



---

Soviet-era science, translated into English

# FINITE SETS AND THE SCALAR PRODUCT

The scalar product is usually introduced as the sum of the products of the components of vectors with the same names

1957

SovietRxiv

---

View the original and related papers at <https://sovietrxiv.org/items/ru-195701.53177>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

**Abstract**

**Full Text**

**MATHEMATICS**

**A. M. MOLCHANOV**

## **FINITE SETS AND THE SCALAR PRODUCT**

*(Presented by Academician M. V. Keldysh, 8 V 1957)*

The scalar product is usually introduced as the sum of the products of the components of vectors with the same names

$$(x, x') = x_1 x'_1 + \dots + x_s x'_s.$$

The purpose of the present note is to show that the scalar product can be obtained by a simple limiting transition, starting from finite sets. More precisely, it will be shown that for subsets of finite sets one can introduce a notion of closeness that passes into the notion of scalar product when the number of elements entering the finite set is increased. In order to formulate precise assertions, it is necessary to introduce several definitions.

**Definition 1.** A partition of a finite set into nonintersecting subsets

$$M = m_1 + m_2 + \dots + m_s$$

is called a **component**.

The subsets  $m_1, m_2, \dots, m_s$  are called the **elements of the component**.

The case of a large number of components is usually of interest, since here limiting relations arise.

**Definition 2.** A **grain** is the intersection of several elements of different components

$$p = m_{i_1}^{(1)} m_{i_2}^{(2)} \dots m_{i_N}^{(N)}.$$

**Definition 3.** Two components of a given system of components ( "structure" ) are called **equivalent** if they consist of the same number of sets and the corresponding grains contain the same number of elements.

For example, if the first and second components are equivalent, then the grains

$$p = m_{i_1}^{(1)} m_{i_2}^{(2)} \dots m_{i_N}^{(N)} \quad \text{and} \quad \tilde{p} = m_{i_2}^{(1)} m_{i_1}^{(2)} \dots m_{i_N}^{(N)},$$

obtained from one another by the permutation of  $i_1$  and  $i_2$ , must contain the same number of elements for all  $i_1, i_2, \dots, i_N$ .

**Definition 4.** The **superposition** ( “mixture” ) of two grains

$$p = m_{i_1}^{(1)} m_{i_2}^{(2)} \dots m_{i_N}^{(N)} \quad \text{and} \quad q = m_{k_1}^{(1)} m_{k_2}^{(2)} \dots m_{k_N}^{(N)}$$

is the grain

$$r = m_{l_1}^{(1)} m_{l_2}^{(2)} \dots m_{l_N}^{(N)},$$

in which each of the sets  $m_{l_1}^{(1)}, m_{l_2}^{(2)}, \dots, m_{l_N}^{(N)}$  coincides with the corresponding set of either the grain  $p$  or the grain  $q$ .

One may, of course, speak of the superposition not only of two grains, but also of a larger number of grains.

**Definition 5.** The **measure of closeness** ( “scalar product” ) of two grains  $p$  and  $q$  is the number of coinciding multipliers

$$(p, q) = \delta_{i_1 k_1} + \delta_{i_2 k_2} + \dots + \delta_{i_N k_N}.$$

Let us note that the measure of closeness of two grains never exceeds  $N$ , and if it is equal to  $N$ , then the grains are identical.

For convenience, in what follows we shall use the additive notation for a grain  $m_{i_1} P_1 + m_{i_2} P_2 + \dots + m_{i_N} P_N$  (instead of the multiplicative notation  $m_{i_1}^{(1)} m_{i_2}^{(2)} \dots m_{i_N}^{(N)}$ ), introducing “projection operators”  $P_1, P_2, \dots, P_N$ . Everywhere in this note it is assumed that all  $N$  components are equivalent, and that the number of elements in each component  $s$  is fixed, while the number of components  $N$  increases without bound. Since the components are equivalent, the additive notation can be simplified by collecting “like” terms for identical (more precisely, equivalent) sets  $m$ :  $p = m_{i_1} p_1 + m_{i_2} p_2 + \dots + m_{i_N} p_N = \sum_m m P(m)$ . Thus a grain is written as an operator-valued function on the elements of a component  $P(m)$ . The measure of closeness of two grains  $p$  and  $q$  is conveniently expressed in terms of the functions  $P(m)$  and  $Q(m)$ :  $(p, q) = \sum_m \text{Sp}[P(m)Q(m)]$ . The superposition of two or several grains is written, as is easy to verify, by means of projection operators  $A_1, A_2, \dots, A_l$  (whose sum is equal to the identity):  $P(m) = A_1 P_1(m) + A_2 P_2(m) + \dots + A_l P_l(m)$ .

