International Journal of Modern Physics B, Vol. 13, No. 23 (1999) 2795–2885

SYMMETRY AND SYMMETRY BREAKING: AN ALGEBRAIC APPROACH TO THE GENETIC CODE*

JOSÉ EDUARDO M. HORNOS†, YVONE M. M. HORNOS‡ and MICHAEL FORGER¶

†*Departamento de Física e Ciências dos Materiais,*
Instituto de Física de São Carlos,
Universidade de São Paulo, Caixa Postal 369,
BR-13560-970 São Carlos, S.P., Brazil

‡*Serviço de Assessoria e Proteção Radiológica (SAPRA),*
Caixa Postal 525, BR-13560-922 São Carlos, S.P., Brazil

¶*Departamento de Matemática Aplicada,*
Instituto de Matemática e Estatística, Universidade de São Paulo,
Caixa Postal 66281, BR-05315-970 São Paulo, S.P., Brazil

Received 12 May 1999

We give a comprehensive review of the algebraic approach to the genetic code originally proposed by two of the present authors, which aims at explaining the degeneracies encountered in the genetic code as the result of a sequence of symmetry breakings that have occurred during its evolution. We present the relevant background material from molecular biology and from mathematics, including the representation theory of (semi) simple Lie groups/algebras, together with considerations of general nature.

Contents

*Work supported by CNP_q (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), Brazil.

1. Introduction

The discovery of the molecular structure of DNA by Watson and Crick in 1953[1,2] is one of the landmarks in the history of science. The substance itself had been known since 1869 and had been recognized as the carrier of genetic information in 1944, but the basis of its biological function, of its stability and of its capability of self-reproduction had remained a mystery. With the advent of the double helix model, genetics entered into a new area and molecular biology was born.

The double helix model of Watson and Crick made it clear that genetic information is stored in DNA in the form of sequences of nucleotides, but it did not by itself provide any hint as to exactly how the information is coded in such a sequence, nor how it is expressed in order to exert its specific biological function. In the language of information theory, this question constitutes the problem of deciphering the *genetic code*. In particular, one of the first major tasks of molecular biology was to find out how the genetic code steers the synthesis of proteins — the most complicated molecules that appear in living organisms, with all their enormous diversity — assembled from amino acids as their basic building blocks.

The first major step in this direction was the discovery by Crick *et al.* in 1961[3] that the genetic code is a triplet code, or in other words, that its elementary unit of information, commonly called a *codon*, is a sequence of 3 nucleic bases, each of which represents uniquely one amino acid (or a termination signal). Subsequent experimental work[4–8] led to a complete classification of the correspondence between codons and amino acids, fully accomplished in 1966 and assembled into a standard table that can be found in any textbook on biochemistry or genetics; it is here reproduced as Table 1.

The importance of this genetic code table for biology can hardly be overestimated and is comparable only to that of the periodic table of elements for chemistry. The two tables show interesting similarities also in other respects. For example, both of them exhibit intriguing regularities, such as the appearance of groups of elements with similar chemical properties, assembled into columns of the periodic table, or the appearance of groups of codons representing the same amino acid, assembled into so-called family boxes of the genetic code table. However, the tables as such do not explain the origin of such regularities. In fact, it took more than 50 years to unveil the reasons why the periodic table of elements is just the way it is, and this understanding was only possible as a result of fundamental advances in physics

Table 1. The standard genetic code (mRNA codons versus amino acids).

	Second base				
First base	U	C	A	G	Third base
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	TERM	TERM	A
	Leu	Ser	TERM	Try	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

such as the discovery of the subdivision of atoms into nucleus and electronic shell and the development of quantum mechanics to explain the organization of the latter. Similarly, the genetic code table has for almost 30 years defied all attempts at explanation, and it does therefore not seem unreasonable to hope that such an explanation will be capable to provide fundamental new insight into molecular biology.

An outstanding feature of the genetic code table is its degeneracy. That degeneracy is unavoidable should be obvious from the fact that there are 64 codons (the number of three-letter words in an alphabet formed by four letters), whereas only 20 amino acids are encountered in all known living organisms. But experience accumulated in more than 50 years of research in physics has revealed, as a golden rule, that *degeneracy is associated with and a consequence of symmetry*. To avoid misunderstandings, it should be pointed out that these symmetries are not necessarily spatial or space-time symmetries; rather, they may refer to transformations in an abstract auxiliary space. It is such a kind of abstract *internal symmetry* that should be related to the degeneracy of the genetic code — a symmetry that acts as an organizing principle for the way in which genetic information is stored and for the way it regulates the process of protein synthesis. This is the spirit of the algebraic approach to deciphering the structure that underlies the genetic code.

The aim of the present review is to give an account of this algebraic approach and to explain some of the background from different areas of science: mathematics, physics, chemistry and biology.

2. The Evolution of Matter

In the course of this century, the notion of evolution — originally introduced into biology through the work of Charles Darwin on "The Origin of Species" — has become one of the most important and universal paradigms of modern science, appearing in practically every area in connection with possible changes in the schemes and patterns into which matter organizes itself.

One prominent example is cosmology, where the transition from a static to a dynamic picture of the universe as a whole is due to Einstein's general theory of relativity and Hubble's subsequent discovery of the red-shift in the spectra of distant galaxies as evidence for the expansion of our universe. Moreover, it has become clear that in the course of this expansion, matter in the universe has undergone and still undergoes a steady process of evolution, proceeding in stages from simpler forms to levels of organization of ever increasing complexity. Roughly speaking, these stages may be characterized as physical evolution, chemical evolution and biological evolution.

Let us begin with a few comments about the physical evolution of the universe. According to the standard (hot) big bang model, this evolution begins, some 10–20 billion years ago, with the big bang itself. In the very first phase (during the first 10^{-44} or 10^{-43} seconds), not only all matter, but even the gravitational field was governed by laws of quantum physics, about which practically nothing is known — except for the fact that the usual concept of space-time as a smooth Lorentz manifold had no meaning and that we have no definite idea as to what should replace it. In the next phase (from about 10^{-44} or 10^{-43} seconds to 10^{-10} seconds after the big bang), the situation is not much better, since we have very little experimental information on the behavior of matter at such high energies. It is only at the end of this period that we enter firm ground, since we may begin to trust the predictions of the standard model of elementary particle physics. The most prominent events afterwards are first of all the breaking of the electroweak symmetry, followed by the condensation of nuclear matter, from a state of a plasma made of (asymptotically free) quarks and gluons to a state of a plasma composed of hadrons (protons, neutrons and other, less stable, strongly interacting particles). Next comes the decoupling of the neutrinos from the remaining particles and the annihilation of electrons and positrons into radiation, followed by the primordial nuclear synthesis, which has produced almost exclusively helium: all this happened during the famous "first three minutes" of the universe. Much later (about 680.000 years after the big bang), we finally record the recombination of the remaining electrons and nuclei into atoms and, as a result, the decoupling of the photons: the universe became transparent.

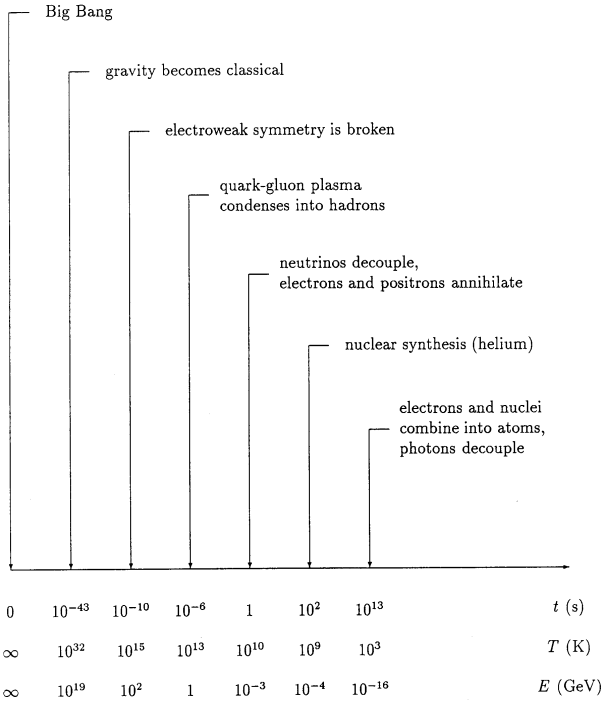


Fig. 1. Physical evolution of the early universe.

A schematic representation of this evolution is given in Fig. 1, where the time scale (in seconds) is supplemented by a temperature scale (in Kelvin) and an energy scale (in gigaelectronvolts). For a more detailed discussion, the reader is referred to the popular book.[9]

The chemical evolution of the universe starts with nucleosynthesis in stars, providing the raw material for the organization of matter at the atomic and molecular level. In our galaxy, this process began about 10 billion years ago (this is the age of the oldest stars in the galaxy) and is continuing up to this day. Other stages of chemical evolution are the formation of molecules and radicals in interstellar clouds.

The biological evolution of life on earth must have begun shortly after the formation of the solar system and in particular of our planet, about 5 billion years ago. The first cell seems to have appeared about 4 billion years ago, since the oldest fossils that we presently know of (the cyanobacteria recently found in Australia) are 3.8 billion years old, whereas evidence for the first animals goes back to only about 1.5 billion years ago. In this scheme, the subject of interest here, namely the evolution of the genetic code for protein synthesis, must have occurred during the earliest phase of the formation of life on earth, slightly less than 4 billion years ago.[10] For a schematic representation, see Fig. 2.

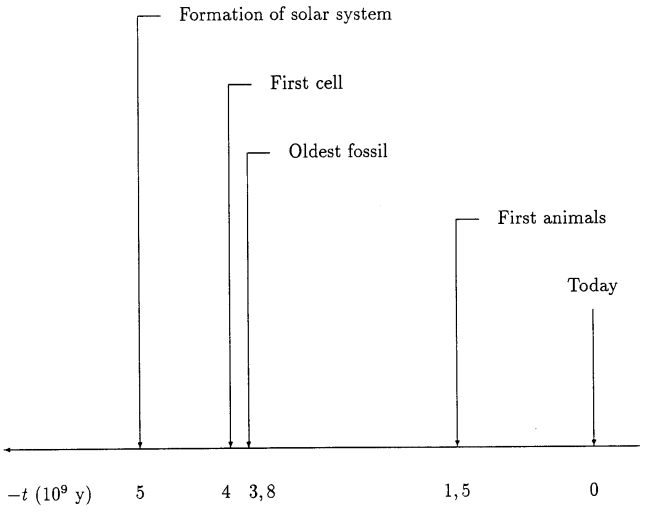


Fig. 2. Early biological evolution on earth.

An important aspect of evolutionary processes, and often one of their most prominent features, is the phenomenon of symmetry breaking: it occurs when an initial state of high symmetry evolves to a subsequent state of lower symmetry. One typical example in cosmology, already mentioned above, is the breaking of the electroweak symmetry, that is, the symmetry between the electromagnetic forces and the weak nuclear forces: this is one of the salient features of the standard model of elementary particle physics, also known as the Glashow–Salam–Weinberg model. Another example from the same area is the process of formation of galaxies (or galaxy clusters) from the "primordial soup" (a hot gas composed mainly of photons, electrons, protons, helium nuclei and neutrinos): during this process, spatial homogeneity of the universe was lost, i.e. the translational symmetry present in the previous stage of evolution was broken.

The notion of evolution is not restricted to (irreversible) processes in real time, but can also be applied to studying the behavior of systems as function of other external parameters, such as temperature, pressure, density, etc. In particular, phase transitions are very often assossiated with changes in symmetry. A common example is the freezing of a substance, where the phase transition from the liquid state to a crystalline solid state is accompanied by the breaking of the continuous translational and/or rotational symmetry to a discrete one.

The main point of the work to be reported here is that the same phenomenon of symmetry breaking has also played an important role in the evolution of the genetic code and that it may actually be used as a guiding principle to analyze, by purely mathematical means, how and through which intermediate steps this evolution has occurred.

3. Basic Building Blocks of Matter

Apart from the many obvious and fundamental differences between physics and biology, the organization of inanimate and of animate matter also exhibits some surprising structural similarities, the main one being the fact that both forms of matter fall into several big classes, each of which is constructed out of just a few basic building locks. In elementary particle physics, these are

- quarks, as constituents of hadronic matter,
- leptons, as constituents of leptonic matter,
- gauge bosons, as carriers of interactions.

In biology, we encounter

- sugars, as constituents of carbohydrates and energy sources,
- lipid acids, as constituents of membranes,
- amino acids, as constituents of proteins,
- nucleic acids, as constituents of the information carriers DNA and RNA.

It is this structural analogy that lends support to the idea that basic concepts which have turned out to be fruitful in one area may very well serve as guiding principles for the other one. One example discussed in the previous section is the notion of evolution which originated in biology and, transferred to physics and chemistry, has led to great progress in the understanding of complex open systems. Conversely, it is to be expected that the ideas of symmetry and symmetry breaking, which have been so enormously successful in physics, will prove to be useful in biology as well.

3.1. *Proteins and amino acids*

The family of amino acids contains more than 200 known variants, but in all forms of life — independently of species — only 20 of them are systematically used in proteins. Occasionally, an unusual amino acid may appear in a protein, but it arises by posterior modification of one of the 20 fundamental amino acids, after the protein has already been synthesized.

The molecular structure of amino acids is very simple. Their common feature is that a carbon atom (the alpha carbon) forms covalent bindings with four different groups: an amino group (NH_2), a carboxyl group (COOH), a hydrogen atom (H) and a radical (R) which is characteristic of the particular amino acid. This schematic structure is shown in Fig. 3. The radical R can vary from the simplest case of glycine, where it is just a hydrogen atom, to more complex structures that may involve aromatic rings or aliphatic chains. Various attempts to correlate physico-chemical properties of amino acids with the genetic code have been made, some of which are clearly relevant. One of these is polarity that plays an important role concerning the behavior of an amino acid in aqueous solutions (hydrophobic/hydrophilic).[11]

A protein is constructed as a polymer, consisting of a chain of amino acids connected through peptide bonds in which the carbon atom in the carboxyl group

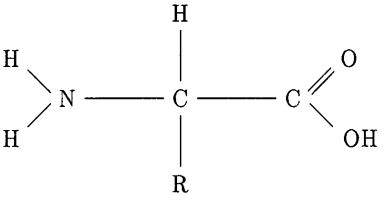


Fig. 3. General molecular structure of amino acids.

of one amino acid binds covalently to the nitrogen atom in the amino group of the next. After the protein is assembled it will fold into a (usually quite complicated) three-dimensional array.

Another remarkable aspect of amino acids is that due to the presence of an asymmetric carbon atom, chiral invariance is broken. Amino acids are divided into two chiral symmetry classes, namely left-handed and right-handed, and it so happens that only the left-handed isomers appear in proteins. The understanding of the mechanism in evolution that guided the choice of the left-handed amino acids is presently an important and vast area of research.

3.2. *DNA, RNA and nucleic bases*

In all forms of life on earth, genetic information is stored in two polymers called DNA (deoxyribonucleic acid) and RNA (ribonucleic acid), composed of structural units called nucleotides. Each of these is made of

- a sugar molecule (deoxyribose in DNA, ribose in RNA),
- a phosphate group,
- one of four different nucleic bases:
  A (adenine), C (cytosine), G (guanine) and
  T (thymine) in DNA or U (uracil) in RNA.

In the regulation of the mechanisms responsible for the sustaining of life and, in particular, the synthesis of proteins, DNA is the primary genetic material, whereas RNA is the secondary genetic material.

The molecular structure of the sugars, which form a five-member ring with four carbon and one oxygen atom, differing only in the substitution of a hydroxyl group attached to one of the carbon atoms (ribose) by a hydrogen atom (deoxyribose), is shown in Fig. 4. Similarly, the molecular structure of the nucleic bases is exhibited in Fig. 5. Cytosine, thymine and uracil are derived from a mother molecule called pyrimidine, a single six-member ring with four carbon and two nitrogen atoms, whereas adenine and guanine are derived from a mother molecule called purine, a double ring made of a six-member ring with four carbon and two nitrogen atoms and a five-member ring with three carbon and two nitrogen atoms, fused along two of the carbon atoms; these mother molecules are shown in Fig. 6. The only difference

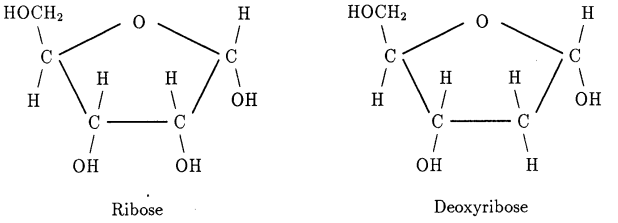


Fig. 4. Molecular structure of ribose and deoxyribose.

