maximum H_{max} of the entropy (12) given constrains

$$\sum_{i=1}^{n} \nu_i \lambda_i = E, \quad \sum_{i=1}^{n} \nu_i = 1 \tag{14}$$

is attained for $\tilde{\nu} = (\tilde{\nu}_1, \dots, \tilde{\nu}_n)$, where $i = 1, \dots, n$. These values will be computed in the proof of Theorem 2.1 below (see also [11]).

Jaynes' entropy concentration theorem says that for any sufficiently small $\epsilon > 0$ the portion of all (ordered) sequences or words $\omega^N = \omega_1 \dots \omega_N$ of the alphabet B^N of the length N satisfying (14) and such that $|\nu_i(\omega^N) - \tilde{\nu}_i| \geq \epsilon$ for some $1 \leq i \leq n$, is not greater than e^{-cN}, where $\tilde{\nu} = (\tilde{\nu}_1, \dots, \tilde{\nu}_n)$ is the point maximizing the Shannon entropy $H(\overline{\nu}) = \sum_{i=1}^{n} -\nu_i \log \nu_i$ given constrains (14) and c is a constant depending of ϵ.

An analogue of this theorem holds for the set of all non-ordered collections $\mathcal{B}^N$.

Theorem 2.1 *1) For any $\epsilon > 0$ the portion of all microstates $\overline{n} \in C_N(\overline{\lambda})$ such that*

$$|\nu_i - \tilde{\nu}_i| \geq \epsilon \tag{15}$$

for some i is not greater than $2^{-c\epsilon^2 N}$ for all sufficiently large N, where $\nu_i = \nu_i(\overline{n})$, the numbers $\tilde{\nu}_i$ are defined by (19) (below) for $i = 1, \dots, n$, and c is a positive constant.

A version of this theorem in the algorithmic complexity language looks as follows:

2) Let $\sigma(N)$ be a non decreasing nonnegative computable function. Then for all microstates $\overline{n} \in C_N(\overline{\lambda})$ except of their portion $2^{-\sigma(N)+O(n \log N)}$ the following property holds

$$NH_{max} - \sigma(N) \leq \mathrm{K}(\overline{n}) \leq NH_{max} + O(n \log N), \tag{16}$$

as $N \to \infty$.

Moreover, for any microstate $\overline{n} \in C_N(\overline{\lambda})$, condition (16) implies the inequality

$$|\nu_i(\overline{n}) - \tilde{\nu}_i| \leq \sqrt{\frac{c(|\beta|\alpha(N) + \sigma(N) + n \log N)}{N}} \tag{17}$$

for all i, where c is a positive constant, the constant β will be defined in the proof of this theorem, $\nu_i(\overline{n}) = N_i/N$ for $i = 1, \dots, n$, and the values of N_i are generated by the microstate $\overline{n}$.

Proof. We remark that the total number of all microstates generating the macrostate maximizing entropy given constraints is exponentially larger than the cardinality of the remaining part of the set of all microstates satisfying (15) for some i.

Recall that, the cardinality of an arbitrary element $\Xi_{\overline{N}}$ of the partition defined by $\overline{N} = (N_1, \dots, N_n)$ is equal to

$$|\Xi_{\overline{N}}| = 2^{NH(\overline{\nu}) + O(n \log N)},$$

where $\nu_i = N_i/N$, $i = 1,\dots,n$. The function H (defined by (12)) is conceive and conditions (14) are linear. Then by Kuhn - Tucker theorem [10] H has a unique maximum H_{max} given constrains (14).

To compute the maximum of (12) given constrains (14), we consider the Lagrange functional

$$L = H(\overline{\nu}) + \beta\sum_{i=1}^{n}\lambda_i\nu_i + \mu(\sum_{i=1}^{n}\nu_i - 1).$$

The necessary conditions for the maximum of (12) are as follows: $\partial L/\partial\nu_i = 0$, $\partial L/\partial\beta = 0$, and $\partial L/\partial\mu = 0$; in particular,

$$\frac{\partial H}{\partial\nu_i} + \beta\lambda_i + \mu = 0, \tag{18}$$

where $i = 1,\dots,n$.

The maximum value H_{max} of entropy is attained for

$$\tilde{\nu}_i = \frac{\rho^{-1}p_i}{2^{-\beta\lambda_i-\mu} - 1} \tag{19}$$

where $i = 1,\dots,n$ and parameters β and μ are determined by (14) (see [11], [17]). Note that the parameters β and μ can be positive or negative, depending on the values of λ_i.

Let $\overline{N} = (N_1,\dots,N_n)$ be an arbitrary macrostate, $\nu_i = N_i/N$ and $\tilde{\nu} = (\tilde{\nu}_1,\dots,\tilde{\nu}_n)$, $\tilde{\nu}_i = \tilde{N}_i/N$. The maximum $H_{max} = H(\tilde{\nu})$ of the entropy corresponds to some macrostate $\tilde{N}$. We have

$$\frac{|\Xi_{\overline{N}}|}{|\Xi_{\tilde{N}}|} = 2^{N(H(\overline{\nu})-H(\tilde{\nu}))+O(n\log N)}. \tag{20}$$

The variation of the entropy at the maximum point given constrains (14) has the form

$$\delta H(\tilde{\nu}) = H(\overline{\nu}) - H(\tilde{\nu}) =$$

$$= \sum_{i=1}^{n}\frac{\partial H}{\partial\nu_i}(\tilde{\nu})(\tilde{\nu})(\nu_i - \tilde{\nu}_i) + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2 H}{\partial\nu_i^2}(\tilde{\nu})(\nu_i - \tilde{\nu}_i)^2 + O((\nu_i - \tilde{\nu}_i)^3)\right),$$

where

$$\frac{\partial^2 H}{\partial\nu_i^2}(\tilde{\nu}) = \frac{-\log e}{\tilde{\nu}_i(1 + \rho\tilde{\nu}_i/p_i)}.$$

Since $\frac{\partial H}{\partial\nu_i}(\tilde{\nu}) = -\beta\lambda_i - \mu$, by (18) and (14) we have

$$\left|\sum_{i=1}^{n}\frac{\partial H}{\partial\nu_i}(\tilde{\nu})(\nu_i - \tilde{\nu}_i)\right| \le |\beta\alpha(N)/N| = o(1)$$

as $N \to \infty$. Hence,

$$\delta H(\tilde{\nu}) = \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2 H}{\partial \nu_i^2}(\tilde{\nu})(\nu_i - \tilde{\nu}_i)^2 + O((\nu_i - \tilde{\nu}_i)^3)\right) + o(1) \tag{21}$$

as $N \to \infty$.

