Chapter 1
Genetic Code:
Emergence of Symmetrical Pattern,
Beginnings of Matrix Genetics

ABSTRACT
This chapter is devoted to symmetrical analysis for genetic code systems. The genetic coding possesses the noise-immunity. Mathematical theories of the noise-immunity coding and discrete signals processing are based on matrix methods of representation and analysis of information. These matrix methods, which are connected closely with relations of symmetry, are borrowed for a matrix analysis of ensembles of molecular elements of the genetic code. This chapter describes a uniform representation of ensembles of genetic multiplets in the form of matrices of a cumulative Kronecker family. The analysis of molecular peculiarities of the system of nitrogenous bases reveals the first significant relations of symmetry in these genetic matrices. It permits to introduce a natural numbering the multiplets in each of the genetic matrices and to give the basis for further analysis of genetic structures. The connection of the numerated genetic matrices with famous matrices of dyadic shifts is demonstrated.

INTRODUCTION AND BACKGROUND
Bioinformatics is defined frequently as the branch of life science that deals with the study of application of information technology to the field of molecular biology. The primary goal of bioinformatics is to increase our understanding of biological processes. The term bioinformatics was coined by Paulien Hogeweg in 1978 for the study of informatics processes in biotic systems.

The genetic code is a key to bioinformatics and to a science about biological self-organizing on the whole. The modern science faces the necessity of understanding and system explanation of mysterious features of ensemble of molecular structures of the genetic code. Why does the genetic alphabet consist
of the four letters? Why does the genetic code encode 20 amino acids? How is the system structure of the molecular genetic code connected with known principles of quantum mechanics, which were developed to explain phenomena on atomic and molecular levels? Why has nature chosen the special code conformity between 64 genetic triplets and 20 amino acids? Can knowledge about the structural essence of the genetic code be useful for mathematical natural sciences on the whole? What kind of mathematical approach should be chosen among many possible approaches to represent and model structuralized ensembles of molecules of the genetic code?

Achievement of deep understanding the genetic code should promote an inclusion of a science about it into the field of the mathematical natural sciences. To provide it, the direction of searches should be based on fundamental mathematical methods and concepts. Methods and principles of symmetry, as well as the matrix analysis, are some of bases of modern mathematical natural sciences. While biological structures are genetically inherited, morphological structures of biological bodies are characterized by many kinds of symmetry. It is known from the history of molecular genetics that investigations of symmetry in genetic molecules have given essential results already. Revelations of new symmetric structures in molecular-genetic systems produce a set of useful heuristic associations due to analogies with known symmetric structures in other scientific fields: quantum mechanics, theory of digital communication and noise-immunity coding, geometry, etc.

Genetic coding possesses the noise-immunity, which allows descendants to be similar to their parents, despite of strong disturbances and noise in the environment of biological molecules. It reminds one of the effective noise-immunity of modern systems of digital communication and signal processing, which is reached by means of special mathematics. The mathematics is based on matrix and symmetric methods of representation and analysis of signals. It’s natural to ask whether it is possible that these mathematical methods, which were developed for digital technique, can be applied in the adequate manner to studying the genetic code?

The objectives of Chapter 1 are the following:

- The explanation of the choice of symmetric and matrix methods of analysis of the genetic code as prospective and adequate methods to investigate and to model structural interrelations among various parts of the integral molecular system of the genetic code;
- The description of the main data about molecular structures of the genetic code;
- The demonstration of the possibility of representation of all sets of genetic multiplets, which differ from each other by their lengths and compositions, in the well-ordered symmetrical form by means of the Kronecker family of the genetic matrices;
- The explanation of the fact that all multiplets in this general matrix form of presentation of their sets can be numerated individually by means of taking into account the symmetrical binary sub-alphabets of the four-letter genetic alphabet;
- The revelation of the connection between the genetic matrices and the matrices of diadic shifts, which are utilized in the theory of discrete signal processing long ago as fundamentals of some special methods of analysis and synthesis of signals.