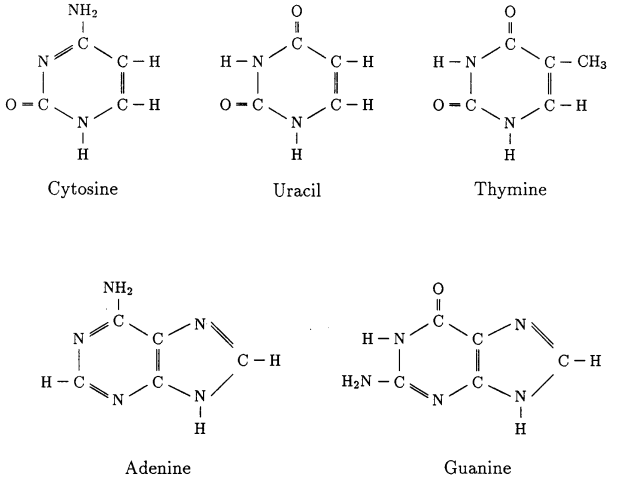


Fig. 5. Molecular structure of nucleic bases.

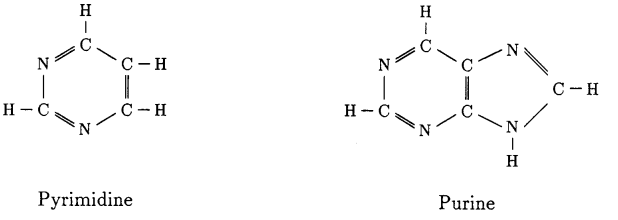


Fig. 6. Pyrimidine and purine.

between thymine and uracil is that the former has a methyl group instead of a hydrogen atom attached to one of the carbon atoms, but this difference does not play a significant role for any of the chemical or biological properties to be discussed in what follows.

The substance DNA has been known since 1869, but it was not recognized as the carrier of hereditary information until 1944.[12] Before reliable chromatographic methods became available, it was believed that the content of the four nucleic

bases encountered in DNA was the same for all life forms. Between 1949 and 1953, a quantitative analysis of the base content was performed by Chargaff and colleagues, who found it to be the same in different tissues from the same species but to vary from species to species. Another important result of this research, known as Chargaff's rule, was the remarkable fact that the ratio of the A-content to T-content and of the C-content to G-content is always one, independent of species. At about the same time, analysis of X-ray diffraction data showed that DNA is a string-like molecule with two regular spacings along the fiber axis. Other data that became available were the dimensions and the stereochemical structure of the purine and pyrimidine bases. However, none of the models for DNA proposed until 1953 provided an explanation for the (highly precise) mechanism of its replication, for its stability or for Chargaff's base content rule.

These questions were convincingly answered by the double helix model of Watson and Crick.[1,2] According to this model, a DNA molecule consists of two strands, formed by two (right-handed) helices coiled together around a common central axis and running in opposite directions. Each helix is composed of a large sequence of nucleotides, consisting of a nucleic base covalently bound to a sugar molecule, and the sugar molecules are interconnected by the phosphate groups to form a strand. The double helix arises through hydrogen bonds coupling the nucleic bases of one strand to the corresponding nucleic bases of the other strand, according to the *Watson-Crick pairing rules*: A pairs only with T and C pairs only with G. This marvelous stereochemical structure can be compared to a staircase in a medieval castle, given the fact that the sugar molecules and the nucleic bases are essentially planar molecules, approximately oriented at right angles to each other: each step in the staircase consists of a pair of two nucleic bases coupled to each other by hydrogen bonds, whereas the sugar molecules provide the railing. See Fig. 7.

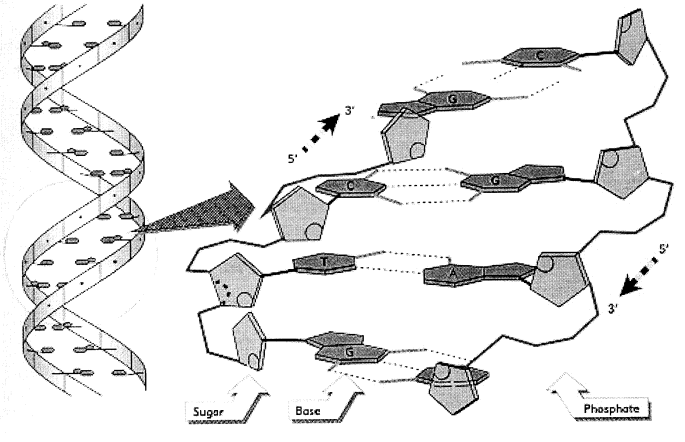


Fig. 7. Schematic structure of the double helix.

The double helix model of Watson and Crick, together with the Watson–Crick pairing rules, not only explains the basic physical and chemical properties of DNA but also provides a mechanism by which genetic information can be replicated with great precision. For example, Chargaff's base content rule is a direct consequence of the Watson–Crick pairing rules. Moreover, the double helical structure enforces the stability of DNA because the two coils can be separated only by completely unwinding them and breaking each hydrogen bond (which requires an energy of the order of 0.1 eV per bond). Finally, the nucleotide sequence in each of the two helices is completely determined by that in the other: this is the basis for the replication process that guarantees genetic continuity between parent and daughter cells, known as semi-conservative self-replication: each parental DNA strand serves as the template for the production of a new complementary daughter strand. The unwinding and complementary copying of the strands, with great fidelity, is achieved through the action of highly specific enzymes (DNA polymerases), together with a number of other regulatory molecules. Other enzymes are responsible for the correct (re-) combination of the two strands.

In RNA, there is only one strand, but base pairing through hydrogen bonds still occurs because the strand can fold back onto itself and form double helical segments of paired nucleic bases, interrupted by loops with unpaired nucleic bases. This kind of structure is common to all forms of RNA, namely

- mRNA — messenger RNA,
- tRNA — transfer RNA,
- rRNA — ribosomal RNA,

despite their different functions in the process of protein synthesis (see below).

The Watson–Crick pairing rules are based on the possibilities of forming hydrogen bonds between nucleic bases and can be understood by realizing the special role played by (a) their molecular size and (b) their "free" nitrogen atom and the groups attached to the two adjacent carbon atoms, taking into account the position and orientation of the binding to the sugar-phosphate backbone.

(a) In order to have base pairs of a well-defined universal size, a pyrimidine must always combine with a purine, because combining a pyrimidine with another pyrimidine would lead to a pair that is too small, whereas combining a purine with another purine would lead to a pair that is too large. It is stereochemically obvious that such pairs cannot fit into the double helical structure of DNA or the paired segments of RNA without disrupting the staircase.

(b) In the case of pyrimidines, the "free" nitrogen atom is the one not bound to the sugar-phosphate backbone, whereas in the case of purines, it is the one not adjacent to one of the two carbon atoms shared by the two fused rings. (See Figs. 5 and 6.) Note also that in one of the two pyrimidines and one of the two purines (T/U and G), this "free" nitrogen atom is hydrogenated, while in the other two (C and A), it is not. In order to form hydrogen bonds, a hydrogenated

"free" nitrogen atom or an amino group attached to one of the two adjacent carbon atoms can act as a donor, whereas a non-hydrogenated "free" nitrogen atom or an oxygen atom attached to one of the two adjacent carbon atoms can act as an acceptor.

Using these criteria, one can construct canonical base pairs $T/U - A$, with two hydrogen bonds and $C - G$, with three hydrogen bonds; they are shown in Fig. 8. Note that the positions and orientations of the bindings to the two sugar-phosphate backbones are always the same. Comparing with the other base pair $T/U - G$ shown in Fig. 9, with two hydrogen bonds, this is not so: there is a slight shift in position. Correspondingly, this possibility does not occur in DNA. However, $U - G$ base pairs do play an important role in codon-anticodon pairing between mRNA and tRNA molecules, according to the wobble rule first proposed by Crick[13] and to be discussed later. Of course, non-conventional base pairings between a pyrimidine and a purine may also occur in DNA, as the result of a tautomeric base change (e.g. as

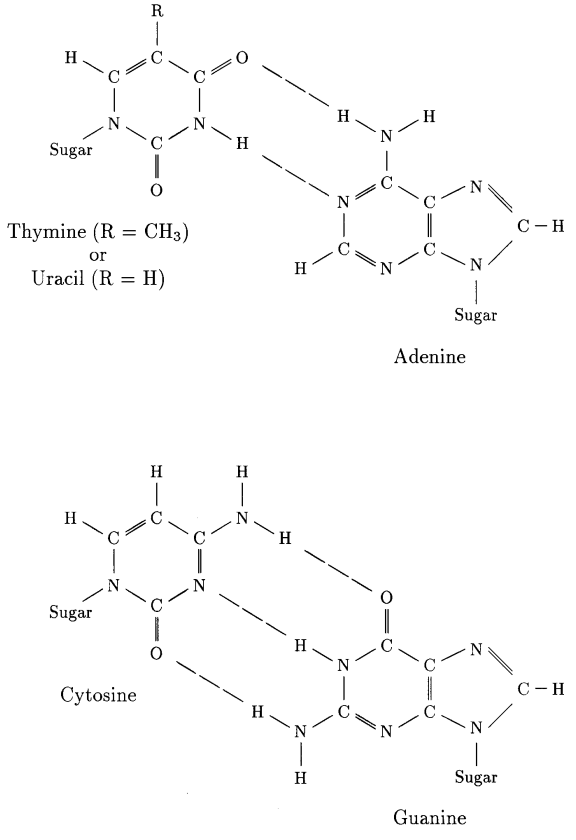


Fig. 8. Watson–Crick pairs of nucleic bases (dashed lines indicate hydrogen bonds).

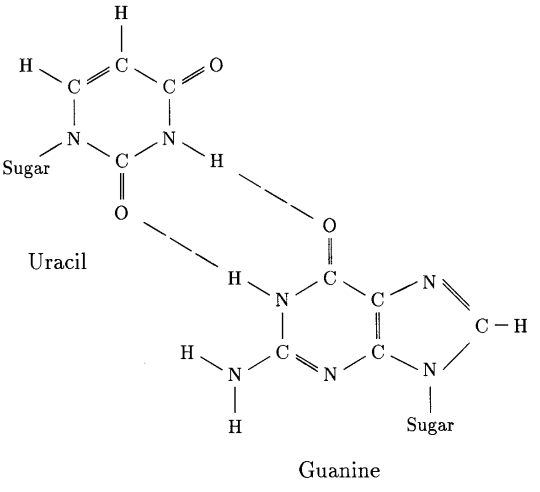


Fig. 9. Wobble pairing between uracil and guanine (dashed lines indicate hydrogen bonds).

the consequence of a point mutation), but we shall disregard such irregularities because they are not the central point of our analysis.

Summarizing, we may infer that the main property of nucleic bases which is at the heart of the special role they play in molecular biology and genetics is that each of them has a specific and well-defined partner to form a base pair. The hydrogen bonds between adenine and thymine/uracil and between cytosine and guanine shown in Fig. 8 represent a specific kind of molecular interaction that will be called Watson–Crick interaction: it is stronger in $C-G$ pairs (3 hydrogen bonds) than in $A-T/U$ pairs (2 hydrogen bonds), the binding energy for each hydrogen bond being of the order of 0.1 eV. We shall also speak of Watson–Crick duality to indicate the fact that nucleic bases come in dual pairs (A with T/U and C with G). The situation is analogous to that in classical mechanics, quantum mechanics or thermodynamics, where dynamical variables always appear in canonically conjugate pairs, such as positions and momenta, angles and angular momenta or temperature and entropy. This suggests introducing a special class of canonical transformations, namely the ones that transform each variable to its canonically conjugate variable. Correspondingly, we shall call the mathematical transformation that takes each nucleic base to its canonically conjugate or Watson–Crick dual nucleic base a Watson–Crick transformation (WCT). A transformation under which a pyrimidine (purine) base goes to the other pyrimidine (purine) base will be called a pyrimidine-purine transformation (PPT).

4. The Process of Protein Synthesis

Protein synthesis in cells implies a processing of the information coded into their DNA to assemble the multitude of proteins needed for the correct functioning of

an organism. Given the fact that in eukaryotic cells, the DNA is contained in the nucleus whereas the synthesis of proteins occurs in the ribosomes, which are organelles located in the cytoplasm, Jacob and Monod postulated the existence of a mediator that transports the genetic information from the DNA in the nucleus to the ribosomes.[14] This mediator was soon after identified to be an RNA molecule and called messenger RNA (mRNA). Therefore, the process of protein synthesis decomposes naturally into two stages: *transcription* and *translation*. This flow of information from DNA to proteins is schematically depicted in Fig. 10.

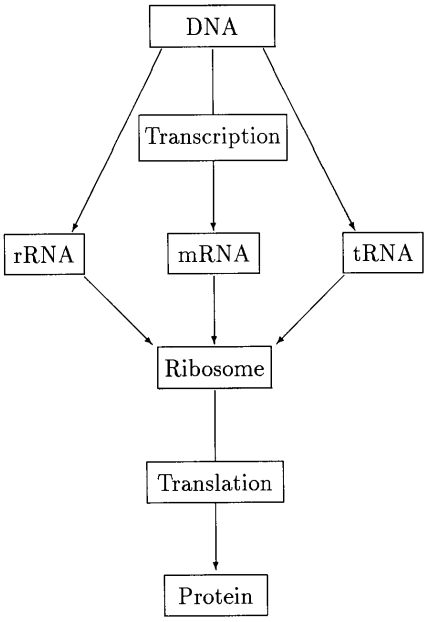


Fig. 10. Transcription and translation.

4.1. *Transcription and translation*

The first step in protein synthesis is transcription, during which the information about the sequence of amino acids of a particular protein, contained in the DNA, is copied to an mRNA molecule, according to the rules of Watson–Crick duality; this mRNA molecule then carries the information to the ribosomes. The transcription process is very similar to the process of semi-conservative self-replication of DNA, where one of the two strands of the parental DNA molecule serves as the template for the synthesis of the dual strand of its descendant. This requires unwinding the double helix and breaking the hydrogen bonds of the parental DNA molecule, which is achieved through specific enzymes and other regulatory molecules, resulting in a copying process of great fidelity and reliability.

An mRNA molecule contains the coding sequence of a specific protein to be synthesized in the ribosome. It is much smaller than the DNA molecule from which it is copied, but may still carry thousands of nucleotides. The concentration of a particular mRNA in the cytoplasm is proportional to the concentration of the corresponding protein.

The second step in protein synthesis is translation, which occurs in the ribosomes: small organelles inside the cytoplasm mainly formed by a few large RNA molecules called ribosomal RNA (rRNA) and a number of well established proteins. The translation process depends crucially on a third class of RNA molecules called transfer RNA (tRNA). A tRNA molecule serves as the carrier of a specific amino acid from the cytoplasm to the ribosomes and simultaneously provides the mechanism to read the information contained in the codons of the mRNA template. It is the smallest type of RNA molecule, containing between 74 and 95 nucleotides. All transfer RNA molecules have a very similar secondary structure in the form of a cloverleaf, formed and stabilized by base pairing between short complementary nucleotide sequences, with one open-ended acceptor arm and three loops (and in some cases a fourth extra loop whose function is not yet known): one of these three loops contains a triplet of unpaired nucleic bases called the *anticodon*. The acceptor arm is that part of the structure that may be charged by one, and only one, of the 20 fundamental amino acids (through the action of one of 20 highly specific enzymes known as aminoacyl synthetases), whereas the anticodon is responsible for the recognition of one or several codons in the mRNA representing the correct amino acid.

The translation process is initiated when the ribosomal complex recognizes the starting signal exhibited on the mRNA and couples to the mRNA molecule, starting to slide along it like a pearl on a thread, in steps of three nucleotides at a time. Whenever an appropriate tRNA molecule, charged with its amino acid and exhibiting the correct anticodon to couple to the codon exhibited on the mRNA, enters the ribosomal complex, the amino acid will be released and linked to the already existing chain of amino acids through the formation of a new peptide bond. Subsequently, the discharged tRNA molecule will also be released and the ribosomal complex moves on to the next codon on the mRNA, to repeat the process, until a termination signal (stop codon) is reached and the mRNA molecule is released from the ribosomal complex as well.

For more details, the reader is referred to standard textbooks on biochemistry and genetics, such as, for example.[15–18]

4.2. *The genetic code*

Independently of the biochemical mechanisms involved in the translation process, the flow of information as such (see Fig. 10) must be governed by a set of well-defined

rules whose existence is a fundamental prerequisite for sustaining any form of life. It is at this point that the genetic code enters the picture.

The genetic code — fully deciphered by 1966 and shown in Table 1 at the level of mRNA codons versus amino acids, as usual — shows some remarkable features. Not only is it highly degenerate, but the distribution of codons representing the same amino acid shows certain regularities, the most prominent one being the weak dependence of the meaning of a codon on the third base. This observation led Crick to formulate the *wobble hypothesis*,[13] according to which the pairing between the third base of a codon and the first base in the corresponding anticodon does not necessarily obey the strict Watson–Crick pairing rules: other pairings are possible, the first and most prominent example being the wobble pairing between G and U (see Fig. 9). This allows, e.g. an anticodon with G in the first position to simultaneously recognize codons with U and with C in the third position, so that the appearance of a tRNA with such an anticodon forces the two codons to have the same meaning. Soon after, it was found that anticodons often contain unusual bases in the first position, such as inosine, which allow for other unconventional pairing rules with the third codon base. There is now an extensive list of wobble rules and of codon-anticodon correspondences in many different kinds of organisms; see Ref. 19 for a recent comprehensive review.