In order to obtain (in the limit, of course) the usual scalar product in a linear vector space, let us consider the sets obtained by taking the union of all grains equivalent to one another (with respect to permutation of components). For brevity we shall call these sets **symmetric grains**. It is clear that each symmetric grain is determined by a numerical function  $n(m) = \text{Sp} P(m)$ , since a permutation of components is the same as a permutation of the operators

$P_i$ . Symmetric grains consist, generally speaking, of a number of grains that increases with  $N$ . If one takes at random, from two symmetric grains, one representative from each and finds the measure of their closeness, then its value will, of course, depend on which representatives were taken. However, for sufficiently large  $N$ , in the overwhelming majority of cases the results obtained are very close. Moreover, if one symmetric grain is given by the function  $n(m)$ , and another by the function  $n'(m)$ , then the measure of closeness of their representatives in the overwhelming majority of cases differs little from the number

$$\frac{1}{N} [n(m_1)n'(m_1) + n(m_2)n'(m_2) + \dots + n(m_s)n'(m_s)].$$

Thus, if one introduces the functions normalized to unity  $a(m) = \frac{1}{N}n(m)$  and  $a'(m) = \frac{1}{N}n'(m)$ , as well as the measure of closeness normalized to unity,

$$\frac{1}{N} \sum_m \text{Sp}[P(m)P'(m)],$$

then as  $N \rightarrow \infty$  there exists a limit of this measure, and it is equal to the usual scalar product of the functions  $a(m)$  and  $a'(m)$  (which, in fact, justifies the very name “scalar product” for the measure of closeness of grains). Analogously, superposition in the limit becomes a linear combination with the sum of weights equal to unity. Therefore the linear space of  $s$ -dimensional vectors (or functions on  $s$  points) with nonnegative coefficients whose sum is equal to unity may be regarded as an asymptotic description of the properties of structures—finite sets divided into a large number of equivalent components. The scalar product then arises naturally as the asymptotics of the measure of closeness. (Let us note in passing that the measure of closeness itself apparently has an analogue in information theory, where the notion of closeness of two messages is introduced. However, it is unclear whether the consideration of equivalent components and symmetric grains corresponds to anything in information theory.)

The proof (and the precise formulation with remainder terms) of the limit theorems mentioned above is obtained by reduction to the theory of summatory functions <sup>(1)</sup> and by then applying the usual probabilistic methods (characteristic functions and the saddle-point method). Let us illustrate this assertion with the example of the scalar product.

Since the scalar product is a function of a pair of grains, the basic set that must be considered is the set of sequences of pairs of elements  $m_{i_1}^{(1)}, m_{k_1}^{(1)}; m_{i_2}^{(2)}, m_{k_2}^{(2)}; \dots; m_{i_N}^{(N)}, m_{k_N}^{(N)}$ . This set is the direct product of  $N$  sets of the form  $(m, m')$ . On each of these sets we define  $2s$  functions as follows:  $f_k(m, m') = \delta_{m_k m}$  and  $f'_k(m, m') = \delta_{m'_k m'}$ ,  $k = 1, 2, \dots, s$ . We next consider the  $2s$  summatory functions  $F_k = f_k^{(1)} + f_k^{(2)} + \dots + f_k^{(N)}$  and  $F'_k = f'_k{}^{(1)} + f'_k{}^{(2)} + \dots + f'_k{}^{(N)}$ . The scalar product is likewise a summatory function  $G = g^{(1)} + g^{(2)} + \dots + g^{(N)}$ , where in each component

$g = f_1 f'_1 + f_2 f'_2 + \dots + f_s f'_s$ . It is clear that  $F_k$  are nothing other than the “occupation numbers” of the sets  $m_k$ , so that fixing the values  $F_k$  is equivalent to specifying a symmetric grain.

Consequently, the asymptotics of the scalar product reduces to the well-known problem in the theory of summatory functions:

Given the values of several ( $2s$ ) summatory functions  $F_k = n(m_k) = Na(m_k)$ ,  $F'_k = n'(m_k) = Na'(m_k)$ ,  $k = 1, 2, \dots, s$ . It is required to find the mean value and the variance of the summatory function  $G$  on the set determined by fixing the values  $F_k$  and  $F'_k$ .

The principal role in solving such a problem is played, as is known <sup>(1)</sup>, by the structural function  $\Omega_N(a_1, a_2, \dots, a_s, a'_1, a'_2, \dots, a'_s)$ , equal to the number of elements in  $M$  on which the functions  $F_k$  and  $F'_k$  take the prescribed values  $Na_k$ ,  $Na'_k$ . An important property possessed by  $\Omega_N$  is that it is the convolution of the structural functions corresponding to the individual components. This circumstance makes it possible, by introducing the characteristic function—the Laplace transform of the structural function—to apply, as was done in <sup>(1)</sup>, the analytic apparatus of probability theory for proving limiting relations. However, it is often convenient to use the saddle-point method directly. Lack of space does not permit us to dwell in greater detail on the details of the arguments. For the same reason we must pass over the very important question of how linear vector spaces with real (and not only positive), and still more with complex, coefficients can arise as limiting formations. Here we can only point out that their appearance is connected with a deeper study, both statistical and structural, of finite sets divided into equivalent components.

*Department of Applied Mathematics  
of the V. A. Steklov Mathematical Institute  
Academy of Sciences of the USSR*

Received  
7 V 1957

## CITED LITERATURE

1. A. Ya. Khinchin, *Mathematical Foundations of Statistical Mechanics*, 1943.

*Note: Figure translations are in progress. See original paper for figures.*

*Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.*