Assume that $|\nu_i - \tilde{\nu}_i| \geq \epsilon$ for some $1 \leq i \leq n$, where ϵ is a sufficiently small positive number. Then by (21) and by the general properties of the function (12) we have for all sufficiently large N

$$H(\overline{\nu}) \leq H(\tilde{\nu}) - (1-\gamma)c_i\epsilon^2, \tag{22}$$

where γ is some positive number such that $0 < \gamma < 1$, and $c_i = -\frac{1}{2}\frac{\partial^2 H}{\partial \nu_i^2}(\tilde{\nu})$.

The total number of all macrostates is polynomial in N. This implies that the total number of all microstates satisfying this inequality decreases exponentially with N. In other words, the cardinality of the union of all sets $\Xi_{\overline{N}}$ such that $|\nu_i - \tilde{\nu}_i| > \epsilon$ for at least one $i = 1, \dots, n$ is not grater than $2^{-c\epsilon^2 N}|\Xi_{\tilde{N}}| < 2^{-c\epsilon^2 N}|\mathcal{C}_N(\overline{\lambda})|$, where c is a positive constant. From this part 1) follows.

The inequalities (16) follow from (3). To prove (17) it suffices to use inequalities

$$NH(\tilde{\nu}) - \sigma(N) \leq \mathrm{K}(\overline{n}) = N(H(\tilde{\nu}) + \delta H(\tilde{\nu})) +$$
$$+O(n\log N) - d_{\Xi(\overline{n})} \leq NH(\tilde{\nu}) - d_{\Xi}(\overline{n}) + |\beta|\alpha(N) +$$
$$+\frac{1}{2}N\sum_{i=1}^{n}\left(\frac{\partial^2 H}{\partial \nu_i^2}(\tilde{\nu})(\nu_i^N - \tilde{\nu}_i)^2 + O((\nu_i^N - \tilde{\nu}_i)^3)\right) + O(n\log N),$$

and $\frac{\partial^2 H}{\partial \nu_i^2}(\tilde{\nu}) < 0$ for all i. ∎

3 Asymptotic relations for frequencies of energy levels

The following theorem is the starting point for the following theorems presenting more tight asymptotic bounds for the frequencies. These bounds are presented in the "worst" case.

Theorem 3.1 *It holds*

$$\max_{\overline{n}\in\mathcal{C}_N(\overline{\lambda})} \mathrm{K}(\overline{n}) = NH_{max} - \log N + \mathrm{K}(N) + O(1). \tag{23}$$

The proof of this theorem is given in Section 4.

The bound (23) is non-computable. Two computable lower bounds, the first one holds for "most" N, the second one, trivial, holds for all N, are presented in Corollary 3.2; they also will be used in the definitions (25) and (27) below.

We will consider limits by the filter of the base

$$B_L^m = \{N : N \geq L, \quad \mathrm{K}(N) \geq \log N - m\},$$

where $m, L = 0, 1, \ldots$. It is easy to see that $B_{L+1}^m \subseteq B_L^m$ and

$$\lim_{m\to\infty} \liminf_{N\to\infty} \frac{|B_L^m \cap \{L, \ldots, L+N-1\}|}{N} = 1$$

for all m and L.

Indeed, taking into account that the number of all programs of length $< \log N' - m$, where $N' \leq L + N - 1$, is not greater than $2^{-m}(L+N-1)$, we obtain

$$\frac{|B_L^m \cap \{L, \ldots, L+N-1\}|}{N} \geq 1 - 2^{-m+1}$$

for all sufficiently large N.

Corollary 3.2 *For any $\epsilon > 0$, $m > 0$ and $N \in B_0^m$*

$$NH_{max} - m - O(1) \leq \max_{\overline{n}\in\mathcal{C}_N(\overline{\lambda})} \mathrm{K}(\overline{n}) \leq NH_{max} + \\ +(1+\epsilon)\log\log N + O(1). \tag{24}$$

The upper bound in (24) is valid for all N. The following trivial lower bound is valid for all N

$$\max_{\overline{n}\in\mathcal{C}_N(\overline{\lambda})} \mathrm{K}(\overline{n}) \geq NH_{max} - \log N - O(1).$$

The bound (24) can be obtained by applying (1) to (23).

Let σ be a nonnegative number. Let us define the set of all *σ-random microstates* locating in the layer (13)

$$\mathrm{Rand}(E, \sigma, N) = \{\overline{n} : \overline{n} \in \mathcal{C}_N(\overline{\lambda}), \mathrm{K}(\overline{n}) \geq NH_{max} - \sigma\}. \tag{25}$$

By Corollary 3.2 for all m and $N \in B_0^m$ the set $\mathrm{Rand}(E, \sigma, N)$ is nonempty when $\sigma \geq m + O(1)$; this set contains all microstates $\overline{n} \in \mathcal{C}_N(\overline{\lambda})$ maximizing the complexity on $\mathcal{C}_N(\overline{\lambda})$. In the following we suppose that $\sigma = \sigma(N) = o(\log\log N)$.