The attempt to explain the degeneracy of the genetic code in this way, on purely biological grounds, has one major drawback: it provides no explanation for the fact that, despite the wealth of diversification observed between different species in the details of the translation process, in particular with regard to the great variety of anticodons and, more generally, of tRNA molecules, the genetic code itself is almost universal. Indeed, the standard genetic code presented in Table 1 was during the first decade after its discovery believed to be strictly universal, and even though we now know that it is not, the deviations found in the form of non-standard codes are small: in each case, the modification affects only a small number of codon-amino acid assignments and applies to a very restricted class of species or to the codes of organelles such as mitochondria and chloroplasts; see Ref. 19. The argument normally employed by biologists and geneticists in this context is the one first put forward by Crick[20] when formulating his famous "frozen accident" hypothesis, according to which the genetic code, after going through a primordial phase of evolution, was at a certain stage frozen into its presently observed form, namely when the protein synthesis machinery in organisms had become so complex that further changes would have become lethal; universality would then be a consequence of the fact that this freezing occurred very early in evolution, even before the bifurcation of life forms into different kingdoms. The analysis of non-standard codes and their origin performed in Ref. 19 could then be interpreted as evidence that the freezing is not complete: some kind of melting can occasionally occur. But even if one is willing to accept all these arguments, the simple statement that the genetic code has been frozen at some stage of its evolution does not provide any hint as to what are the laws that governed that evolution before the freezing occurred.

The "frozen accident" hypothesis in its extreme form states that this primordial evolution was entirely a matter of "chance". A simple statistical argument first worked out by Bertman and Jungk[21] shows, however, that the number of possible genetic codes is of the order of 10^{71}, most of which do not exhibit any kind of regular structure at all, so that the appearance of a genetic code with marked regularities is extremely unlikely.

In view of these arguments, it is a challenge to identify the laws that have governed the early evolution of the genetic code.

The algebraic approach to the genetic code addresses exactly this problem, based on the idea that the observed degeneracy of the genetic code is a reflection of a primordial symmetry which in the course of the evolution of the genetic code was broken in a sequence of steps. One of the main advantages of this approach is that requirements of compatibility with some symmetry cut down radically on the number of possibilities mentioned above, leading to a non-negligible probability for the present genetic code to be just the way it is. In this sense, the algebraic approach is compatible with the idea of freezing.

Finally, we would like to point out that the decision of applying group theoretical techniques to analyze the degeneracy of the genetic code is based on the experience accumulated in physics, where these techniques are useful for analyzing a large variety of phenomena, ranging from particle physics to molecular vibrations.

5. The Use of Symmetries in Physics

The concept of symmetry has been one the most important guiding principles for the development of theoretical physics in the 20th century and has been used in a wide variety of contexts. Its many variants can partly, and very roughly, be classified in terms of the following contrasting notions.

- Spatial or space-time symmetries/internal symmetries.
  Symmetries that are realized through transformations such as translations, rotations and reflections in three-dimensional physical space or, in the case of relativistic physics where space and time merge into a single space-time continuum, through transformations in four-dimensional space-time, are usually referred to as spatial symmetries or as space-time symmetries, respectively. In contrast to these, internal symmetries are realized through transformations in an abstract internal space that has nothing to do with physical space or space-time, being instead related to the dynamical variables of the theory under consideration.
- Continuous symmetries/discrete symmetries.
  Symmetries can be distinguished according to whether the corresponding group of transformations is continuous, such as the group of spatial translations (parametrized by vectors) or rotations (parametrized, e.g. by the Euler angles), or is discrete, such as the group of spatial translations or rotations in a crystal lattice or a reflection group. The simplest reflection group consists of just two elements, the identity and a nontrivial reflection (i.e. a nontrivial transformation

which, when applied twice, gives back the identity): a prominent example is chiral symmetry.

- Local symmetries/global symmetries.
  A very important extension of the usual symmetry concept arises from the idea that the parameters characterizing a specific element within a given symmetry group may depend on the point in space or space-time where the symmetry transformation is performed. This notion of "gauging a symmetry" by allowing the symmetry operations to be defined locally rather than globally, in the sense that different transformations may be performed at different points, is at the heart of *gauge theories*, which occupy a central position in modern (classical and quantum) field theory. However, gauge transformations do not really represent symmetry transformations in the strict sense of the word, since they do not relate observable quantities. Instead, their presence reflects the fact that the system under consideration is being described in terms of redundant, non-observable quantities (such as the vector potentials in electrodynamics), and the amount of redundancy in the choice of these variables is controlled by the principle of gauge invariance: the observable, physical content of the theory is precisely its gauge invariant part. Therefore, our use of the term "symmetry" in this paper will always refer to global symmetries, not to local ones.
- Broken symmetries/exact symmetries.
  One speaks of a broken symmetry when a symmetry is only approximate, that is, there is a deviation from the exact symmetry which is however sufficiently small for it to remain clearly perceptible.
- Hidden symmetries/manifest symmetries.
  Many important physical systems possess dynamical symmetries which cannot be inferred directly from the original formulation of the theory: they are hidden, rather than manifest. Hidden symmetries are a typical feature of *integrable systems* and can be viewed as the real reason behind their integrability. Of course, when passing to a different description, e.g. by a judiciously chosen transformation to new dynamical variables, a hidden symmetry may become manifest. Conversely, the identification of a hidden symmetry is in many cases the decisive step towards finding such a transformation which, in turn, often allows to determine the exact solution of the theory. The importance and usefulness of this point of view becomes apparent when one realizes that *all* dynamical systems in physics known to admit an exact solution are examples of integrable systems which owe their integrability to the presence of some hidden symmetry!

Historically, the systematic use of symmetry principles in modern physics in the 20th century begins with Einstein's theory of relativity. Special relativity is based on a judicious analysis of the symmetry that underlies the different possible choices of inertial frames (Galilei versus Lorentz invariance). Similarly, the starting point for general relativity was Einstein's attempt to extend this line of reasoning to include non-inertial frames, together with his insight that this naturally brings gravity into

the picture, according to the equivalence principle. In a different way, symmetry arguments play an important role in general relativity up to the present day, due to the fact that all known exact solutions of Einstein's equations admit nontrivial groups of isometries: these are crucial for being able to reduce Einstein's equations to a system of equations that can be solved explicitly.

Notwithstanding the importance of symmetry considerations in relativity, the decisive impetus for their systematic use in physics, as well as for the further development of the underlying mathematical machinery, came with the advent of quantum theory. Apart from the immediate application of group-theoretical techniques to quantum mechanical problems, systematically exposed in textbooks by van der Waerden,[22] Weyl[23,24] and Wigner,[25] a first milestone of theoretical nature was Wigner's classification of elementary particles in terms of irreducible unitary representations of the Poincaré group.[26] An equally and perhaps even more important line of development is of phenomenological nature, using group theory to classify exact or approximate degeneracies between experimentally observed states. This approach originates with Heisenberg's postulate of SU(2) isospin invariance of the nuclear forces, which some 30 years later was extended by Gell-Mann and Ne'eman to the postulate of SU(3) flavor invariance of the strong interactions ("eightfold way"). Since then, group-theoretical techniques have become an indispensable tool in the study of the fundamental constituents of matter and of their interactions. More recently, Iachello has in the same spirit initiated the application of group-theoretical techniques to problems in nuclear physics (interacting boson model and interacting boson-fermion model) and in molecular physics (vibron model).

A common feature of these "phenomenological symmetries" is that they are broken (rather than exact) and to a large extent hidden (rather than manifest), as well as internal (rather than spatial or space-time).

The history of hidden symmetries can be traced back to the discovery of the dynamical SO(4) symmetry in the Kepler or Coulomb problem, originally found in 1926 by Pauli and used to explain the additional degeneracy of the energy levels of the hydrogen atom with total angular momentum[27]: it extends the standard SO(3) symmetry due to rotational invariance of the Coulomb potential — a manifest spatial symmetry of the same nature as the manifest space-time symmetries encountered in relativity — in an unexpected way. The idea can be most conveniently explained within the Hamiltonian formulation of mechanics. For the motion of a point particle in an arbitrary central potential $V(r)$, rotational invariance of the potential can be expressed by the statement that the Hamiltonian H of the theory commutes with the generators of the rotation group SO(3), which are just the three components of angular momentum L. In classical mechanics, conservation of L forces the motion to be confined to the plane orthogonal to L and to satisfy Kepler's second law (the area law), whereas in quantum mechanics, it implies degeneracy of the energy levels with the magnetic quantum number m. What singles out the Coulomb potential $1/r$ is the presence of an additional hidden symmetry, in the sense that in this — and only this — case, the Hamiltonian H commutes with

three additional quantities, namely the three components of the so-called Runge–Lenz vector M. Geometrically, this vector lies in the plane orthogonal to the angular momentum vector, whereas algebraically, the two vectors — after a suitable (energy-dependent) renormalization of the latter — generate the group SO(4). In classical mechanics (Kepler problem), conservation of M (in addition to that of L) means that the bounded trajectories of the particle are closed, periodic curves (ellipses), with the Runge–Lenz vector pointing towards the pericenter (point of minimal distance to the center), whereas in quantum mechanics, it implies degeneracy of the bound state energy levels not only with the magnetic quantum number m but also with the angular momentum quantum number l.

It is also interesting to observe what happens when the dynamical SO(4) symmetry in the Kepler or Coulomb problem is slightly broken by a small perturbation of the potential, such as in the effective potential for the geodesic motion of a point particle in the Schwarzschild solution (which contains a $1/r^3$ contribution) or for the dynamics of the valence electron of an alkali atom. The results is that in classical mechanics, the trajectories are no longer closed or periodic, whereas in quantum mechanics, the additional degeneracy of the energy levels with l is removed: however, the deviation from closed periodic orbits or from degenerate energy levels is small to the extent that the perturbation is weak. (Thus perihelion rotation in celestial mechanics is really a symmetry breaking phenomenon!) A similar and even more well known phenomenon occurs when the rotational SO(3) symmetry is broken by angular dependent terms, such as in the effective potential for the geodesic motion of a point particle in the Kerr solution or for the dynamics of the valence electron of a hydrogen or alkali atom in the presence of an external electric or magnetic field. In this case, the classical trajectories are no longer planar, whereas the quantum energy levels cease to be degenerate with m, leading to the well known splitting of spectral lines observed in the Stark effect and Zeeman effect.

The first "phenomenological symmetry" to appear in physics was isospin. Mathematically, the SU(2) *isospin* group is identical with the universal covering group of the usual rotation group SO(3), known as the SU(2) *spin* group, but unlike the latter, it represents an internal symmetry, with an entirely different physical interpretation: it expresses the principle of *isospin invariance* of the strong interactions, according to which all hadrons (strongly interacting particles) are organized into isospin multiplets and all members of the same isospin multiplet participate in strong interactions in exactly the same way. The origin of this principle was the observation that nuclear forces are charge independent, i.e. the experimental fact that after subtracting the contribution from the Coulomb interaction between protons, the forces between two protons, between a proton and a neutron and between two neutrons are practically the same: this led Heisenberg in 1932 to consider proton and neutron, with respect to nuclear forces, as merely two different states of one and the same particle, for which the term nucleon was coined. After Yukawa had in 1935 postulated the existence of mesons, or more precisely of the charged pions π^+

and π^-, as mediators of the nuclear forces, and after Kummer had in 1938 come to the conclusion that there must also be a neutral pion π^0 in order for the pions to form an isospin triplet, interacting with the nucleons which form an isospin doublet, there emerged the first model for the strong interactions, based on isospin invariance. Of course, isospin invariance is not exact, but is broken by electromagnetic as well as weak interactions.

Unfortunately, it soon became evident that this first model for the strong interactions was far from complete, due to the existence of many other hadrons (strongly interacting particles), which started to appear in the late 1940's and were found and identified in great numbers in the 1950's, after the first particle accelerators had gone into operation. Most hadrons are extremely short-lived resonances, with lifetimes of the order of 10^{-24} s, but some are much more stable, with lifetimes typically of the order of 10^{-10} to 10^{-11} s: this led physicists to postulate that these could, in particle-antiparticle pairs, be produced through strong interactions but could not, by themselves, decay through strong interactions. The postulate was formalized by the introduction of a new quantum number called strangeness, conserved in strong interactions and electromagnetic interactions but not in weak interactions, and by attributing to each hadron its proper value of strangeness.

The breakthrough came in 1961 when Gell-Mann and Ne'eman realized that not only do hadrons of the same spin and the same parity appear in isospin multiplets, but that in a two-dimensional diagram with the third component of isospin as the first coordinate and strangeness as the second coordinate, these isospin multiplets combine to form multiplets under a larger group now known in physics as the SU(3) *flavor* group. Thus the principle of isospin invariance was extended to the principle of *flavor invariance* of the strong interactions, according to which all hadrons are organized into multiplets under the flavor group (such multiplets are in group theory known as weight diagrams) and all members of the same multiplet participate in strong interactions in exactly the same way. Moreover, the approach offered a natural explanation for the observed particle spectrum by postulating the existence of sub-nuclear particles, called *quarks*, coming in three species (flavors) — up (u), down (d) and strange (s) — such that all baryons (hadrons of half-integer spin) are bound states of three quark and all mesons (hadrons of integer spin) are bound states of a quark and an antiquark. The strong interactions between hadrons are from this point of view simply the residue of the strong interactions between quarks (just as the van der Waals forces between atoms or molecules are the residue of the Coulomb interactions between the nuclei and electrons involved) and these strong interactions are flavor blind. Just like isospin invariance, flavor invariance is however not exact, being broken by electromagnetic as well as weak interactions. An even more important contribution to symmetry breaking comes from the mass difference between the quarks: in a good approximation, we have $m_u = m_d \ll m_s$, which implies the Gell-Mann–Okubo mass formula.

The main problem of the *quark model* is, of course, that the quarks themselves are not observed in nature, at least not in the form of freely propagating particles: they appear to be permanently confined within the hadrons. Finding a convincing mechanism to explain this phenomenon of quark confinement, within the context of the standard model of elementary particle physics, is still one of the big open problems of theoretical physics. Nevertheless, the quark model itself is now firmly established as one of the great triumphs of physics in this century and has definitely promoted group theory to the status of a basic tool of science: it was able to provide an organizing principle for the "particle zoo" and, at the same time, unveil the organization of matter at the sub-nuclear level.[28]

The success of the quark model has stimulated the development of other models based on group-theoretical methods, such as the *interacting boson model* (IBM)[29] in nuclear physics and the *vibron model*[30] in molecular physics. The basic idea of these *algebraic models* is to generate spectra and transition amplitudes for nuclear and molecular systems, respectively, by the use of Lie algebras — more precisely, of unitary Lie algebras and their irreducible representations by totally symmetric tensors, which are a natural offspring of a description of effective physical degrees of freedom in terms of *boson operators*.

In general terms, the main difficulty encountered in theoretical nuclear physics as well as theoretical molecular physics is that one is almost always dealing with complex systems, with a large number of degrees of freedom. In the case of nuclear physics, the situation is aggravated by the fact that not even the basic laws of interaction between nucleons are fully understood; in particular, their derivation from quantum chromodynamics remains a challenge. But even in molecular physics, where an approach based on first principles — namely the Schrödinger equation for the system of nuclei and electrons of which the molecule under consideration is composed — is possible, such "*ab initio*" calculations often turn out to be unpractical or unsatisfactory, due to the errors implied by the approximations that have to be performed. Therefore, it is usually a more successful strategy to formulate a phenomenological model in terms of effective degrees of freedom. The algebraic approach provides a strategy for devising such phenomenological models.

The starting point of the algebraic approach is to associate the quantum states of the system with the vectors in an irreducible representation of some spectrum generating Lie algebra $\mathfrak{g}$. This means that the algebraic Hamiltonian is to be constructed as a polynomial function in the generators of $\mathfrak{g}$. The ultimate criterion for the selection of both the spectrum generating Lie algebra and the algebraic Hamiltonian is of course agreement with experiment. However, in view of the overwhelmingly large number of possibilities, physical considerations must be used as guiding principles. For example, the most general algebraic Hamiltonian, being an arbitrary polynomial

$$H = H_0 + \sum_i \epsilon^i T_i + \sum_{i,j} \epsilon^{ij} T_i T_j + \cdots + \sum_{i_1,\ldots,i_p} \epsilon^{i_1 \ldots i_p} T_{i_1} \cdots T_{i_p} + \cdots , \qquad (1)$$

in the generators T_i of $\mathfrak{g}$, contains an enormous (in fact, arbitrarily large) number of free parameters and therefore has little predictive power. Among these, there is a set of special operators, namely the Casimir operators for subalgebras of the given spectrum generating algebra, from which one wishes to construct the Hamiltonian by linear combination. To guarantee that the various Casimir operators appearing in such a linear combination commute among themselves, so that they can be simultaneously diagonalized and their simultaneous eigenvalues may be used to label the states, one uses chains of subalgebras

$$\mathfrak{g} \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_k \,. \tag{2}$$

Moreover, it is required that all admissible chains terminate in a copy of the Lie algebra $\mathfrak{so}(3)$ of the rotation group SO(3), which is therefore an invariance group of the Hamiltonian, followed by a copy of the Lie algebra $\mathfrak{so}(2)$ of the rotation group SO(2) around an arbitrarily chosen fixed axis to guarantee that the standard angular momentum quantum numbers L (referring to $\mathfrak{so}(3)$) and M (referring to $\mathfrak{so}(2)$) appear among the state labels. If the Hamiltonian is constructed as a linear combination of Casimir operators within a single such chain, it is referred to as a *dynamical symmetry Hamiltonian*. In this case, one obtains a closed formula for the energy spectrum in terms of the representation labels. Otherwise, one may start from such a dynamical symmetry Hamiltonian as a first approximation and try to improve the agreement with the experimental data by including other, non-invariant terms; a particular example are the Majorana terms arising from Casimir operators of a different chain.

Another guiding principle is the construction of the Lie algebra $\mathfrak{g}$ in terms of boson operators, which leads to the Lie algebra $\mathfrak{u}(r)$: given r operators b_j and their hermitean adjoints $b_j^\dagger$ satisfying canonical commutation relations

$$[b_j, b_k] = 0 \,, \qquad [b_j, b_k^\dagger] = \delta_{jk} \,, \qquad [b_j^\dagger, b_k^\dagger] = 0 \,, \tag{3}$$

the generators

$$X_{jk} = b_j^\dagger b_k \,, \tag{4}$$

satisfy the commutation relations of the Lie algebra $\mathfrak{u}(r)$:

$$[X_{jk}, X_{lm}] = \delta_{kl} X_{jm} - \delta_{jm} X_{kl} \,. \tag{5}$$

These generators act on the Fock space obtained by applying products of the creation operators $b_j^\dagger$ to the ground state $|0\rangle$, and the "N-particle subspace" spanned by vectors of the form

$$b_{j_1}^\dagger \cdots b_{j_N}^\dagger |0\rangle \tag{6}$$

carries the totally symmetric representation of order N, of highest weight $(N, 0, \ldots, 0)$, of the Lie algebra $\mathfrak{u}(r)$ spanned by the generators X_{jk}. In order to guarantee rotational invariance of the entire setup, it is furthermore imposed that

the boson operators transform according to a representation of the ordinary rotation group SO(3), which can be implemented by using a double index notation $j = (l, m)$, with $m = -l, \ldots, l$. Thus in general, we encounter a single scalar boson operator $b_{0,0}$, a triple of vector boson operators $b_{1,-1}$, $b_{1,0}$, $b_{1,1}$, a quintuplet of tensor boson operators $b_{2,-2}$, $b_{2,-1}$, $b_{2,0}$, $b_{2,1}$, $b_{2,2}$ and so on, plus their hermitean adjoints. In nuclear physics, the appearance of quadrupole deformations has suggested the use of five tensor boson operators $b_{2,-2}$, $b_{2,-1}$, $b_{2,0}$, $b_{2,1}$, $b_{2,2}$ which, together with a scalar boson operator $b_{0,0}$, give rise to $\mathfrak{u}(6)$ as the basic Lie algebra of the interacting boson model. In molecular physics, the predominant dipolar nature of covalent chemical bonds has motivated the use of three vector boson operators $b_{1,-1}$, $b_{1,0}$, $b_{1,1}$ which, together with a scalar boson operator $b_{0,0}$, give rise to $\mathfrak{u}(4)$ as the basic Lie algebra of the vibron model.

The motivation for the use of boson operators to describe the effective physical degrees of freedom, in the interacting boson model[29] introduced by Feshbach, Iachello and Arima[31–34] as well as in the vibron model[30] introduced by Iachello and Levine,[35,36] is to be sought in the physical picture underlying the former, which is closely analogous to that behind the BCS theory of superconductivity. It starts out from the hypothesis that, in nuclei with an even total number of nucleons, nuclear interactions between nucleons cause these to form pairs which, being bosons, may under appropriate circumstances condense into a common ground state — just as in certain metals, phonon interactions between electrons in the electron gas cause these to form Cooper pairs which, being bosons, will at sufficiently low temperatures condense into the superconducting ground state; excited states may then be described as the result of applying boson creation operators to the ground state. Similarly, it is tempting to consider the covalent chemical bonds between atoms, which are formed by (one or several) electron pairs, as bosonic degrees of freedom and to describe excited states (in particular, vibrational states) of molecules as the result of applying boson operators to the ground state. Of course, this kind of reasoning is rather vague and it is clear that more work will be needed in order to gain a better understanding of the microscopic foundations of the resulting models. For the time being, the main argument for taking them seriously is their excellent agreement with experiment.

As the simplest example from molecular physics, let us consider the case of diatomic molecules. The spectrum of energy levels for such a molecule is described by two quantum numbers, a vibrational quantum number ν and an angular momentum quantum number L, with energy levels belonging to the same value of ν gathered into bands whose "head" is composed of a large number of rotational lines differing by only a few cm^{-1}. In a crude first approximation, these levels are organized harmonically: this is the reason why excited states in molecular physics are often referred to as overtones. Closer examination, of course, reveals deviations from the harmonic predictions; these anharmonicities increase as the energy grows towards the dissociation limit. One possibility to incorporate such deviations is to use Hamiltonians based on more sophisticated interatomic potentials, for example

potentials derived from a Taylor expansion around the equilibrium configuration beyond quadratic order or the intrinsically nonlinear Morse potential which, to a good approximation, describes the vibrational spectrum of fluoride acid HF. An alternative method is the phenomenological description of rotation-vibration spectra given by the celebrated Dunham expansion, which in the simplest case of a diatomic molecule reads

$$E(n,L) = \sum_{i,k} \Delta_{ik} \left(\nu + \frac{1}{2}\right)^i \left[L(L+1)\right]^k , \tag{7}$$

where the coefficients Δ_{ik} are obtained by fitting to the experimentally observed spectrum. However, this approach provides no information about wave functions or transition amplitudes.

In the vibron model, the strategy described above, applied to diatomic molecules, leads to consider the space of relevant quantum states as an irreducible representation of the Lie algebra $\mathfrak{u}(4)$ and to construct the Hamiltonian from Casimir operators for subalgebras contained in descending chains, starting with the entire Lie algebra $\mathfrak{u}(4)$ and terminating in the Lie algebra $\mathfrak{so}(3)$ of the rotation group SO(3) (followed by a copy of the Lie algebra $\mathfrak{so}(2)$ of the rotation group SO(2) around an arbitrarily chosen fixed axis, as mentioned before). There are two such chains. Written together with the group theoretical labels for the irreducible representations of each subalgebra, they are

$$\begin{array}{ccccccc} \mathfrak{u}(4) & \supset & \mathfrak{u}(3) & \supset & \mathfrak{so}(3) & \supset & \mathfrak{so}(2) \\ N & & n & & L & & M \end{array} \tag{8}$$

and

$$\begin{array}{ccccccc} \mathfrak{u}(4) & \supset & \mathfrak{so}(4) & \supset & \mathfrak{so}(3) & \supset & \mathfrak{so}(2) \\ N & & \omega & & L & & M \end{array} \tag{9}$$

Here, N is the tensorial degree of the totally symmetric representation of $\mathfrak{u}(4)$ to be used: it is a characteristic of the chemical bond and hence of the molecule to be described. Moreover, n runs from 0 to N, whereas ω runs in steps of 2 from 0 (for N even) or 1 (for N odd) up to N, while L and M are the standard angular momentum quantum numbers. The corresponding dynamical symmetry Hamiltonians are

$$H = H_0 + \alpha C_2(\mathfrak{u}(3)) + \beta C_1(\mathfrak{u}(3)) + \gamma C_2(\mathfrak{so}(3)) \tag{10}$$

for the first chain and

$$H = H_0 + \alpha C_2(\mathfrak{so}(4)) + \beta C_2(\mathfrak{so}(3)) \tag{11}$$

for the second chain. Evaluating the Casimir operators leads to a closed formula for the energy of each state in terms of the quantum numbers introduced above, namely

$$E = E_0 + \alpha n(n-3) + \beta n + \gamma L(L+1) \tag{12}$$

for the first chain and

$$E = E_0 + \alpha\omega(\omega + 2) + \beta L(L + 1) \tag{13}$$

for the second chain, which turns out to be the relevant one for the physical applications. It should be noted that the Lie algebra $\mathfrak{so}(4)$, being isomorphic to the direct sum of two copies of the simple Lie algebra $\mathfrak{so}(3) \simeq \mathfrak{su}(2)$, has two independent quadratic Casimir operators, namely the sum $L_1^2 + L_2^2$ and the difference $L_1^2 - L_2^2$ of the quadratic Casimir operators of its two constituents. However, only the first of them contributes to the Hamiltonian, since the second vanishes on totally symmetric representations. Moreover, there is a simple relation between the group theoretical label ω ($\omega = l_1 + l_2$) and the vibrational quantum number ν, namely

$$\nu = \frac{1}{2}(N - \omega)\,, \tag{14}$$

implying that the energy E depends quadratically on ν. The free parameters N, $E_0 = E_0(N)$, α and β are determined by fitting to the experimentally observed spectrum, minimizing the least square deviation. A schematic spectrum resulting from this procedure is shown in Fig. 11.

Essentially the same procedure can be applied to polyatomic molecules. Here, one associates one copy of the Lie algebra $\mathfrak{u}(4)$ to each chemical bond: the spectrum

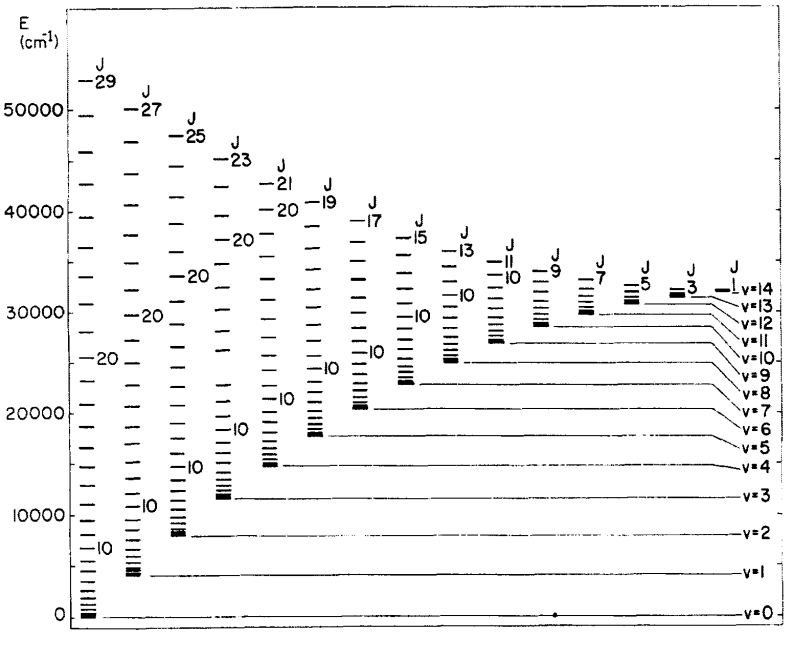
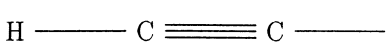


Fig. 11. Schematic spectrum for the $\mathfrak{so}(4)$ chain in the vibron model.

generating algebra is then the direct sum of these Lie algebras. As an illustrative example, consider monofluoracetylene:



(15)

The relevant symmetry chain, together with the appropriate group theoretical labels, is

$$\begin{array}{ccccccccc} \mathfrak{u}_1(4)\oplus\mathfrak{u}_2(4)\oplus\mathfrak{u}_3(4) & \supset & \mathfrak{so}_1(4)\oplus\mathfrak{so}_2(4)\oplus\mathfrak{so}_3(4) & & & & & & \\ (N_1,N_2,N_3) & & (\omega_1,\omega_2,\omega_3) & & & & & & \\ & \supset & \mathfrak{so}_{13}(4)\oplus\mathfrak{so}_2(4) & \supset & \mathfrak{so}_{132}(4) & \supset & \mathfrak{so}_{132}(3) & \supset & \mathfrak{so}_{132}(2) \\ & & (\eta_1,\eta_2) & & (\tau_1,\tau_2) & & L & & M \end{array} \tag{16}$$

and the corresponding dynamical symmetry Hamiltonian becomes

$$\begin{aligned} H = {} & H_0 + \alpha_1 C_2(\mathfrak{so}_1(4)) + \alpha_2 C_2(\mathfrak{so}_2(4)) + \alpha_3 C_2(\mathfrak{so}_3(4)) \\ & + \alpha_{13} C_2(\mathfrak{so}_{13}(4)) + \alpha_{132} C_2(\mathfrak{so}_{132}(4)) + \beta C_2(\mathfrak{so}_{132}(3))\,. \end{aligned} \tag{17}$$

Evaluating the Casimir operators leads to a closed formula for the energy of each state in terms of the quantum numbers introduced above:

$$\begin{aligned} E = {} & E_0 + \alpha_1\omega_1(\omega_1+2) + \alpha_2\omega_2(\omega_2+2) + \alpha_3\omega_3(\omega_3+2) \\ & + \alpha_{13}[(\eta_1+1)^2+\eta_2^2] + \alpha_{132}[(\tau_1+1)^2+\tau_2^2] + \beta L(L+1)\,. \end{aligned} \tag{18}$$

As before, only one of the two independent quadratic Casimir operators of each of the $\mathfrak{so}(4)$ Lie algebras has been included. The others would give contributions corresponding to unphysical couplings between the modes, whose coefficients must therefore be equal to zero to a very good approximation. The remaining free parameters N_1, N_2, N_3, $E_0 = E_0(N_1,N_2,N_3)$, α_1, α_2, α_3, α_{13}, α_{132} and β are, once again, determined by fitting to the experimentally observed spectrum, minimizing the least square deviation. For a comparison between the resulting theoretically predicted spectrum of monofluoracetylene and the experimentally observed one, based on 173 lines, see Ref. 37.

Compared with other models of molecular physics, the vibron model is characterized by a significant reduction in the number of free parameters, together with an unrivaled agreement with experiment.

Summarizing the discussion, we may say that the application of symmetry principles constitutes one of the most important and successful tools of quantum physics. One fundamental lesson to be drawn is that *degeneracy results from and is a signal of symmetry*, whereas the *removal of degeneracy results from and is a signal of symmetry breaking*. As it stands, this statement is a general postulate

of quantum physics, but it certainly has a much wider range of applicability. In particular, the algebraic approach to the genetic code proposed in Refs. 38 and 39 (see also Refs. 40 and 41) is based on the hypothesis that the same postulate may be applied to problems in biology and that it is legitimate to postpone the task of providing its microscopic foundation, based on biological and/or chemical considerations, to a posterior stage of the analysis. In this sense, the algebraic approach to the genetic code has been inspired by the procedure adopted in the vibron model. One of the major questions posed by such a transfer from physics to biology is what biological interpretation — beyond that of providing the correct degeneracies — should be attributed to the analogue of the Hamiltonian operator, whose role in physical models is as obvious as it is fundamental. This remains a challenge.

6. The Mathematical Theory of Symmetries

The composition of complicated objects out of just a few basic and relatively simple building blocks is a common feature in many areas of science. All the enormous variety of matter observed in nature appears as the result of putting together atoms, which in turn are composed of protons, neutrons and electrons. In particle physics, the full spectrum of "elementary" particles is obtained by forming bound states of a few fundamental particles (quarks and leptons). In biology, proteins are assembled from 20 fundamental amino acids, DNA from 4 fundamental nucleic bases and so on.

The same kind of phenomenon can be observed in mathematics, where the structure of the basic building blocks is determined not by experimental information, but by classification theorems. A simple example is the notion of number. The idea can be made precise through the notion of a number field or, somewhat more generally, of a division algebra: one of the classification theorems in this area of algebra states that division algebras over the reals exist only in one, two and four dimensions, corresponding to the real, complex and quaternionic numbers, respectively. (If one gives up the requirement of associativity, there appears the additional possibility of a division algebra in eight dimensions, corresponding to the octonionic numbers.) Therefore, any attempt to construct number fields (even non-commutative or non-associative) in other dimensions is doomed to failure. One might say that classification theorems in mathematics are mandatory in the sense of explicitly exhibiting all possibilities of realizing certain ideas.

Other important examples of such theorems are the classification of simple Lie algebras due to Élie Cartan, completed in 1914, and the more recently achieved classification of simple finite groups.

The notion of a group is the simplest mathematical concept to realize symmetries. By definition, a *group* is just a set G in which there is defined a binary operation, called the *product* in G, with the following properties: (a) it is *associative*, (b) there exists a distinguished element e in G, called the *unit*, which acts as

the identity under multiplication (from either side) and (c) each element g in G has a unique (two-sided) *inverse* g^{-1} in G. This purely algebraic definition is sufficient for handling finite groups, whereas for infinite groups it must be supplemented by additional hypotheses, mostly of topological or analytical nature. For example, if we impose that the set G, apart from being a group, also carries the structure of a *topological space* and require the product to be a continuous map from $G \times G$ to G and the inversion to be a continuous map from G to G, we obtain the notion of a *topological group*. (Very roughly speaking, a topological space is a set in which it makes sense to speak of open and closed subsets, of neighborhoods and of properties such as connectedness and compactness.) Similarly, if we impose that the set G, apart from being a group, also carries the structure of a smooth *manifold* and require the product to be a smooth map from $G \times G$ to G and the inversion to be a smooth map from G to G, we arrive at the notion of a *Lie group*. (Intuitively, a smooth manifold is a set which can locally be parametrized by coordinates such that coordinate changes are smooth, and the smoothness conditions on multiplication and inversion mean that the coordinates of the product of two elements and of the inverse of an element should be smooth functions of the coordinates of the original elements.)