Theorem 3.3 *For any m and for any nondecreasing numerical function $\sigma(N)$ such that $\sigma(N) \geq m + O(1)$ and $\sigma(N) = o(\log\log N)$ as $N \to \infty$, the following asymptotic relation holds*

$$\limsup_{B_N^m} \sup_{\overline{n}\in\mathrm{Rand}(E,\sigma(N),N)} \sum_{i=1}^{n} \left(\frac{N_i - N\tilde{\nu}_i}{\sqrt{N\tilde{\nu}_i(1+\rho\tilde{\nu}_i/p_i)\log\log N}} \right)^2 = 1, \tag{26}$$

where the numbers $N_i = N_i(\overline{n})$ are defined by (9) and the values $\tilde{\nu}_i$ are defined by (19) for $i = 1, \ldots, n$.

The proof of this theorem is given in Section 4.

Let us consider the second computable lower bound for the maximum of the complexity given constrains (14). Let us define

$$\mathrm{Rand}'(E,\sigma,N) = \{\overline{n} : \overline{n} \in \mathcal{C}_N(\overline{\lambda}), \mathrm{K}(\overline{n}) \geq NH_{max} - \log N - \sigma\}. \tag{27}$$

The following theorem holds.

Theorem 3.4 *For any nonnegative number σ*

$$\limsup_{N\to\infty} \sup_{\overline{n}\in\mathrm{Rand}'(E,\sigma,N)} \sum_{i=1}^{n} \left(\frac{N_i - N\tilde{\nu}_i}{\sqrt{N\tilde{\nu}_i(1+\rho\tilde{\nu}_i/p_i)\mathrm{K}(N)}} \right)^2 = 1,$$

where the notations are the same as in Theorem 3.3.

See the sketch of the proof of this theorem in Section 4.

The following corollary asserts that the maximum of the complexity on $\mathcal{C}_N(\overline{\lambda})$ is attained at a microstate $\overline{n}$ "random" in the set $\Xi_{\overline{N}}$ representing the corresponding macrostate $\overline{N} = (N_1, \dots, N_n)$; this macrostate is also "random" (has a general position) in the set defined by the inequalities (28) below.

For any $m > 0$ and N define

$$\mathrm{Max}(E,m,N) = \{\overline{n} : \mathrm{K}(\overline{n}) \geq \max_{\overline{k}\in\mathcal{C}_N(\overline{\lambda})} \mathrm{K}(\overline{k}) - m\}.$$

Corollary 3.5 *(from Theorem 3.1). For any $m > 0$ and $\overline{n} \in \mathrm{Max}(E,m,N)$, the relations $d_\Xi(\overline{n}) = O(1)$ and $\mathrm{K}(N_1,\dots,N_n|N,\mathrm{K}(N)) = \frac{1}{2}(n-2)\log N + O(1)$ hold as $N \to \infty$. Also, for some $c \geq 1$*

$$1 \leq \limsup_{N\to\infty} \sup_{\overline{n}\in\mathrm{Max}(E,m,N)} \sum_{i=1}^{n} \left(\frac{N_i - N\tilde{\nu}_i}{\sqrt{N\tilde{\nu}_i(1+\rho\tilde{\nu}_i/p_i)}} \right)^2 \leq c, \tag{28}$$

where we use the same notations as in Theorem 3.3.

The proof of this corollary see in Section 4.

4 Proofs of Theorems 3.1, 3.3 and 3.4

Proof of Theorem 3.3. For each N choose a microstate $\overline{n} \in \mathcal{C}_N(\overline{\lambda})$ such that

$$\mathrm{K}(\overline{n}) \geq NH_{max} - o(\log\log N). \tag{29}$$

By (11) and (21) we have

$$\mathrm{K}(\overline{n}) = NH_{max} - \sum_{i=1}^{n} \left(\frac{(N_i - N\tilde{\nu}_i)^2}{N\tilde{\nu}_i(1+\rho\tilde{\nu}_i/p_i)} + o(N(\delta\tilde{\nu}_i)^2) \right) -$$

$$-\frac{1}{2}n\log N - d_\Xi(\overline{n}) + \mathrm{K}(N_1,\dots,N_n|N,\mathrm{K}(N)) + \mathrm{K}(N) + O(1). \tag{30}$$

Let $\epsilon > 0$ be a sufficiently small number. By (1)

$$\mathrm{K}(N) \le \log N + (1 + \frac{1}{4}\epsilon)\log\log N + O(1). \tag{31}$$

By (29), (30) and (31) we obtain

$$\begin{aligned}(1 - \frac{1}{4}\epsilon)\sum_{i=1}^{n} \frac{(N_i - N\tilde{\nu}_i)^2}{N\tilde{\nu}_i(1 + \rho\tilde{\nu}_i/p_i)} &\le \mathrm{K}(N_1, \dots, N_n|N, \mathrm{K}(N)) - \\ &-\frac{1}{2}(n-2)\log N + (1 + \frac{1}{4}\epsilon)\log\log N + o(\log\log N).\end{aligned} \tag{32}$$

Suppose that for some positive integer number L the following inequality holds

$$\sum_{i=1}^{n} \frac{(N_i - N\tilde{\nu}_i)^2}{N\tilde{\nu}_i(1 + \rho\tilde{\nu}_i/p_i)} \le (1 - \frac{1}{4}\epsilon)^{-1}L + o(\log\log N). \tag{33}$$

Since each number $\tilde{\nu}_i$, $i = 1, \dots, n$, is computable, using relation (33) we can effectively find the corresponding interval of integer numbers with center in $N\tilde{\nu}_i$ containing the number N_i. The length of the program computing N_i is bounded by the length of this interval and by the number $i \le n$. Then we have

$$\mathrm{K}(N_i|L, N) \le \frac{1}{2}\log N + \frac{1}{2}\log L + o(\log\log N). \tag{34}$$

Hence, taking into account constrains (14) on $(N_1, \dots, N_n)$ we obtain the inequality

$$\begin{aligned}\mathrm{K}(N_1, \dots, N_n|L, N) - \frac{1}{2}(n-2)\log N &\le \\ &\le \frac{1}{2}(n-2)\log L + o(\log\log N).\end{aligned}$$

Using the standard inequality $\mathrm{K}(x|L) \ge \mathrm{K}(x) - K(L) - O(1)$, we obtain

$$\begin{aligned}\mathrm{K}(N_1, \dots, N_n|N, \mathrm{K}(N)) - \frac{1}{2}(n-2)\log N &\le \\ &\le \frac{1}{2}n\log L + o(\log\log N).\end{aligned} \tag{35}$$

Since (33) holds for L equal to the maximum of the integer part of the number

$$\mathrm{K}(N_1, \dots, N_n|N, \mathrm{K}(N)) - \frac{1}{2}(n-2)\log N + (1 + \frac{1}{4}\epsilon)\log\log N$$

and the number one, we obtain by (35)

$$L \le \frac{1}{2}n\log L + (1 + \frac{1}{4}\epsilon)\log\log N + o(\log\log N). \tag{36}$$

Then we have

$$\mathrm{K}(N_1,\dots,N_n|N,\mathrm{K}(N)) - \frac{1}{2}(n-2)\log N + (1+\frac{1}{4}\epsilon)\log\log N) \le$$
$$\le (1+\frac{1}{2}\epsilon)\log\log N + o(\log\log N). \quad (37)$$

By (32) we have for any $\epsilon > 0$

$$\sum_{i=1}^{n} \frac{(N_i - N\tilde{\nu}_i)^2}{N\tilde{\nu}_i(1+\rho\tilde{\nu}_i/p_i)} \le (1+\epsilon)\log\log N + o(\log\log N). \quad (38)$$

Hence,

$$\limsup_{B_N^m} \sup_{\overline{n}\in \mathrm{Rand}(E,\sigma(N),N)} \sum_{i=1}^{n} \frac{(N_i - N\tilde{\nu}_i)^2}{N\tilde{\nu}_i(1+\rho\tilde{\nu}_i/p_i)\log\log N} \le (1+\epsilon) \quad (39)$$

for any $\epsilon > 0$. Since ϵ is an arbitrary positive real number, we can replace $1+\epsilon$ in (39) on 1.

Now, let us prove that the lower bound in (26) is also valid. The intersection of the simplex Π_N^n with the layer $\mathcal{C}_N(\overline{\lambda})$ contain the center of the ellipsoid

$$\{(N_1,\dots,N_n) : \sum_{i=1}^{n} \frac{(N_i - N\tilde{\nu}_i)^2}{N\tilde{\nu}_i(1+\rho\tilde{\nu}_i/p_i)\log\log N} \le 1\}, \quad (40)$$

where, temporarily, N_i, $i = 1,\dots,n$, are real variables.

Let ϵ be a sufficiently small positive real number. The volume of the layer

$$1-\epsilon \le \sum_{i=1}^{n} \frac{(N_i - N\tilde{\nu}_i)^2}{N\tilde{\nu}_i(1+\rho\tilde{\nu}_i/p_i)\log\log N} \le 1 - \frac{\epsilon}{2}, \quad (41)$$

is equal to $c(\epsilon)V$, where V is the volume of the whole ellipsoid (40), and the constant $c(\epsilon)$ depends on ϵ, but does not depend on N. The similar equality holds for volume of the intersection of the simplex Π_N^n with the layer $\mathcal{C}_N(\overline{\lambda})$ and with the ellipsoid (40). The similar equality also holds for volume of the intersection of the simplex Π_N^n with the layer $\mathcal{C}_N(\overline{\lambda})$ and with the layer (41). Since the volume of any ellipsoid is proportional to the product of lengths of its semi-axes, the total number of all n-tuples $(N_1,\dots,N_n)$ of positive integer numbers locating in the intersection of the layer (41) with the simplex Π_N^n and with the layer $\mathcal{L}^N$ is proportional to $(N\log\log N)^{\frac{1}{2}(n-2)}$. Choose an n-tuple of positive integer numbers $(N_1,\dots,N_n)$ of the general position locating in this intersection. We have

$$\mathrm{K}(N_1,\dots,N_n|N,\mathrm{K}(N)) = \frac{1}{2}(n-2)\log N + \log c(\epsilon) + o(\log\log N)$$

for this n-tuple. Let $\Xi_{\overline{N}}$ be the corresponding macrostate. Then for any $\overline{n} \in \Xi_{\overline{N}}$ such that $d_{\Xi}(\overline{n}) = O(1)$ the following inequality holds

$$\begin{aligned} \mathrm{K}(\overline{n}) \geq NH_{max} - (1 - \frac{1}{2}\epsilon)\log\log N - \frac{1}{2}n\log N + \\ +\mathrm{K}(N_1, \dots, N_n | N, \mathrm{K}(N)) + K(N) + o(\log\log N) = \\ = NH_{max} + \mathrm{K}(N) - \log N - \log\log N + \frac{1}{2}\epsilon\log\log N + \\ + \log c(\epsilon) + o(\log\log N). \end{aligned} \tag{42}$$

We have for this n-tuple $(N_1, \dots, N_n)$ of general position

$$\sum_{i=1}^{n} \frac{(N_i - N\tilde{\nu}_i)^2}{N\tilde{\nu}_i(1 + \rho\tilde{\nu}_i/p_i)\log\log N} \geq 1 - \epsilon. \tag{43}$$

By (2) the inequality $\mathrm{K}(N) \geq \log N + \log\log N$ holds for infinitely many N.