The prototypes of Lie groups are the set $\mathrm{GL}(n, \mathbb{R})$ of all invertible real $(n \times n)$-matrices and the set $\mathrm{GL}(n, \mathbb{C})$ of all invertible complex $(n \times n)$-matrices, in which the product is simply given by matrix multiplication.

A second concept of central importance in the mathematical formulation of symmetries is that of a *representation* of a group G on a vector space V, which is by definition a homomorphism T from G into the group of invertible linear transformations on V, or in other words, a prescription which assigns to every element g of G an invertible linear transformation $T(g)$ on V in such a way that T takes the product of two elements in G to the product of the corresponding linear transformations. We shall be dealing exclusively with finite-dimensional representations, so that choosing a basis of V, we may represent linear transformations by square matrices, and the product of linear transformations corresponds to the product of matrices. A group will in general have many different representations, which is one of the reasons why it has become customary in mathematics to distinguish between abstract groups and concrete representations.

The elementary building blocks for group representations (and similarly for representations of other mathematical entities such as algebras) are the irreducible representations. Observe that two representations T_1 and T_2 of the same group G on vector spaces V_1 and V_2 can be joined together to form a representation T_3 of G on a vector space V_3 called their *direct sum* (one reason for this terminology being that the dimension of V_3 is the sum of the dimensions of V_1 and of V_2), where for each g in G, the matrix of $T_3(g)$ is block diagonal, formed by inserting the matrices of $T_1(g)$ and of $T_2(g)$ along the diagonal: then the spaces V_1 and V_2 appear as nontrivial subspaces of V_3 which are invariant under the action of G. Representations formed

by this process are composite, or in mathematical terms, reducible. More precisely, a representation is called *reducible* if its carrier space contains some nontrivial invariant subspace, *irreducible* if it does not and *completely reducible* if it can be reduced to a direct sum of irreducible representations.

Returning to the question of how to encode the concept of symmetry in mathematical language, we must also take into account that many symmetries observed in nature are not exact: most of them are broken. Generically, we speak of a broken symmetry when a symmetry is only approximate, that is, there occurs a deviation from the exact symmetry which is however sufficiently small for it to remain clearly perceptible. In fact, the process of *symmetry breaking* is a phenomenon pervading all areas of science, a typical example being the breaking of *chiral symmetry* which plays an equally important role in physics, chemistry and biology.

With the above mathematical terminology at our disposal, we are able to formulate these intuitive notions in precise mathematical terms.

- An *exact symmetry* is described abstractly by a group G and realized concretely by a set of matrices which, taken together, form a representation of G on some finite-dimensional vector space V, chosen according to the application we have in mind. For convenience, we shall assume in the following that this representation is irreducible, or in the terminology used in many areas of science other than mathematics (such as physics or chemistry), that the vectors of some (arbitrarily chosen) basis of V form a single *multiplet* under G. (Otherwise, the representation must be split into its irreducible constituents, which leads to an entire set of multiplets.)
- A *broken symmetry* is described by fixing, in addition, a subgroup H of G representing the *residual symmetry*, i.e. that part of the symmetry which has remained intact during the breaking. (As a pictorial example, we show in Fig. 12 a disk in the plane, which has an obvious symmetry under the rotation group SO(2) in two-dimensions: in (a) the symmetry is completely broken by chopping off a randomly chosen piece of the disk, so the residual symmetry group is reduced to the trivial subgroup, whereas in (b) the symmetry is only partly broken to produce a regular hexagon, so the residual symmetry group is that of the regular hexagon, i.e. the group of rotations by angles which are multiples of 60°.) Then the given

(a) (b)

Fig. 12. Breaking the rotational symmetry of a disk; (a) without residual symmetry; (b) with hexagonal residual symmetry.

irreducible representation of G on V, when restricted to H, breaks into several irreducible representations of H on subspaces of V, that is, the single multiplet under G breaks up into several multiplets under H — a phenomenon commonly known as *branching*.

More generally, the idea that symmetry breaking often occurs in several steps, rather than in a single stroke, can be implemented by supposing that G comes with a sequence of subgroups $G_1, \ldots, G_k$ which form a descending chain

$$G \supset G_1 \supset \cdots \supset G_k \,, \tag{19}$$

leading to a whole sequence of successive branchings, where at each step an irreducible representation of the previous group breaks up into several irreducible representations of the next group in the chain.

Briefly, we may say that symmetries are associated with *degeneracies*, whereas symmetry breaking leads to the *lifting of degeneracies*.

Finally, we may inquire about the converse problem, which is the following. Given just a set of multiplets, find a group G and a descending chain of subgroups $G_1, \ldots, G_k$ such that the given set of multiplets can be arranged into an irreducible representation of that group and be reproduced by reduction through that chain of subgroups.

Examples of this "spectroscopic approach" towards the identification of symmetries are abundant in physics. In fact, in some areas such as quantum theory, whose objects are not directly observable to us, it is often the only available method to experimentally verify the presence of symmetries. To mention just one elementary example, consider the degeneracy of energy levels in a central potential due to rotational symmetry and its lifting by application of an external field, resulting in a splitting of spectral lines, such as in the Stark and Zeeman effects.

In the following, we shall present the basic results from the theory of Lie groups and Lie algebras which are pertinent to our analysis of the degeneracies of the genetic code, according to the strategy described above. Details can be found in standard textbooks, such as.[42–48]

6.1. *Lie groups, Lie algebras and the Cartan classification*

Despite the fact that Lie groups are infinite (to the extent that their elements are labeled by continuous parameters), their representation theory is accessible by means of purely algebraic techniques, just as that of discrete groups. The basic reason is that representations of Lie groups are essentially equivalent to representations of Lie algebras, which are purely algebraic entities.

A *Lie algebra* is, by definition, a vector space $\mathfrak{g}$ in which there is defined a binary operation $[.,.]$, called the *commutator* in $\mathfrak{g}$, with the following properties: (a) it is *bilinear*, (b) it is *antisymmetric*, (c) it satisfies the *Jacobi identity*

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0 \,, \qquad \text{for all } X, Y, Z \in \mathfrak{g} \,. \tag{20}$$

We shall always assume, without further mention, that Lie algebras are finite-dimensional, although the definition can of course be extended to the infinite-dimensional case and although there are important examples of such infinite-dimensional Lie algebras, such as Kac–Moody algebras and the Virasoro algebra, for which there already exists an extensive mathematical theory: they do however not play any role in our work. In terms of an arbitrary basis $\{T_1, \ldots, T_N\}$ of $\mathfrak{g}$, the commutator is given by a set of coefficients f_{ij}^k, called the *structure constants* of $\mathfrak{g}$ (with respect to the given basis), according to the formula[a]

$$[T_i, T_j] = f_{ij}^k T_k \,. \tag{21}$$

The prototypes of Lie algebras are the set $\mathfrak{gl}(n, \mathbb{R})$ of all real $(n \times n)$-matrices and the set $\mathfrak{gl}(n, \mathbb{C})$ of all complex $(n \times n)$-matrices, which under addition and multiplication by scalars form an n^2-dimensional real vector space and an n^2-dimensional complex vector space, respectively, and in which the commutator is derived from matrix multiplication by the usual formula

$$[X, Y] = XY - YX \,. \tag{22}$$

This allows us to define a (real or complex) *representation* of an arbitrary Lie algebra $\mathfrak{g}$ on a vector space V as a homomorphism t from $\mathfrak{g}$ to the Lie algebra of linear transformations on V, or in other words, a prescription which assigns to every element X of $\mathfrak{g}$ a linear transformation $t(X)$ on V in such a way that t is linear and takes the commutator of two elements in $\mathfrak{g}$ to the commutator of the corresponding linear transformations. We shall be dealing exclusively with finite-dimensional representations, so that choosing a basis of V, we may represent linear transformations by square matrices, and the commutator of linear transformations corresponds to the commutator of matrices.

An important theorem of Ado states that every Lie algebra admits (at least) one representation which is *faithful*, i.e. such that no nonzero element of $\mathfrak{g}$ is taken to the zero matrix. This means that every Lie algebra is a subalgebra of $\mathfrak{gl}(n, \mathbb{R})$ or of $\mathfrak{gl}(n, \mathbb{C})$, for suitable n, or in other words, all Lie algebras are matrix Lie algebras. However, such a realization is often impractical due to the fact that n^2 may be very large (as compared to the dimension of $\mathfrak{g}$).

Every Lie group G gives rise to a Lie algebra $\mathfrak{g}$ which, in mathematical terms, may be defined to be simply its tangent space at the unit element e. The most practical way to think about this concept is to view elements X of $\mathfrak{g}$ as formal derivatives of curves in G, taken at the point e:

$$X = \frac{d}{ds} g(s)\Big|_{s=0} \quad \text{provided } g(0) = e \,. \tag{23}$$

For matrix groups, this is more than a formal prescription, as may be demonstrated by considering a few examples. To begin with, any real or complex $(n \times n)$-matrix X can be obtained by differentiating an appropriate curve in $\mathrm{GL}(n, \mathbb{R})$ or $\mathrm{GL}(n, \mathbb{C})$,

[a] We follow Einstein's summation convention.

respectively (take, for example, $g(s) = 1 + sX$ for $|s|$ sufficiently small), which means that the Lie algebra of $\mathrm{GL}(n,\mathbb{R})$ and of $\mathrm{GL}(n,\mathbb{R})$ is $\mathfrak{gl}(n,\mathbb{R})$ and $\mathfrak{gl}(n,\mathbb{C})$, respectively. Less trivial examples are obtained by imposing constraints to define matrix Lie groups (closed subgroups of $\mathrm{GL}(n)$) and their corresponding matrix Lie algebras (Lie subalgebras of $\mathfrak{gl}(n)$). For instance, the formula

$$\frac{d}{ds}\det(g(s))\Big|_{s=0} = \mathrm{trace}\left(\frac{d}{ds}g(s)\Big|_{s=0}\right) \quad \text{provided } g(0) = e$$

shows that the Lie algebra $\mathfrak{sl}(n)$ of the group $\mathrm{SL}(n)$ of all matrices of determinant one consists of all matrices of trace zero, whereas the formulas

$$\frac{d}{ds}(g(s)^{\mathrm{T}}g(s))\Big|_{s=0} = \left(\frac{d}{ds}g(s)\Big|_{s=0}\right)^{\mathrm{T}} + \frac{d}{ds}g(s)\Big|_{s=0} \quad \text{provided } g(0) = e$$

and

$$\frac{d}{ds}(g(s)^{\dagger}g(s))\Big|_{s=0} = \left(\frac{d}{ds}g(s)\Big|_{s=0}\right)^{\dagger} + \frac{d}{ds}g(s)\Big|_{s=0} \quad \text{provided } g(0) = e$$

(where superscript T denotes transpose and dagger † denotes hermitean adjoint) show that the Lie algebra $\mathfrak{o}(n)$ of the group $\mathrm{O}(n)$ of all orthogonal matrices consists of all antisymmetric matrices[b] and the Lie algebra $\mathfrak{u}(n)$ of the group $\mathrm{U}(n)$ of all unitary matrices consists of all antihermitean matrices. For other examples, the reader is referred to the extensive list in Ref. 45, pp. 444–446.

Conversely, we may pass from the Lie algebra $\mathfrak{g}$ of a given Lie group G back to the group by the so-called *exponential map*, which is a map $\exp\colon \mathfrak{g} \to G$ essentially defined by the requirement that it takes straight lines in $\mathfrak{g}$ through the origin into *one-parameter subgroups* of G. For matrix Lie algebras and matrix Lie groups, it can be shown that this requirement is fulfilled by the usual matrix exponential, which justifies the terminology. It also explains why elements of $\mathfrak{g}$ are usually called *generators* of G: they generate the one-parameter subgroups of G, through the exponential map.

Another fundamental fact in the theory is the theorem that every Lie algebra arises from the above construction, or in other words, every Lie algebra is the Lie algebra of some Lie group. It must be emphasized, however, that the correspondence between Lie groups and Lie algebras is not one-to-one, except if one imposes additional conditions of topological nature.

The first problem is that a Lie group may split into several disconnected pieces. An example is provided by the real orthogonal group $\mathrm{O}(n)$, which splits into two connected components: the rotation group $\mathrm{SO}(n)$ of all orthogonal transformations of determinant $+1$, which is the connected component containing the unit matrix 1, and a second piece consisting of all orthogonal transformations of determinant

[b]It is customary to write $\mathrm{O}(n)$, rather than $\mathrm{O}(n,\mathbb{R})$, for the real orthogonal group, whereas the notation $\mathrm{O}(n,\mathbb{C})$ is maintained for the complex orthogonal group. An analogous convention holds for the corresponding Lie algebras.

-1, combining a rotation with a reflection. In general, when a Lie group G is not connected, the connected component of G containing the unit e is a closed normal subgroup G^0 of G: then G^0 and G have the same Lie algebra $\mathfrak{g}$, and the exponential map takes $\mathfrak{g}$ into G^0.[c] Moreover, the quotient G/G^0 is a discrete group. This implies that the representation theory of general Lie groups G falls into two substantially distinct sectors: the representation theory of connected Lie groups G^0, which can be handled by Lie algebra methods, and the representation theory of discrete groups G/G^0, which as a matter of principle is beyond the reach of such techniques.

The second problem is that even within the class of connected Lie groups, there are in general several different alternatives that lead to the same Lie algebra. The most elementary example is provided by the real line and the unit circle, whose common Lie Algebra is the one-dimensional real line, equipped with the trivial (identically vanishing) commutator, reflecting the fact that these groups are Abelian. (Incidentally, these two exhaust the class of one-dimensional Lie groups.) Among the possible alternatives, there is however a distinguished choice, since for every Lie algebra $\mathfrak{g}$, there exists a unique Lie group G with Lie algebra $\mathfrak{g}$ which, apart from being connected, is *simply connected*: it is the *universal covering group* of all other connected Lie groups with Lie algebra $\mathfrak{g}$, in the following sense: any connected Lie group $\tilde{G}$ with Lie algebra $\mathfrak{g}$ can be written in the form $\tilde{G} = G/\tilde{Z}$, where $\tilde{Z}$ is a discrete subgroup of the *center* Z of G. In essence, this means that there is a one-to-one correspondence between connected Lie groups, modulo discrete central subgroups, and Lie algebras.

As far as representations are concerned, it can be shown without much effort that the formula

$$t(X) = \frac{d}{ds} T(g(s))\Big|_{s=0} \quad \text{for} \quad X = \frac{d}{ds} g(s)\Big|_{s=0} \tag{24}$$

provides a prescription which associates to each representation T of a Lie group G a representation t of its Lie algebra $\mathfrak{g}$, on the same vector space. An even more useful formula is

$$T(\exp(sX)) = \exp(s\, t(X)) \quad \text{for all } s \in \mathbb{R}\,,\ X \in \mathfrak{g}\,. \tag{25}$$

Conversely, the fundamental theorem about the relation between representations of Lie groups and representations of Lie algebras states that every representation t of a Lie algebra $\mathfrak{g}$ can be obtained in this way from a uniquely determined representation T of the corresponding connected and simply connected Lie group G. The representations of any other connected Lie group $\tilde{G} = G/\tilde{Z}$ with Lie algebra $\mathfrak{g}$ are obtained as a subclass, because a representation of G will correspond to a representation of $\tilde{G} = G/\tilde{Z}$ if and only if it is trivial on the subgroup $\tilde{Z}$.

These facts justify the common practice in physics to use the notions of (connected) Lie groups and of Lie algebras interchangeably: it is in a sense just an

[c] In particular, the Lie algebras $\mathfrak{o}(n)$ and $\mathfrak{so}(n)$ are identical (indeed, antisymmetric matrices are automatically traceless) and it is customary to write $\mathfrak{so}(n)$ rather than $\mathfrak{so}(n)$.

abuse of language which very rarely leads to errors. A similar abuse of language is standard practice even among pure mathematicians, who call a Lie algebra $\mathfrak{g}$ *compact* if and only if at least one of its associated connected Lie groups is compact. Obviously this does not mean that $\mathfrak{g}$ is compact in the usual topological sense — being a vector space, it is not. As it turns out, a Lie algebra $\mathfrak{g}$ is compact if and only if it admits a positive definite inner product $(.,.)$ which is invariant under the action of the corresponding compact connected Lie group G on $\mathfrak{g}$ (by means of the so-called adjoint representation), or to put it differently, if

$$([Z,X],Y)+(X,[Z,Y])=0 \quad \text{for all } X,Y,Z\in\mathfrak{g}\,. \tag{26}$$

To make this clearer, let us write down the explicit definition of the adjoint representation Ad of G and the corresponding adjoint representation ad of $\mathfrak{g}$, both of which act on the Lie algebra $\mathfrak{g}$ itself as the carrier space of the representation:

$$\begin{aligned}
\mathrm{Ad}(g)X &= \frac{d}{ds}(gg(s)g^{-1})\Big|_{s=0} \quad \text{for} \quad X=\frac{d}{ds}g(s)\Big|_{s=0}\,,\\
\mathrm{ad}(Z)X &= \frac{d}{ds}\mathrm{Ad}(g(s))X\Big|_{s=0} \quad \text{for} \quad Z=\frac{d}{ds}g(s)\Big|_{s=0}\,,
\end{aligned} \tag{27}$$

so

$$\mathrm{Ad}(\exp(sZ))=\exp(s\,\mathrm{ad}(Z)) \quad \text{for all } s\in\mathbb{R},\ Z\in\mathfrak{g}\,. \tag{28}$$

Then it turns out that

$$\mathrm{ad}(Z)X=[Z,X] \quad \text{for all } Z,X\in\mathfrak{g}\,, \tag{29}$$

so the invariance condition stated above amounts to requiring the matrices $\mathrm{ad}(Z)$ to be antisymmetric and (hence) the matrices $\mathrm{Ad}(g)$ to be orthogonal.