Hence, for any microstate $\overline{n}$ of sufficiently large size N satisfying (43), the inequality $\mathrm{K}(\overline{n}) \geq NH_{max}$ holds. ∎

Sketch of the proof of Theorem 3.4. The proof of this theorem is similar to the proof of Theorem 3.3, where in the second part for any $\epsilon > 0$ we can take an n-tuple of positive integer numbers $(N_1, \dots, N_n)$ of the general position locating in the intersection of the layer

$$1 - \epsilon \leq \sum_{i=1}^{n} \frac{(N_i - N\tilde{\nu}_i)^2}{N\tilde{\nu}_i(1 + \rho\tilde{\nu}_i/p_i)\mathrm{K}(N)} \leq 1 - \frac{\epsilon}{2} \tag{44}$$

with the simplex Π_N^n and with the layer $\mathcal{C}_N(\overline{\lambda})$; this n-tuple satisfies

$$\mathrm{K}(N_1, \dots, N_n | N, \mathrm{K}(N)) \geq \frac{1}{2}(n-2)\log N + O(1).$$

∎

Proof of Theorem 3.1. Let an n-tuple $(N_1, \dots, N_n)$ has a general position, i.e., it lies in the intersection of the layer

$$1 - \epsilon \leq \sum_{i=1}^{n} \frac{(N_i - N\tilde{\nu}_i)^2}{N\tilde{\nu}_i(1 + \rho\tilde{\nu}_i/p_i)} \leq 1 - \frac{\epsilon}{2} \tag{45}$$

with the simplex Π_N^n and with the layer $\mathcal{C}_N(\overline{\lambda})$, i.e. such that

$$\mathrm{K}(N_1, \dots, N_n | N, \mathrm{K}(N)) \geq \frac{1}{2}(n-2)\log N + O(1).$$

In this case, we have by (22) that for any microstate $\overline{n} \in \Xi_{\overline{N}}$ such that $d_{\Xi}(\overline{n}) = O(1)$ the following inequality holds

$$\max_{\overline{n} \in \mathcal{C}_N(\overline{\lambda})} \mathrm{K}(\overline{n}) \geq NH_{max} - \log N + \mathrm{K}(N) - O(1). \tag{46}$$

The left - hand part of the relation (23) follows from this inequality for all $N \in B_0^m$.

Assume that the maximum of the Kolmogorov complexity given constrains (14) is attained on a microstate $\overline{n}$. As was proved above, in this case the inequality

$$\mathrm{K}(\overline{n}) \geq NH_{max} - \log N + \mathrm{K}(N) + O(1)$$

holds. By the proof of the first part of Theorem 3.3 we obtain

$$\mathrm{K}(N_1, \dots, N_n | N, \mathrm{K}(N)) - \frac{1}{2}(n-2)\log N = O(1). \qquad (47)$$

Applying (30) and taking into account (47) we obtain

$$\mathrm{K}(\overline{n}) \leq NH_{max} - \log N + \mathrm{K}(N) + O(1).$$

The right - hand inequality of (23) follows from this relation. Theorem is proved. ∎

Proof of Corollary 3.5. By

$$\mathrm{K}(\overline{n}) = NH_{max} - \log N + \mathrm{K}(N) + O(1)$$

and by the representation (30) using some ideas of the proof of Theorem 3.3 it is easy to obtain (28) and also (47). ∎

5 Appendix: Applications to finance

5.1 A combinatorial model of the securities market

We consider a simple combinatorial model of securities market similar to that considered in [14]. We apply to this model some ideas of statistical mechanics considered in Section 2.

There are M securities (stocks, bonds, etc.) indexed by positive integer numbers $i = 1, \dots, M$. We assume that the value of investments are measured relative to the total market value. We define a *market unit* as the $1/N$ of the total value of the securities market, where N is a sufficiently large positive integer number. In that follows, the numbers N and M will be of the same order. By the choice of the market unit, if the total value of the market instruments in dollars (rubles, etc.) increases (decreases), then the value of our market unit also increases (decreases) in the same proportion and in the same currency.

We consider the discrete time scale $t = 1, 2, \dots$. For example, a time moment can represent the beginning of a trading period (day, month) at the securities market. The distribution of total value N of the market between all securities at time moment $t-1$ can be represented by a microstate that is a vector $\overline{n}^{t-1} = (n_1^{t-1}, \dots, n_M^{t-1})$, where n_i^{t-1} is the number of market units invested in shares of the ith security, $i = 1, \dots, M$. We assume that n_i^{t-1} are positive integer numbers. The total sum of these market units is $n_1^{t-1} + \dots + n_M^{t-1} = N$. We call the microstate $\overline{n}$ the *market portfolio* and N the total value of the market portfolio.

By the definition of the market unit, the total value of the market portfolio does not change with time, always being equal to N. The vector $\overline{\epsilon}^t = (\epsilon_1^t, \ldots, \epsilon_M^t)$ gives relative *rates of return* for all M securities at moment t. Each ϵ_i^t satisfies the inequality $-1 \le \epsilon_i^t < \infty$ and has the following meaning. If at time moment $t-1$ the total value of all shares of ith security equals n_i^{t-1} market units, then at time moment t this total value becomes equal to $(1+\epsilon_i^t)n_i^{t-1}$ market units. Here $\epsilon_i^t n_i^{t-1}$ is the gain (or loss) at time t from investing n_i^{t-1} market units in the ith security at time $t-1$. Note that ϵ_i^t are real numbers, and thus the numbers $(1+\epsilon_i^t)n_i^{t-1}$ are not necessarily integer. By this reason, the requirement 1) below is formulated with an approximate equality.

Let at time moment $t-1$ the market portfolio is represented according to a vector $\overline{n}^{t-1}$; in particular, $n_1^{t-1} + \cdots + n_M^{t-1} = N$. At the next time moment t, the marked defines a vector of relative returns $\overline{\epsilon}^t$. Let us point out two main properties of our model:

1) Conservation of the total market value: the total value of all securities at time moment $t-1$ "approximately" does not change at the next time moment t, i.e., we have $(1+\epsilon_1^t)n_1^{t-1} + \cdots + (1+\epsilon_M^t)n_M^{t-1} \approx N$. This approximate equality follows from the definition of the market unit. The equality is not exact since our market unit is discrete. Thus, this condition should be

$$n_1^{t-1}\epsilon_1^t + \cdots + n_M^{t-1}\epsilon_M^t = o(N) \tag{48}$$

as $N \to \infty$. The accuracy $o(N)$ is the largest possible for the results of this section to hold. Replacing the equality sign by an approximate relation also allows us possibility to avoid problems caused by the incommensurability of the numbers ϵ_i.