Lie algebras fall naturally into two disjoint classes: *semisimple* Lie algebras, which can be obtained by taking direct sums of *simple* Lie algebras and *solvable* Lie algebras. A Lie algebra $\mathfrak{g}$ is called simple if it does not contain any nontrivial ideals, and it is called solvable if it admits a finite descending chain of subalgebras $\mathfrak{g}_1,\ldots,\mathfrak{g}_k$, $\mathfrak{g}\supset\mathfrak{g}_1\supset\cdots\supset\mathfrak{g}_k$, such that each subalgebra is an ideal in the preceding one and such that the quotients $\mathfrak{g}/\mathfrak{g}_1,\ldots,\mathfrak{g}_{k-1}/\mathfrak{g}_k$, as well as the last subalgebra $\mathfrak{g}_k$, are Abelian.[d] According to a theorem of Levi and Malcev, every Lie algebra $\mathfrak{g}$ decomposes into the so-called semidirect sum of a semisimple subalgebra and the *radical* of $\mathfrak{g}$, i.e. the maximal solvable ideal in $\mathfrak{g}$. Semisimple Lie algebras are characterized by the fact that their *Killing form*, i.e. the symmetric bilinear form defined by

$$\mathrm{Kill}(X,Y)=\mathrm{trace}(\mathrm{ad}(X)\ \mathrm{ad}(Y)) \quad \text{for all } X,Y\in\mathfrak{g}\,, \tag{30}$$

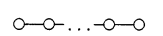
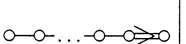
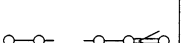
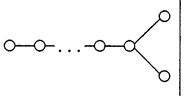
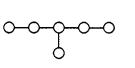
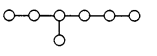
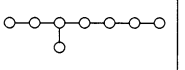
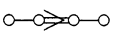
[d] An ideal $\mathfrak{i}$ in a Lie algebra $\mathfrak{g}$ is a subspace which is somewhat more restricted than a subalgebra: it is characterized by the condition that commutators between elements of $\mathfrak{i}$ and elements of $\mathfrak{g}$ (and not only elements of $\mathfrak{i}$) belong to $\mathfrak{i}$. The quotient $\mathfrak{g}/\mathfrak{i}$ of a Lie algebra by an ideal is in a natural way again a Lie algebra.

is non-degenerate: they can, as mentioned before, be decomposed into direct sums of simple Lie algebras, which have been completely classified by Cartan. Solvable Lie algebras, which contain as a particular subclass the *nilpotent* Lie algebras, are apparently much more complicated: their classification is up to the present day an open mathematical problem and we do not use them in our work.

To explain Cartan's classification of the simple Lie algebras, a word is in order about the relation between real and complex Lie algebras. First of all, any real Lie algebra can in a rather trivial way be extended to a complex Lie algebra, called its *complexification*, by augmenting the coefficient field from $\mathbb{R}$ to $\mathbb{C}$. In terms of an arbitrary basis of the given real Lie algebra, this amounts to passing from linear combinations with real coefficients to linear combinations with complex coefficients, without any change in the basis vectors. Conversely, a real Lie algebra is called a *real form* of a given complex Lie algebra if the latter is isomorphic to its complexification. It is important to realize that a complex Lie algebra may have several non-isomorphic real forms, but that representation theory is insensitive to this distinction (provided we think in terms of representations on complex vector spaces): a representation of a real Lie algebra on a complex vector space extends in a canonical way to a representation of its complexification, and a representation of a complex Lie algebra on a complex vector space yields a representation of any of its real forms, by restriction. For semisimple Lie algebras, a celebrated theorem of Hermann Weyl asserts that every complex semisimple Lie algebra has a unique *compact real form*: its construction is achieved by what has become known as "Weyl's unitary trick", and it is uniquely characterized by the property that the Killing form is negative definite on it. As a corollary, one infers that it also has a unique *normal real form*, but other real forms may exist as well. Cartan's procedure was to classify first the complex simple Lie algebras, which is an easier task due to the fact that the complex numbers form an algebraically closed field, and then find their different real forms. The result is presented in Table 2, where we list the four *classical series* of complex simple Lie algebras, as well as their compact and normal real forms, and the five *exceptional* simple Lie algebras, together with their dimensions; the index r refers to the rank (see below). For the sake of completeness, the Dynkin diagram (see below) is also shown.

The first classical series, the A-series, is related to the special linear group and to the special unitary group, whereas two other classical series, the B-series and the D-series, are related to the orthogonal group in odd and in even dimensions, respectively. The remaining classical series, the C-series, is related to the symplectic group and is up to the present day the least understood, in many respects. (We remark in passing that the term "symplectic" was coined by Hermann Weyl, some 70 years ago, as a synonym for "complicated": things don't seem to have changed that much!) Interestingly, it is just one of the symplectic groups which provides the most promising candidate for the symmetry group that organizes the genetic code, which is why we shall discuss the C-series in some detail in Sec. 6.2 below.

Table. 2. Cartan classification of simple Lie algebras.

Cartan Label	Complex Algebra	Compact Real Form	Normal Real Form	Dimension	Dynkin Diagram
A_r	$\mathfrak{sl}(r+1,\mathbb{C})$	$\mathfrak{su}(r+1)$	$\mathfrak{sl}(r+1,\mathbb{R})$	$r(r+2)$	
B_r	$\mathfrak{so}(2r+1,\mathbb{C})$	$\mathfrak{so}(2r+1)$	$\mathfrak{so}(r+1,r)$	$r(2r+1)$	
C_r	$\mathfrak{sp}(2r,\mathbb{C})$	$\mathfrak{sp}(2r)$	$\mathfrak{sp}(2r,\mathbb{R})$	$r(2r+1)$	
D_r	$\mathfrak{so}(2r,\mathbb{C})$	$\mathfrak{so}(2r)$	$\mathfrak{so}(r,r)$	$r(2r-1)$	
E_6				78	
E_7				133	
E_8				248	
F_4				52	
G_2				14	

One important common feature of all complex semisimple Lie algebras is that they admit a special basis, called the *Cartan–Weyl basis*, associated with the choice of a *Cartan subalgebra*. A Cartan subalgebra of a semisimple Lie algebra $\mathfrak{g}$ is an Abelian subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ whose elements, when acting on $\mathfrak{g}$ via the adjoint representation ad, are semisimple, or in other words, diagonalizable: a standard theorem of linear algebra asserts that they are then simultaneously diagonalizable. Such Cartan subalgebras exist in any complex semisimple Lie algebra $\mathfrak{g}$, and they are all conjugate; their common dimension is called the *rank* of $\mathfrak{g}$. Fixing one of them, call it $\mathfrak{h}$, we can decompose the Lie algebra $\mathfrak{g}$ into joint eigenvectors X_α under the action of $\mathfrak{h}$ (via ad), collecting the corresponding eigenvalues into linear functionals α on $\mathfrak{h}$, according to the formula

$$[H, X_\alpha] = \alpha(H) X_\alpha \quad \text{for all } H \in \mathfrak{h}\,. \tag{31}$$

The nonzero linear functionals α on $\mathfrak{h}$ for which there exist nonzero eigenvectors X_α in $\mathfrak{g}$ satisfying this equation are called *roots* and the corresponding eigenvectors X_α are called *root generators*. As it turns out, $\mathfrak{h}$ itself is precisely the joint eigenspace associated with the zero functional, whereas the joint eigenspaces associated with

the roots are all one-dimensional. A Cartan–Weyl basis of $\mathfrak{g}$ is now simply a basis of $\mathfrak{g}$ adapted to this decomposition: it consists of $r = \text{rank } \mathfrak{g}$ generators $H_1, \ldots, H_r$ that form a basis of $\mathfrak{h}$, plus $\dim \mathfrak{g} - \text{rank } \mathfrak{g}$ generators E_α associated with the roots α, satisfying the following commutation relations:

$$[H_j, H_k] = 0\,, \tag{32}$$

$$[H_j, E_\alpha] = \alpha_j E_\alpha\,, \tag{33}$$

$$[E_\alpha, E_{-\alpha}] = (E_\alpha, E_{-\alpha}) H_\alpha\,, \tag{34}$$

$$[E_\alpha, E_\beta] = N_{\alpha,\beta} E_{\alpha+\beta} \quad \text{for } \alpha + \beta \neq 0\,. \tag{35}$$

Here the coefficient α_j is the value of the root α (viewed as a linear form on $\mathfrak{h}$) on the basis vector H_j and the coefficient $N_{\alpha,\beta}$ is zero whenever $\alpha + \beta$ is not a root, whereas the vectors H_α belonging to $\mathfrak{h}$ are defined by

$$(H_\alpha, H) = \alpha(H) \quad \text{for all } H \in h\,; \tag{36}$$

their exact value depends on the normalization convention for the bilinear form (.,.) on $\mathfrak{g}$ employed (a conveniently chosen multiple of the Killing form).

The same strategy of simultaneous diagonalization of the generators in the Cartan subalgebra can be applied to arbitrary representations, instead of the adjoint. Given any representation t of $\mathfrak{g}$ on some (finite-dimensional) complex vector space V, we can decompose V into joint eigenvectors v_λ under the action of $\mathfrak{h}$ (via t), collecting the corresponding eigenvalues into linear functionals λ on $\mathfrak{h}$, according to the formula

$$t(H) v_\lambda = \lambda(H) v_\lambda \quad \text{for all } H \in \mathfrak{h}\,. \tag{37}$$

The linear functionals λ on $\mathfrak{h}$ for which there exist nonzero eigenvectors v_λ in V satisfying this equation are called *weights* and the corresponding eigenvectors v_λ are called *weight vectors*. There is however no information on the multiplicities (dimensions of the joint eigenspaces) of the weights for a general representation: their calculation can be a formidable task.

The roots of a semisimple Lie algebra of rank r form a finite set spanning an r-dimensional Euclidean space. (Technically speaking this is the real subspace of $\mathfrak{h}^*$ formed by the linear functionals on $\mathfrak{h}$ which take real values on the intersection of $\mathfrak{h}$ with the normal real form of $\mathfrak{g}$ or equivalently, which take imaginary values on the intersection of $\mathfrak{h}$ with the compact real form of $\mathfrak{g}$, but this will not be of much importance for what follows.) This *root diagram* has a very interesting geometry. For example, it admits a large number of reflection symmetries: it is symmetric about the origin (i.e. if α is a root, so must be $-\alpha$) and symmetric under reflection in all planes orthogonal to some root (i.e. if α and β are roots, so must be $s_\alpha(\beta) = \beta - 2((\beta, \alpha)/(\alpha, \alpha))\alpha$, which is the result of reflecting β in the plane orthogonal to α). As it turns out, these reflections generate a finite group, which is called the *Weyl group* $W(\mathfrak{g})$ of $\mathfrak{g}$. Moreover, when $\mathfrak{g}$ is simple, there are at most two possible values for the length of a root: either all roots have the same length, or there occur

roots of exactly two distinct lengths, the *long roots* and the *short roots*. In any case, two roots of the same length can be connected by a Weyl group transformation. The Weyl group invariance of the root diagram also allows to subdivide the roots into positive and negative, according to whether their orthogonal projection onto some fixed root is positive or negative. Once such a subdivision has been adopted, the positive roots may be decomposed into linear combinations, with non-negative integer coefficients, of r uniquely determined *simple roots* $\alpha_1, \ldots, \alpha_r$ whose lengths and relative angles completely characterize the full root diagram: this information is encoded into the so-called *Dynkin diagram* of $\mathfrak{g}$, in which r vertices are connected to one another by 0, 1, 2, or 3 lines according to whether the corresponding simple roots form an angle of 90°, 120°, 135° or 150°, respectively, with an arrow pointing from the longer to the shorter of the two simple roots (in case they have different length).

Dual to the simple roots $\alpha_1, \ldots, \alpha_r$ are the *fundamental weights* $\lambda_1, \ldots, \lambda_r$, defined by the formula

$$\frac{2(\lambda_j, \alpha_k)}{(\alpha_j, \alpha_k)} = \delta_{jk} \,. \tag{38}$$

The duality can be understood in the sense of lattices, just like in solid state physics, as a duality between the *root lattice* formed by integral linear combinations of the simple roots and the *weight lattice* formed by integral linear combinations of the fundamental weights, but we shall not enter into details of this correspondence. What is important, however, is that one of the fundamental properties of root systems implies that the root lattice is contained in the weight lattice as a sublattice: roots are special weights.

We are now ready for the last theorem of central importance: the classification of irreducible representations of semisimple Lie algebras by their highest weight. First, every irreducible representation t of a semisimple Lie algebra is a highest weight representation, i.e. there exists a vector v_Λ in the carrier space V of the representation, unique up to a scalar multiple, which is

(a) a weight vector with weight Λ,

$$t(H)v_\Lambda = \Lambda(H)v_\Lambda \quad \text{for all } H \in \mathfrak{h} \,, \tag{39}$$

(b) annihilated by the *raising operators* $t(E_\alpha)$ (α a positive root),

$$t(E_\alpha)v_\Lambda = 0 \quad \text{for positive roots } \alpha \,, \tag{40}$$

(c) cyclic under the action of the *lowering operators* $t(E_\alpha)$ (α a negative root): this means that every vector in the carrier space V of the representation can be written as a linear combination of vectors obtained by applying products of such lowering operators to v_Λ.

Second, the highest weight of an irreducible representation of $\mathfrak{g}$ belongs to the positive cone in the weight lattice of $\mathfrak{g}$, or in other words, it is a linear combination

of the fundamental weights with non-negative integer coefficients. Finally, the classification theorem states that every vector belonging to the positive cone in the weight lattice of $\mathfrak{g}$, or in other words, every linear combination of the fundamental weights with non-negative integer coefficients, appears as the highest weight of an irreducible representation of $\mathfrak{g}$ which is determined uniquely (up to equivalence). In short, irreducible representations of semisimple Lie algebras of rank r are classified by their highest weight

$$\Lambda = \sum_{k=1}^{r} m_k \lambda_k \,, \tag{41}$$

or simply by a sequence $(m_1, \ldots, m_r)$ of r non-negative integers.

As an important application, we mention *Weyl's dimension formula* which gives the dimension of an irreducible representation of highest weight Λ:

$$\dim V_\Lambda = \prod_{\alpha>0} \frac{(\Lambda+\rho,\alpha)}{(\rho,\alpha)} \,. \tag{42}$$

Here the product is over all positive roots and the vector ρ is defined to be half the sum of the positive roots,

$$2\rho = \sum_{\alpha>0} \alpha \,; \tag{43}$$

it plays an important role in many formulas of representation theory. For the simple Lie algebras $A_1 = C_1$, A_2, C_2 and A_3, C_3, the result is collected in Table 3.

An interesting consequence of Weyl's dimension formula is the following *monotonicity* of the dimension as a function of the highest weight. Suppose that

$$\Lambda = \sum_{k=1}^{r} m_k \lambda_k \quad \text{and} \quad \Lambda' = \sum_{k=1}^{r} m'_k \lambda_k \tag{44}$$

are two highest weights satisfying

$$\Lambda \leq \Lambda' \quad \text{in the sense that} \quad m_k \leq m'_k \quad \text{for } k = 1, \ldots, r \,. \tag{45}$$

Then

$$\dim V_\Lambda \leq \dim V_{\Lambda'} \,. \tag{46}$$

Given the highest weight of an irreducible representation of a semisimple Lie algebra, one may recursively construct all other weights by subtracting positive roots. In particular, successive subtraction of the same positive root α from a given weight λ leads to a *string* of weights, and the first important point is to know where the string ends. An elementary and very useful criterion in this respect is that when α and λ are such that $\lambda + \alpha$ is not a weight, then the α-string through λ is of the form $\{\lambda, \lambda - \alpha, \ldots, \lambda - m\alpha\}$, where the integer m is given by

$$m = \frac{2(\lambda,\alpha)}{(\alpha,\alpha)} \,.$$

Table 3. Dimension formula for irreducible representations of some classical Lie algebras of low rank.