2) The second property follows from the theory of algorithmic complexity. First, let us make condition (48) more precise. Let us fix some non-decreasing unbounded function $\alpha(N)$ such that $\alpha(N) = o(N)$ as $N \to \infty$. Let $\mathcal{D}_N(\overline{\epsilon}^t)$ be a set of all vectors $\overline{n}$ with non-negative integer coordinates satisfying the conditions

$$\begin{aligned} n_1 + \cdots + n_M &= N, \\ |n_1\epsilon_1^t + \cdots + n_M\epsilon_M^t| &\le \alpha(N). \end{aligned} \tag{49}$$

Let m be an arbitrary positive integer number. Then, for $(1-2^{-m})|\mathcal{D}_N(\overline{\epsilon}^t)|$ elements $\overline{n}$ of the set $\mathcal{D}_N(\overline{\epsilon}^t)$ we have the inequality

$$\log|\mathcal{D}_N(\overline{\epsilon}^t)| - m \le \mathrm{K}(\overline{n}|\mathcal{D}_N(\overline{\epsilon}^t)) \le \log|\mathcal{D}_N(\overline{\epsilon}^t)| + c, \tag{50}$$

where $\mathrm{K}(\overline{n}|\mathcal{D}_N(\overline{\epsilon}^t))$ is the conditional Kolmogorov complexity of the market portfolio $\overline{n}$, and c is a constant.

Passing from (50) to an analytical bound can be made using some "integration" of information on the base of some sufficient statistics like it was done in Section 2.

We consider a statistical microstructure of securities and their rates of return. Assume that securities are divided into "integrated" groups having the same return rates. It is natural to form these groups by joining all jth securities with close values of ϵ_j. We assign the same rate of return to all securities from one integrated group. Let there be n such groups, $n \le M$. Let the ith integrated group $\mathcal{G}_i$ contain G_i

securities with the same rate of return λ_i^t. Denote $\overline{\lambda}^t = (\lambda_1^t, \ldots, \lambda_n^t)$. By definition $\sum_{i=1}^{n} G_i = M$. Also denote $p_i = G_i/M$, $i = 1, \ldots, n$, $\overline{p} = (p_1, \ldots, p_n)$, and let $\rho = N/M$ be the "investment density".

We assume that the investment density ρ, the number n of integrated groups, and their densities p_i, $i = 1, \ldots, n$, are constants independent of N. We assume that $p_i > 0$ for all i. These conditions are crucial for defining the integrated groups $\mathcal{G}_i$, since in this case Theorem 2.1 from Section 3 can be applied. In the following, we will study asymptotic relations as $N \to \infty$.

Let $N_i = \sum_{j \in \mathcal{G}_i} n_j$ be the total value of all shares of all securities of the ith integrated group (they have the same rates of return λ_i^t), where $i = 1, \ldots, n$. The corresponding macrostate $\Xi_{\overline{N}}$, where $\overline{N} = (N_1, \ldots, N_n)$, is called the *macro portfolio*. Denote $\nu_i = N_i/N$, $i = 1, \ldots, n$, $\overline{\nu} = (\nu_1, \ldots, \nu_n)$. Requirement (49) is then replaced by the requirement

$$N_1 + \cdots + N_n = N,$$
$$|\sum_{i=1}^{n} \lambda_i^t N_i| \le \alpha(N). \tag{51}$$

Denote by $C_N(\overline{\lambda})$ the set of all portfolios $\overline{n}$ generating the macro portfolio $\overline{N}$.

We consider nontrivial distributions of investments, i.e., we assume that the intersection of the hyperplanes $\sum_{i=1}^{n} \lambda_i^t x_i = 0$ and $\sum_{i=1}^{n} x_i = 1$ (x_i are real numbers) for $x_i \ge 0$ is nonempty, and therefore, for all sufficiently large N, the set (51) contains vectors $(N_1, \ldots, N_n)$ with integer coordinates.

5.2 Complex investment portfolios

By an investment portfolio, we mean any vector of nonnegative integer numbers $(k_1, \ldots, k_{M'})$, where the number k_i is interpreted as the value of all shares of the ith security (the number of market units invested in the ith security), where $i = 1, \ldots, M'$ and M' is the total number of all securities among which the total value of the portfolio $K = \sum_{i=1}^{M'} k_i$ is distributed.

We consider a one-step variant of the market model introduced in Section 5.1. At the beginning of a trading period, an investor distributes the total value of his capital K, measured in market units, in form of a portfolio $(k_1, \ldots, k_{M'})$. At the end of the trading period, this total value changes by $\delta K = \sum_{i=1}^{M'} \epsilon_i k_i$, where ϵ_i is the rate of return for the ith security. The investor's capital change δK can have a large absolute value. This means that the owner of the portfolio can gain profit or incur losses. The total value of the investor portfolio becomes $K^{+1} = K + \delta K$.

First, we consider the case when the investor bets by buying some portfolio of stocks at the beginning of the trading period (usually it is a day, month, or year).

Market presents the vector of relative returns $\bar{\epsilon}$ at the end of this period. After that, the investor calculates his gain or loss. We refer to this case as to the *long position*. In case of the *short position* (or short selling) the investor bets by selling his portfolio at the beginning date and by buying it at the closing date [1]; after that, he also calculates his gain or loss.

Assume that we can define the integrated groups $\mathcal{G}_i$, $i = 1,\dots,n$, of financial instruments satisfying conditions given in Section 5.1. Let $\overline{\lambda}$ be a vector of rates of return for these groups.

Let $\overline{n}$ be a market portfolio and $\overline{N} = (N_1,\dots,N_n)$ be the corresponding macro portfolio. Define $\nu(\overline{n}) = (\nu_1(\overline{n}),\dots,\nu_n(\overline{n}))$, where $\nu_i(\overline{n}) = N_i/N$ and $i = 1,\dots,n$, and $H(\nu(\overline{n}))$ is defined by (12). Let $H_{max} = H(\tilde{\nu})$, where $\tilde{\nu}$ be a vector of frequencies maximizing the entropy given constrains

$$|\sum_{i=1}^{n}\nu_i\lambda_i| \le \alpha(N), \tag{52}$$

$$\sum_{i=1}^{n}\nu_i = 1 \tag{53}$$

and given the vector of return rates $\overline{\lambda}$. We call H_{max} the *market entropy*; this value is relates to the period $[t-1,t]$ of time.