Cartan Label	Complex Algebra	Compact Real Form	Highest Weight	Dimension
A_1	$\mathfrak{sl}(2,\mathbb{C})$	$\mathfrak{su}(2)$	$2s$	$2s+1$
A_2	$\mathfrak{sl}(3,\mathbb{C})$	$\mathfrak{su}(3)$	(m_1,m_2)	$\frac{1}{2}(m_1+1)(m_2+1)(m_1+m_2+2)$
A_3	$\mathfrak{sl}(4,\mathbb{C})$	$\mathfrak{su}(4)$	(m_1,m_2,m_3)	$\frac{1}{12}(m_1+1)(m_2+1)(m_3+1)$ $\times(m_1+m_2+2)(m_2+m_3+2)$ $\times(m_1+m_2+m_3+3)$
C_2	$\mathfrak{sp}(4,\mathbb{C})$	$\mathfrak{sp}(4)$	(m_1,m_2)	$\frac{1}{6}(m_1+1)(m_2+1)(m_1+m_2+2)$ $\times(m_1+2m_2+3)$
C_3	$\mathfrak{sp}(6,\mathbb{C})$	$\mathfrak{sp}(6)$	(m_1,m_2,m_3)	$\frac{1}{720}(m_1+1)(m_2+1)(m_3+1)$ $\times(m_1+m_2+2)(m_2+m_3+2)$ $\times(m_2+2m_3+3)$ $\times(m_1+m_2+m_3+3)$ $\times(m_1+m_2+2m_3+4)$ $\times(m_1+2m_2+2m_3+5)$

A second useful fact is that the weight diagram as a whole is invariant under the action of the Weyl group $W(\mathfrak{g})$ of $\mathfrak{g}$, so it decomposes naturally into *Weyl group orbits*; moreover, the multiplicities are also invariant under $W(\mathfrak{g})$ and are therefore constant along the Weyl group orbits. In particular, the weights belonging to the Weyl group orbit of the highest weight are the longest weights that appear in the weight diagram and their multiplicity is necessarily equal to 1.

Once the complete weight diagram has been found, the major problem in the explicit construction of the representation is the appearance of weights of multiplicity > 1. The first task is to calculate the multiplicities themselves: this is an issue which has been completely solved through explicit formulas, due to Freudenthal and to Kostant. Unfortunately, these formulas are often difficult to handle in practice and, in this sense, quite cumbersome. The second task is to find appropriate bases for the weight spaces of dimension > 1 on which the action of the generators in $\mathfrak{g}$ is as simple as possible. The action of the generators in the Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ on the weight spaces being completely degenerate, this requires resorting to other techniques.

It is at this point that the idea of a chain of subalgebras comes into play. More specifically, consider a sequence of reductive[e] Lie subalgebras $\mathfrak{g}_1,\ldots,\mathfrak{g}_k$ of $\mathfrak{g}$ which

[e]A Lie Algebra is called reductive if it decomposes into the direct sum of an Abelian subalgebra, which is precisely its center, and a semisimple subalgebra.

form a chain

$$\mathfrak{g} \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_k \supset \mathfrak{h}\,, \tag{47}$$

descending from $\mathfrak{g}$ all the way down to its Cartan subalgebra $\mathfrak{h}$. (We may imagine that we have put $\mathfrak{g} = \mathfrak{g}_0$ and $\mathfrak{h} = \mathfrak{g}_{k+1}$.) Given an irreducible representation of $\mathfrak{g}$ of highest weight Λ, it will in general break into several irreducible representations of $\mathfrak{g}_1$, each of which breaks into several irreducible representations of $\mathfrak{g}_2$ and so on, until finally we have a breaking into irreducible representations of $\mathfrak{h}$ which, $\mathfrak{h}$ being Abelian, are all one-dimensional: their quantum numbers are exactly the weights. However, nothing prevents us from augmenting the scheme by simultaneously diagonalizing not only the action of $\mathfrak{h}$ but also that of all invariant operators (Casimir operators) for each of the subalgebras $\mathfrak{g}_1, \ldots, \mathfrak{g}_k$, so that the weight vectors belonging to weights of multiplicity > 1 are further restricted by requiring that they belong to a definite irreducible representation of each of the intermediate subalgebras $\mathfrak{g}_1, \ldots, \mathfrak{g}_k$: they may therefore be labeled by a set of k highest weights $\Lambda_1, \ldots, \Lambda_k$ for $g_1, \ldots, \mathfrak{g}_k$, respectively, in addition to their usual weight λ. The main question is then under what conditions on the chain of subalgebras and for which representations of the original algebra this scheme leads to a complete characterization of all weight vectors (up to scalar multiples, of course). If it does not, the additional labels $\Lambda_1, \ldots, \Lambda_k$ introduced before are clearly insufficient — a situation usually referred to as the *missing label problem*. For the classical Lie algebras in the A-series, it has been shown by Gel'fand and Zetlin[49] that the additional labels derived from the so-called *canonical chain*,

$$\mathfrak{sl}(r+1, \mathbb{C}) \supset \mathfrak{sl}(r, \mathbb{C}) \oplus \mathbb{C} \supset \mathfrak{sl}(r-1, \mathbb{C}) \oplus \mathbb{C}^2 \supset \cdots \supset \mathfrak{sl}(2, \mathbb{C}) \oplus \mathbb{C}^{r-1} \supset \mathbb{C}^r\,, \tag{48}$$

are sufficient to completely remove the degeneracies between weight vectors in arbitrary irreducible representations; a similar result holds for the B-series and the D-series. For the classical Lie algebras in the C-series, on the other hand, the *canonical chain*,

$$\begin{aligned}\mathfrak{sp}(2r, \mathbb{C}) \supset \mathfrak{sp}(2r-2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C}) \supset \mathfrak{sp}(2r-4, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})^2 \supset \cdots \\ \supset \mathfrak{sp}(4, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})^{r-1} \supset \mathfrak{sl}(2, C)^r \supset \mathbb{C}^r\,,\end{aligned} \tag{49}$$

where $\mathfrak{sl}(2, \mathbb{C})^p$ denotes the direct sum of p copies of $\mathfrak{sl}(2, \mathbb{C})$, is not sufficient to completely remove the degeneracies between weight vectors in arbitrary irreducible representations. This problem has recently been addressed by Cerkaski[50] who was able to specify a complete set of labels (although he did not exhibit the action of the generators on the states resulting from his prescription).

6.2. *The symplectic symmetry*

The *complex symplectic group* $\mathrm{Sp}(2r, \mathbb{C})$ is the set of invertible complex $(2r \times 2r)$-matrices preserving the quadratic antisymmetric form ω defined by

$$\omega(z, w) = z^{\mathrm{T}} \omega w = z_1 w_{r+1} + \cdots + z_r w_{2r} - z_{r+1} w_1 - \cdots - z_{2r} w_r\,, \tag{50}$$

where superscript T denotes transpose, z and w are vectors in $\mathbb{C}^{2r}$ and ω is the (2×2)-block matrix

$$\omega = \begin{pmatrix} 0 & 1_r \\ -1_r & 0 \end{pmatrix} , \tag{51}$$

with 1_r denoting the unit $(r \times r)$-matrix. Its Lie algebra $\mathfrak{sp}(2r, \mathbb{C})$ is the set of complex $(2r \times 2r)$-matrices X satisfying

$$X^{\mathrm{T}}\omega + \omega X = 0 . \tag{52}$$

These matrices have the general (2×2)-block form

$$X = \begin{pmatrix} A & B \\ C & -A^{\mathrm{T}} \end{pmatrix} , \tag{53}$$

where A, B and C are complex $(r \times r)$-matrices, B and C being symmetric. The (complex) dimension of this Lie algebra is $r(2r + 1)$. The *unitary symplectic group* $\mathrm{Sp}(2r)$ and its Lie algebra $\mathfrak{sp}(2r)$ are obtained by imposing that the matrices involved should in addition be unitary and antihermitean, respectively; the second condition means that

$$X^\dagger + X = 0 \qquad \text{or} \qquad A^\dagger + A = 0 , \qquad B^\dagger + C = 0 = C^\dagger + B . \tag{54}$$

Similarly, the *real symplectic group* $\mathrm{Sp}(2r, \mathbb{R})$ and its Lie algebra $\mathfrak{sp}(2r, \mathbb{R})$ are obtained by imposing that the matrices involved should all be real, rather than complex. It is easily seen that $\mathfrak{sp}(2r)$ is the compact real form of $\mathfrak{sp}(2r, \mathbb{C})$ (in fact the group $\mathrm{Sp}(2r) = \mathrm{Sp}(2r, \mathbb{C}) \cap \mathrm{U}(2r)$ is compact) and $\mathfrak{sp}(2r, \mathbb{R})$ is the normal real form of $\mathfrak{sp}(2r, \mathbb{C})$: note that it is this normal real form (not the compact one) which appears naturally in Hamiltonian mechanics. All of them have (real) dimension $r(2r + 1)$.

The standard invariant bilinear form on $\mathfrak{sp}(2r, \mathbb{C})$ is defined by

$$(X, Y) = \mathrm{trace}(XY) \quad \text{for all } X, Y \in \mathfrak{sp}(2r, \mathbb{C}) ; \tag{55}$$

it is equal to $1/(2r + 2)$ times the Killing form of $\mathfrak{sp}(2r, \mathbb{C})$ and is negative definite on $\mathfrak{sp}(2r)$, as usual.

The commutation relations for the symplectic Lie algebra can be obtained by direct matrix manipulations. A particularly convenient choice of basis is obtained upon introducing the standard basis of $\mathfrak{gl}(r)$ consisting of the $(r \times r)$-matrices E_{ij} having 1 in the ith row and jth column and 0 everywhere else:

$$(E_{ij})_{kl} = \delta_{ik}\delta_{jl} . \tag{56}$$

Product and commutator of two such matrices are given by

$$E_{ij}E_{kl} = \delta_{jk}E_{il} , \qquad [E_{ij}, E_{kl}] = \delta_{jk}E_{il} - \delta_{il}E_{kj} . \tag{57}$$

The generators in $\mathfrak{sp}(2r,\mathbb{C})$ can then be written in terms of the following $(2r\times 2r)$-matrices, expressed in (2×2)-block form:

$$A_{ij}=\begin{pmatrix} E_{ij} & 0\\ 0 & -E_{ji}\end{pmatrix}\quad (1\le i,j\le r)\,, \tag{58}$$

$$B_{ij}=\begin{pmatrix} 0 & E_{ij}+E_{ji}\\ 0 & 0\end{pmatrix}\quad (1\le i\le j\le r)\,, \tag{59}$$

$$C_{ij}=\begin{pmatrix} 0 & 0\\ E_{ij}+E_{ji} & 0\end{pmatrix}\quad (1\le i\le j\le r)\,. \tag{60}$$

They have the following commutation relations:

$$[A_{ij},A_{kl}]=\delta_{jk}A_{il}-\delta_{il}A_{kj}\,, \tag{61}$$

$$[A_{ij},B_{kl}]=+\delta_{jk}B_{il}+\delta_{jl}B_{ik}\,, \tag{62}$$

$$[A_{ij},C_{kl}]=-\delta_{ik}C_{jl}-\delta_{il}C_{jk}\,, \tag{63}$$

$$[B_{ij},B_{kl}]=0=[C_{ij},C_{kl}]\,, \tag{64}$$

$$[B_{ij},C_{kl}]=\delta_{ik}A_{jl}+\delta_{jk}A_{il}+\delta_{il}A_{jk}+\delta_{jl}A_{ik}\,. \tag{65}$$

The diagonal generators spanning the Cartan subalgebra $\mathfrak{h}$ are

$$H_j=A_{jj}\quad (1\le j\le r)\,, \tag{66}$$

whereas the A_{ij} with $i\neq j$, the B_{ij} and the C_{ij} are root generators; the corresponding roots can be read off directly from the commutation relations

$$[H_j,A_{kl}]=(\delta_{jk}-\delta_{jl})A_{kl}\,, \tag{67}$$

$$[H_j,B_{kl}]=+(\delta_{jk}+\delta_{jl})B_{kl}\,, \tag{68}$$

$$[H_j,C_{kl}]=-(\delta_{jk}+\delta_{jl})C_{kl}\,. \tag{69}$$

In terms of the basis $\{e_1,\ldots,e_r\}$ dual to the basis $\{H_1,\ldots,H_r\}$,[f] this means that the root system

$$\Delta=\Delta_l\cup\Delta_s \tag{70}$$

of $\mathfrak{sp}(2r,\mathbb{C})$ is given by

$$\Delta_l=\{\pm 2e_k/1\le k\le r\}\,, \tag{71}$$

the set of long roots (of length $\sqrt{2}$), with associated generators

$$\alpha=+2e_k\implies E_\alpha=B_{kk}\,, \tag{72}$$

$$\alpha=-2e_k\implies E_\alpha=C_{kk}\,, \tag{73}$$

[f]These two bases are orthonormal except for an overall normalization factor of $\sqrt{2}$: $(H_i,H_j)=\mathrm{trace}(H_iH_j)=2\delta_{ij}$ and hence $(e_i,e_j)=\delta_{ij}/2$.

and

$$\Delta_s = \{\pm e_k \pm e_l / 1 \leq k < l \leq r\}\,, \tag{74}$$

the set of short roots (of length 1), with associated generators

$$\alpha = +e_k + e_l \implies E_\alpha = B_{kl}\,, \tag{75}$$

$$\alpha = -e_k - e_l \implies E_\alpha = C_{kl}\,, \tag{76}$$

$$\alpha = +e_k - e_l \implies E_\alpha = A_{kl}\,, \tag{77}$$

$$\alpha = -e_k + e_l \implies E_\alpha = A_{lk}\,. \tag{78}$$

The ordering in this root system will be chosen such that

$$\Delta^+ = \Delta_l^+ \cup \Delta_s^+\,, \tag{79}$$

where

$$\Delta_l^+ = \{2e_k / 1 \leq k \leq r\}\,, \tag{80}$$

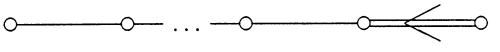
$$\Delta_s^+ = \{e_k \pm e_l / 1 \leq k < l \leq r\}\,. \tag{81}$$

It corresponds to the prescription that the positive (negative) roots are precisely those roots α for which the corresponding generators E_α are upper (lower) triangular matrices.

With this choice of ordering, the simple roots $\alpha_1, \ldots, \alpha_r$ are given by:

$$\alpha_1 = e_1 - e_2\,, \ldots, \qquad \alpha_{r-1} = e_{r-1} - e_r\,, \qquad \alpha_r = 2e_r\,. \tag{82}$$

Note that the simple roots α_i and α_j are orthogonal except when i and j are adjacent (i.e. except when $|i-j| = 1$) and that α_i and α_{i+1} form an angle of 120° when $1 \leq i \leq r-2$ and an angle of 135° when $i = r-1$, the last simple root being longer than the previous ones; this corresponds to the following Dynkin diagram:



(83)

Note also that the vector ρ which is half the sum of the positive roots (and which appears, for example, in the Weyl dimension formula) is equal to

$$\rho = re_1 + (r-1)e_2 + \cdots + 2e_{r-1} + e_r\,. \tag{84}$$

For $r = 2$, there are 8 roots, localized in the 4 corners and in the midpoints of the 4 edges of a square rhombic in the plane spanned by e_1 and e_2 (see Fig. 13), while for $r = 3$, there are 18 roots, localized in the 6 corners and in the midpoints of the 12 edges of a regular octahedron in the space spanned by e_1, e_2 and e_3 (see Fig. 14).

Dually, the fundamental weights $\lambda_1, \ldots, \lambda_r$ are given by:

$$\begin{aligned} &\lambda_1 = e_1\,, \qquad \lambda_2 = e_1 + e_2\,, \ldots\,, \\ &\lambda_{r-1} = e_1 + \cdots + e_{r-1}\,, \qquad \lambda_r = e_1 + \cdots + e_r\,. \end{aligned} \tag{85}$$

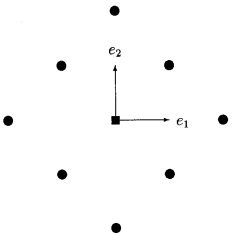


Fig. 13. Root diagram of $\mathfrak{sp}(4)$.

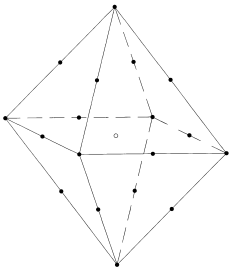


Fig. 14. Root diagram of $\mathfrak{sp}(6)$.

Forming linear combinations of $\lambda_1, \ldots, \lambda_r$ with integer coefficients gives the same set of vectors as forming linear combinations of $e_1, \ldots, e_r$ with integer coefficients: this means that the weight lattice of $\mathfrak{sp}(2r, \mathbb{C})$ is an orthogonal lattice. In particular, weights

$$\lambda = \sum_{k=1}^{r} m_k \lambda_k = \sum_{k=1}^{r} a_k e_k \tag{86}$$

of representations of $\mathfrak{sp}(2r, \mathbb{C})$ will in the following be written as r-tuples $(m_1, \ldots, m_r)$ or as r-tuples $(a_1, \ldots, a_r)$. One advantage of the orthogonal basis $\{e_1, \ldots, e_r\}$ (as compared to the non-orthogonal basis formed by the fundamental weights) is that the action of the Weyl group takes a particularly simple form. As is well known, the Weyl group of $\mathfrak{sp}(2r, \mathbb{C})$ is $\mathbb{Z}_2^r \times S_r$, where the nontrivial element in the ith $\mathbb{Z}_2$ factor acts through reflection along the ith coordinate axis (taking e_i to $-e_i$ and e_j to e_j for $j \neq i$), whereas the elements of the group S_r act through permutation of the coordinates with respect to this basis.