Let $\delta(N)$ be a nonnegative rational-valued function such that $\delta(N) = o(1)$ as $N \to \infty$ and $\liminf\limits_{N\to\infty} \delta(N)/(\log N/N) = \infty$.

Theorem 2.1 (where $E = 0$) asserts that the vast majority of market portfolios $\overline{n}$ satisfy

$$\sum_{i=1}^{n} N_i = N,$$

$$|\sum_{i=1}^{n}\lambda_i N_i| \le \alpha(N), \tag{54}$$

$$\mathrm{K}(\overline{n}) \ge NH_{max} - N\delta(N). \tag{55}$$

Here $H_{max} = H(\tilde{\nu})$ is the entropy of the security market at a given trading period. The parameter β defined in the proof of Theorem 2.1 plays the important role in our market analysis. We define by analogy with thermodynamics the *market temperature* T such that $\beta = -\frac{1}{T}$. It was mentioned in the proof of Theorem 2.1 that the temperature T can be both positive or negative (and also infinite).

Let at the beginning of the trading period, the total value K of a portfolio be distributed among M' securities divided into groups $\mathcal{G}'_i$, $i = 1,\dots,n$, where $\mathcal{G}'_i \subseteq \mathcal{G}_i$. Moreover, the proportions of this distribution are the same as for the market as a whole, i.e., $\rho = K/M'$ and $p_i = |\mathcal{G}'_i|/M'$. Such securities can be obtained by dividing the initial M securities into "small" groups such that each small group contains N/K initial securities and is contained in some integrated group $\mathcal{G}_i$. We assume for simplicity that N is divisible by K and such a partition is possible. Thus, the partition into small groups is a refinement of the partition $\mathcal{G}_i$, $i = 1, \dots, n$. After that, we can effectively choose one security from each small group; their total number is M'. The portfolio $\overline{k} = (k_1, \dots, k_{M'})$ is defined by a distribution of its total value among these chosen securities. Then elements of new integrated group $\mathcal{G}'_i$, $i = 1, \dots, n$, are securities chosen from the small groups whose union is equal to $\mathcal{G}'_i$. The former return rate λ_i is assigned to each security of this integrated group.

The portfolio $\overline{k}$ and the corresponding macro portfolio $(K_1, \dots, K_n)$ are defined in the same way as in Section 5.1.

Recall that the total value of any portfolio at the end of the trading period is

$$K^{+1} = \sum_{i=1}^{n} K_i(1 + \lambda_i).$$

In the next theorem we formalize an intuitive idea that an investment distributed in a sufficiently complex way changes its value in the same way as the market as a whole.

Theorem 5.1 *Suppose that the rates of return λ_i, where $-1 \le \lambda_i < \infty$ for $i = 1, 2 \dots, n$, be given; in this case the entropy H_{max} of the securities market and the market temperature T are defined. Let also, the portfolio $\overline{k}$ with a value K satisfies*

$$\frac{\mathrm{K}(\overline{k})}{K} \ge H_{max} - \delta(K). \tag{56}$$

1) Let $T > 0$. Then any investment portfolio $\overline{k}$ satisfying the condition (56) is asymptotically nonrisk in case of "dealing for a rise":

$$\liminf_{K \to \infty} \frac{K^{+1}}{K} \ge 1.$$

2) Let $T < 0$. Then any investment portfolio $\overline{k}$ satisfying the condition (56) is asymptotically nonrisk in case of "dealing for a fall":

$$\limsup_{K \to \infty} \frac{K^{+1}}{K} \le 1.$$

3) Let $T = \infty$. Then any investment portfolio $\overline{k}$ satisfying the condition (56) is asymptotically nonrisk in both cases (dealing for a rise and dealing for a fall):

$$\lim_{K \to \infty} \frac{K^{+1}}{K} = 1. \tag{57}$$

Proof. In part 1) of the theorem we have $\beta < 0$. At the beginning of the trading period, we choose the investment portfolio $\overline{k} = (k_1, \dots, k_{M'})$ satisfying (56). Let $\overline{K} = (K_1, \dots, K_n)$ be the corresponding macro-portfolio; let $\nu(\overline{k}) = (K_1/K, \dots, K_n/K)$. In this case inequality

$$\mathrm{K}(\overline{k}) \le KH(\nu(\overline{k})) + o(K) \tag{58}$$

holds as $K \to \infty$.

Let ϵ be a sufficiently small positive number. Assume that at the end of the trading period the total value of the portfolio decreases,

$$\sum_{i=1}^{n} K_i\lambda_i \le -\epsilon K \tag{59}$$

for some sufficiently large K. Let $\overline{\nu} = \nu(\overline{k}) = (\nu_1, \dots, \nu_n)$, where $\nu_i = K_i/K$ and $i = 1, \dots, n$. Recall that the maximum of the entropy is attained at the vector $\tilde{\nu} = (\tilde{\nu}_1, \dots, \tilde{\nu}_n)$. After that, we follow the proof of Theorem 2.1. For all sufficiently large K, we have

$$\delta H(\tilde{\nu}) = H(\overline{\nu}) - H(\tilde{\nu}) =$$

$$= \sum_{i=1}^{n} \frac{\partial H}{\partial \nu_i}(\tilde{\nu})(\nu_i - \tilde{\nu}_i) + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2 H}{\partial \nu_i^2}(\tilde{\nu})(\nu_i - \tilde{\nu}_i)^2 + O((\nu_i - \tilde{\nu}_i)^3)\right) \le$$

$$\le \beta\epsilon + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2 H}{\partial \nu_i^2}(\tilde{\nu})(\nu_i - \tilde{\nu}_i)^2 + O((\nu_i - \tilde{\nu}_i)^3)\right) + o(1) \le -c, \tag{60}$$

where c is a positive constant. This inequality follows from Theorem 2.1 and (18), (59). Hence, an analogous to (22) inequality follows:

$$H_{max} - H(\overline{\nu}) \ge c > 0.$$

This inequality contradicts to inequalities (56) and (58) for all sufficiently large K. This contradiction shows that, for any arbitrary small $\epsilon > 0$ and for all sufficiently large K we have

$$\sum_{i=1}^{n} K_i\lambda_i > -\epsilon K.$$

From this part 1) of the theorem follows. Part 2) of the theorem is proved analogously.

To prove part 3), let us first show that $\liminf\limits_{K\to\infty} \frac{K^{+1}}{K} \ge 1$. Assume that, for some sufficiently small $\epsilon > 0$ and for some sufficiently large K, inequality (59) holds. Rewrite (59) as $\sum\limits_{i=1}^{n} \nu_i\lambda_i < -\epsilon$. Recall that $\sum\limits_{i=1}^{n} \tilde{\nu}_i\lambda_i = 0$. From this, it follows that $|\nu_i - \tilde{\nu}_i| > r\epsilon$ for some i, where r is a positive constant. Then by (60) (for $\beta = 0$) we obtain that, for all sufficiently large K the inequality $H(\tilde{\nu}) - H(\overline{\nu}) \ge c > 0$ holds, where c is a positive constant. After that, we obtain a contradiction as in the proof of part 1). The inequality $\limsup\limits_{K\to\infty} \frac{K^{+1}}{K} \le 1$ can be proved analogously. ∎

Condition 1) of the theorem can be interpreted as follows. Suppose that an investor is dealing for a rise. He forms a sufficiently complex portfolio $\overline{k}$ satisfying (56) at the beginning of the trading period. Suppose also that this investor predicts that the market temperature at a given trading period will be positive: $T > 0$. Then he does not bear a loss (or received a gain) at the end of the trading period if his prediction will valid.

The interpretation of the condition 2) is analogous. We suppose that an investor is dealing for a fall. In this case the investor's predicts that $T < 0$. In this case the investor used the short-selling of a sufficiently complex portfolio and does not loss if his prediction will valid.

It is easy to see that, the sign of T depends on the position of the maximum point of entropy computed under the single constraint (53). For $T > 0$ (for $\beta < 0$) the maximum point of $H(\overline{\nu})$ does not lie below the hyperplane (53); it is easy to see that this happens if $\sum\limits_{i=1}^{n} \lambda_i p_i \geq 0$, i.e., if the mean value of the return rate of all securities does not decrease at the end of the trading period.

For $T < 0$ (for $\beta > 0$) the maximum point of $H(\overline{\nu})$ does not lie above the hyperplane (53); it is easy to see that this happens if $\sum\limits_{i=1}^{n} \lambda_i p_i \leq 0$, i.e., if the mean value of the return rate of all securities does not increase at the end of the trading period. For $T = \infty$ (for $\beta = 0$) the maximum point lies on the hyperplane (53) and the corresponding sum is zero.

In cases 1)-3), the investor is required to predict the sign of the mean value of the return rates for the market as a whole; he does not required for predictions of the return rates for each security. In other words, the investor is *dealing in the mean rise* or the investor is *dealing in the mean fall.*

Received: August 2006. Revised: Oct. 2006.

References

[1] CONNOLLY K.B., *Buying and Selling Volatility.* Wiley. 1997.

[2] COVER T.M., THOMAS J.A., *Elements of Information Theory.* Wiley. 1991.

[3] CSISZÁR I., KÖRNER J., *Information Theory: Coding Theorems for Discrete Memoryless Systems.* New York: Academic, 1981.

[4] GÁCS P., TROMP J., VITANYI P., *Algorithmic Statistics.* IEEE Trans. Inform. Theory. 2001. V. 20. N5. P.1-21.

[5] JAYNES E.T., *Papers on Probability, Statistics and Statistical Physics.* Kluwer Academic Publisher. 2nd edition. 1989.

[6] KOLMOGOROV A.N., *Three Approaches to the Quantitative Definition of Information.* Problems Inform. Transmission. 1965. V.1 N1. P.4–7.

[7] KOLMOGOROV A.N., *The Logical Basis for Information Theory and Probability Theory.* IEEE Trans. Inf. Theory. 1968. IT. V.**14**. P.662–664.

[8] Kolmogorov A.N., *Combinatorial Basis of Information Theory and Probability Theory.* Russ. Math. Surveys. 1983. V.**38** P.29–40.

[9] Kolmogorov A.N., Uspensky V.A., *Algorithms and Randomness.* Theory Probab. Applic. 1987. V.**32**. P.389–412.

[10] Künzi H.P., Krelle W., *Nichtlineare Programmierung.* Berlin Göttingen Heidelberg: Springer, 1962.

[11] Landau L.D., Lifshitz E.M., *Statistical Physics.* Vol. **1**, Oxford, New York: Pergamon, 1980.

[12] Li M., Vitányi P., *An Introduction to Kolmogorov Complexity and Its Applications.* 2nd ed. New York: Springer–Verlag. 1997.

[13] Victor Maslov and Vladimir V'yugin., *Maximum Entropy Principle in Non-Ordered Setting,* Lecture Notes in Computer Science, Volume **3244**, 2004, P.221-233, Springer-Verlag Heidelberg. Algorithmic Learning Theory: 15th International Conference, ALT 2004, Padova, Italy, October 2-5, 2004. Proceedings Editors: Shai Ben-David, John Case, Akira Maruoka.

[14] Shafer G., Vovk V., *Probability and Finance.* It's Only a Game! New York: Wiley, 2001.

[15] Shiryaev A.N., *Probability.* Berlin: Springer. 1984.

[16] Stratonovich R.L., *Teoriya informatsii (Information Theory).* Moscow.: Sovetskoe radio. 1975.

[17] V'yugin V.V., Maslov V.P., *Concentration Theorems for Entropy and Free Energy.* Probl. Inf. Trans. 2005. V.41 N2. P. 134-149.