Concluding this subsection, we note two interesting subalgebras of $\mathfrak{sp}(2r, \mathbb{C})$, both of which are maximal semisimple subalgebras;[51,52] see also Ref. 47. One of them is the simple subalgebra $\mathfrak{sl}(r, \mathbb{C})$ spanned by the A_{ij}, the other is the semisimple subalgebra $\mathfrak{sp}(2r-2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$, where $\mathfrak{sp}(2r-2, \mathbb{C})$ is spanned by the matrices A_{ij}, B_{ij} and C_{ij} with i and j running from 1 to $r-1$ only, whereas $\mathfrak{sl}(2, \mathbb{C})$ is spanned by A_{rr}, B_{rr} and C_{rr}.

7. The Search for Symmetries in the Genetic Code

The first step in the search for symmetries in the genetic code consists in selecting a simple Lie algebra $\mathfrak{g}$ and an irreducible representation of $\mathfrak{g}$ on a vector space of dimension 64: such a representation will in the following be referred to as a *codon representation*. The reason for requiring the representation to be irreducible is that a reducible representation is a composed object and can be expressed as the sum of irreducible components. The use of a reducible representation would correspond not to the starting point of the symmetry breaking process but rather to a posterior stage, in which some kind of breaking has already occurred. Similarly, only simple Lie algebras are considered because they are the building blocks for the construction of semisimple Lie algebras. A semisimple Lie algebra is the direct sum of simple Lie algebras and would correspond to a composite symmetry.

Selecting a general Lie algebra that is not even semisimple would not extend the spectrum of options. We recall from Sec. 6 that according to the Levi–Malcev theorem, a general Lie algebra decomposes into the semidirect sum of a semisimple subalgebra and its radical. Moreover, any irreducible representation of such a Lie algebra can under this decomposition be written as the tensor product of an irreducible representation of its semisimple subalgebra with a one-dimensional representation of its radical (see, e.g. Ref. 42, p. 226). As the latter does not contribute to dimensions and multiplicities when it comes to computing branching rules, we may just as well assume right from the start, and without any loss of generality, that the original Lie algebra is semisimple.

7.1. *The classification of codon representations*

The determination of all codon representations is based on the Cartan classification theorem explained in Sec. 6, according to which the four series of classical Lie algebras, $A_r = \mathfrak{su}(r+1)$, $B_r = \mathfrak{so}(2r+1)$, $C_r = \mathfrak{sp}(2r)$ and $D_r = \mathfrak{so}(2r)$, together with the five exceptional Lie algebras E_6, E_7, E_8, F_4 and G_2, exhaust the class of simple Lie algebras.[g] The result is that only the simple Lie algebras $\mathfrak{su}(2)$, $\mathfrak{su}(3)$, $\mathfrak{su}(4)$, $\mathfrak{sp}(4)$, $\mathfrak{sp}(6)$, $\mathfrak{so}(13)$, $\mathfrak{so}(14)$, G_2 and obviously, $\mathfrak{su}(64)$, $\mathfrak{so}(64)$ and $\mathfrak{sp}(64)$, admit a codon representation.

This can most conveniently be seen by inspecting the tables of dimensions of irreducible representations of the classical and exceptional Lie algebras. The existence of only a finite number of codon representations is a consequence of the monotonous growth of the dimensions of irreducible representations of simple Lie algebras with their rank: algebras of high rank admit only high-dimensional representations. In particular, there is for each of the classical series a maximum rank beyond which all representations have dimension higher than 64. Therefore, the list of codon representations is finite.

[g]In this section, we shall express the classical Lie algebras in terms of their compact real forms.

To illustrate this point and for the convenience of the reader, we list in Table 4 all irreducible representations of the classical Lie algebras, in terms of their highest weights, together with their dimension d, up to and including the first whose dimension exceeds 64, following the tables of McKay and Patera.[53] In the case of the A-series, note that these come in complex conjugate pairs, with highest weights $(a_1, a_2, \ldots, a_{r-1}, a_r)$ and $(a_r, a_{r-1}, \ldots, a_2, a_1)$: for simplicity, only one of them is shown, even though they are in general inequivalent (except when $a_1 = a_r, a_2 = a_{r-1}, \ldots$).

Inspection of Table 4 provides convincing evidence for certain regularities which can in fact be proved with mathematical rigor by using the monotonicity of the dimension as a function of the highest weight mentioned in the previous section. First of all, the lowest-dimensional irreducible representation is always the first fundamental or defining representation, with highest weight $(1, 0, 0, \ldots, 0)$, and in the case of the A-series, its complex conjugate, with highest weight $(0, \ldots, 0, 0, 1)$. This not only identifies the defining representation of $\mathfrak{su}(64)$ and its complex conjugate, as well as the defining representations of $\mathfrak{so}(64)$ and of $\mathfrak{sp}(64)$, as codon representations, but also shows that there are no codon representations of $\mathfrak{su}(n)$, $\mathfrak{so}(n)$ or $\mathfrak{sp}(n)$ for $n \geq 65$. Similarly, except for a few small values of n, the next-to-lowest-dimensional irreducible representations is the second fundamental representation, with highest weight $(0, 1, 0, \ldots, 0)$, and in the case of the A-series, its complex conjugate, with highest weight $(0, \ldots, 0, 1, 0)$: more precisely, this is true provided that $n \geq 4$ for $\mathfrak{su}(n)$, $n \geq 15$ for $\mathfrak{so}(n)$ and $n \geq 4$ for $\mathfrak{sp}(n)$. This representation is built from antisymmetric tensors of rank 2 over the vectors in the defining representation and has dimension $n(n-1)/2$ in the case of $\mathfrak{su}(n)$ or $\mathfrak{so}(n)$ and dimension $n(n-1)/2-1$ in the case of $\mathfrak{sp}(n)$. None of these numbers is a power of 2, so these irreducible representations do not qualify as codon representations. Moreover, the argument shows that there can be no codon representations of $\mathfrak{su}(n)$ when $n \geq 12$, of $\mathfrak{so}(n)$ when $n \geq 15$ and of $\mathfrak{sp}(n)$ when $n \geq 12$, except the defining representations mentioned above. Finally, the only exceptional Lie algebra which admits a 64-dimensional irreducible representation is G_2,[53] and so we arrive at Table 5 as the complete list of codon representations.

We conclude this subsection with a few words on the codon representations of the high rank algebras $\mathfrak{su}(64)$, $\mathfrak{so}(64)$ and $\mathfrak{sp}(64)$ listed in the lower part of Table 5, which have been excluded from our search. They have a large number of generators and an enormous number of possible chains and can therefore reproduce practically any degeneracy, so that a symmetry breaking scheme based on one of these algebras is devoid of predictive power. For example, in the case of $\mathfrak{su}(64)$, one can reproduce *any* distribution of multiplets by breaking to a subalgebra of the form $\mathfrak{su}(n_1) \oplus \cdots \oplus \mathfrak{su}(n_k)$ with k equal to the number of multiplets and $n_1, \ldots, n_k$ (with $n_1 + \cdots + n_k = 64$) equal to their dimensions, a subalgebra which incidentally can be realized as the (semisimple part of the) stability algebra of a specific column vector in the 64-dimensional representation space, having mutually distinct entries λ_1 in the first n_1 rows, $\ldots$, λ_k in the last n_k rows.

Table 4a. Highest weights and dimensions of low-dimensional irreducible representations of the classical Lie algebras of low rank: A_r.

$A_2 - \mathfrak{su}(3)$		$A_3 - \mathfrak{su}(4)$		$A_4 - \mathfrak{su}(5)$		$A_5 - \mathfrak{su}(6)$		$A_6 - \mathfrak{su}(7)$	
Highest Weight	d	Highest Weight	d	Highest Weight	d	Highest Weight	d	Highest Weight	d
(0, 1)	3	(0, 0, 1)	4	(0, 0, 0, 1)	5	(0, 0, 0, 0, 1)	6	(0, 0, 0, 0, 0, 1)	7
(0, 2)	6	(0, 1, 0)	6	(0, 0, 1, 0)	10	(0, 0, 0, 1, 0)	15	(0, 0, 0, 0, 1, 0)	21
(1, 1)	8	(0, 0, 2)	10	(0, 0, 0, 2)	15	(0, 0, 1, 0, 0)	20	(0, 0, 0, 0, 0, 2)	28
(0, 3)	10	(1, 0, 1)	15	(1, 0, 0, 1)	24	(0, 0, 0, 0, 2)	21	(0, 0, 0, 1, 0, 0)	35
(0, 4)	15	(0, 0, 3)	20	(0, 0, 0, 3)	35	(1, 0, 0, 0, 1)	35	(1, 0, 0, 0, 0, 1)	48
(1, 2)	15	(0, 1, 1)	20	(0, 0, 1, 1)	40	(0, 0, 0, 0, 3)	56	(0, 0, 0, 0, 0, 3)	84
(0, 5)	21	(0, 2, 0)	20	(0, 1, 0, 1)	45	(0, 0, 0, 1, 1)	70		
(1, 3)	24	(0, 0, 4)	35	(0, 0, 2, 0)	50				
(2, 2)	27	(1, 0, 2)	36	(0, 0, 0, 4)	70				
(0, 6)	28	(0, 1, 2)	45						
(1, 4)	35	(0, 3, 0)	50						
(0, 7)	36	(0, 0, 5)	56						
(2, 3)	42	(0, 2, 1)	60						
(0, 8)	45	(**1, 1, 1**)	**64**						
(1, 5)	48	(1, 0, 3)	70						
(0, 9)	55								
(2, 4)	60								
(1, 6)	63								
(**3, 3**)	**64**								
(0, 10)	66								

Table 4a. (*Continued*)

$A_7 - \mathfrak{su}(8)$		$A_8 - \mathfrak{su}(9)$		$A_9 - \mathfrak{su}(10)$		$A_{10} - \mathfrak{su}(11)$		$A_{11} - \mathfrak{su}(12)$	
Highest Weight	d	Highest Weight	d	Highest Weight	d	Highest Weight	d	Highest Weight	d
(0, 0, 0, 0, 0, 0, 1)	8	(0, 0, 0, 0, 0, 0, 0, 1)	9	(0, 0, 0, 0, 0, 0, 0, 0, 1)	10	(0, 0, 0, 0, 0, 0, 0, 0, 0, 1)	11	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1)	12
(0, 0, 0, 0, 0, 1, 0)	28	(0, 0, 0, 0, 0, 0, 1, 0)	36	(0, 0, 0, 0, 0, 0, 0, 1, 0)	45	(0, 0, 0, 0, 0, 0, 0, 0, 1, 0)	55	(0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0)	66
(0, 0, 0, 0, 0, 0, 2)	36	(0, 0, 0, 0, 0, 0, 0, 2)	45	(0, 0, 0, 0, 0, 0, 0, 0, 2)	55	(0, 0, 0, 0, 0, 0, 0, 0, 0, 2)	66		
(0, 0, 0, 0, 1, 0, 0)	56	(1, 0, 0, 0, 0, 0, 0, 1)	80	(1, 0, 0, 0, 0, 0, 0, 0, 1)	99				
(1, 0, 0, 0, 0, 0, 1)	63								
(0, 0, 0, 1, 0, 0, 0)	70								

Table 4b. Highest weights and dimensions of low-dimensional irreducible representations of the classical Lie algebras of low rank: B_r.

$B_3 - \mathfrak{so}(7)$		$B_4 - \mathfrak{so}(9)$		$B_5 - \mathfrak{so}(11)$		$B_6 - \mathfrak{so}(13)$		$B_7 - \mathfrak{so}(15)$	
Highest Weight	d	Highest Weight	d	Highest Weight	d	Highest Weight	d	Highest Weight	d
(1, 0, 0)	7	(1, 0, 0, 0)	9	(1, 0, 0, 0, 0)	11	(1, 0, 0, 0, 0, 0)	13	(1, 0, 0, 0, 0, 0, 0)	11
(0, 0, 1)	8	(0, 0, 0, 1)	16	(0, 0, 0, 0, 1)	32	(**0, 0, 0, 0, 0, 1**)	**64**	(0, 1, 0, 0, 0, 0, 0)	105
(0, 1, 0)	21	(0, 1, 0, 0)	36	(0, 1, 0, 0, 0)	55	(0, 1, 0, 0, 0, 0)	78		
(2, 0, 0)	27	(2, 0, 0, 0)	44	(2, 0, 0, 0, 0)	65				
(0, 0, 2)	35	(0, 0, 1, 0)	84						
(1, 0, 1)	48								
(3, 0, 0)	77								

Table 4c. Highest weights and dimensions of low-dimensional irreducible representations of the classical Lie algebras of low rank: C_r.

$C_2 - \mathfrak{sp}(4)$		$C_3 - \mathfrak{sp}(6)$		$C_4 - \mathfrak{sp}(8)$		$C_5 - \mathfrak{sp}(10)$		$C_6 - \mathfrak{sp}(12)$	
Highest Weight	d	Highest Weight	d	Highest Weight	d	Highest Weight	d	Highest Weight	d
(1, 0)	4	(1, 0, 0)	6	(1, 0, 0, 0)	8	(1, 0, 0, 0, 0)	10	(1, 0, 0, 0, 0, 0)	12
(0, 1)	5	(0, 1, 0)	14	(0, 1, 0, 0)	27	(0, 1, 0, 0, 0)	44	(0, 1, 0, 0, 0, 0)	65
(2, 0)	10	(0, 0, 1)	14	(2, 0, 0, 0)	36	(2, 0, 0, 0, 0)	55		
(0, 2)	14	(2, 0, 0)	21	(0, 0, 0, 1)	42	(0, 0, 1, 0, 0)	110		
(1, 1)	16	(3, 0, 0)	56	(0, 0, 1, 0)	48				
(3, 0)	20	(**1, 1, 0**)	**64**	(3, 0, 0, 0)	120				
(0, 3)	30	(1, 0, 1)	70						
(2, 1)	35								
(4, 0)	35								
(1, 2)	40								
(0, 4)	55								
(5, 0)	56								
(**3, 1**)	**64**								
(1, 3)	80								

Table 4d. Highest weights and dimensions of low-dimensional irreducible representations of the classical Lie algebras of low rank: D_r.

$D_4 - \mathfrak{so}(8)$		$D_5 - \mathfrak{so}(10)$		$D_6 - \mathfrak{so}(12)$		$D_7 - \mathfrak{so}(14)$		$D_8 - \mathfrak{so}(16)$	
Highest Weight	d	Highest Weight	d	Highest Weight	d	Highest Weight	d	Highest Weight	d
(1, 0, 0, 0)	8	(1, 0, 0, 0, 0)	10	(1, 0, 0, 0, 0, 0)	12	(1, 0, 0, 0, 0, 0, 0)	14	(1, 0, 0, 0, 0, 0, 0, 0)	16
(0, 0, 1, 0)	8	(0, 0, 0, 1, 0)	16	(0, 0, 0, 0, 1, 0)	32	(**0, 0, 0, 0, 0, 1, 0**)	**64**	(0, 1, 0, 0, 0, 0, 0, 0)	120
(0, 0, 0, 1)	8	(0, 0, 0, 0, 1)	16	(0, 0, 0, 0, 0, 1)	32	(**0, 0, 0, 0, 0, 0, 1**)	**64**		
(0, 1, 0, 0)	28	(0, 1, 0, 0, 0)	45	(0, 1, 0, 0, 0, 0)	66	(0, 1, 0, 0, 0, 0, 0)	91		
(2, 0, 0, 0)	35	(2, 0, 0, 0, 0)	54						
(0, 0, 2, 0)	35	(0, 0, 1, 0, 0)	120						
(0, 0, 0, 2)	35								
(1, 0, 1, 0)	56								
(1, 0, 0, 1)	56								
(0, 0, 1, 1)	56								
(3, 0, 0, 0)	112								
(0, 0, 3, 0)	112								
(0, 0, 0, 3)	112								

Table 5. Codon representations of simple Lie algebras.

Cartan Label	Simple Lie Algebra	Highest Weight
A_1	$\mathfrak{su}(2)$	63
A_2	$\mathfrak{su}(3)$	(3, 3)
C_2	$\mathfrak{sp}(4)$	(3, 1)
G_2		(1, 1)
A_3	$\mathfrak{su}(4)$	(1, 1, 1)
C_3	$\mathfrak{sp}(6)$	(1, 1, 0)
B_6	$\mathfrak{so}(13)$	(0, 0, 0, 0, 0, 1)
		(0, 0, 0, 0, 0, 1, 0)
D_7	$\mathfrak{so}(14)$	(0, 0, 0, 0, 0, 0, 1)
C_{32}	$\mathfrak{sp}(64)$	(1, 0, . . . , 0)
D_{32}	$\mathfrak{so}(64)$	(1, 0, . . . , 0)
A_{63}	$\mathfrak{su}(64)$	(1, 0, . . . , 0)
		(0, . . . , 0, 1)

7.2. *Symmetry breaking through chains of subalgebras*

Starting from one of the codon representations listed in the upper part of Table 5, the first step of the analysis consists in establishing its branching rules under the reduction of the original simple Lie algebra to any of its maximal subalgebras. For this purpose, it is sufficient to consider maximal semisimple subalgebras, because a possible nontrivial center does not contribute to dimensions or branching rules. Maximal semisimple subalgebras of simple Lie algebras have been classified by Dynkin[51,52]; see also Ref. 47: those of the simple Lie algebras which appear in the upper part of Table 5 are shown in Table 6. Note that almost all these maximal subalgebras are associated with irreducible Riemannian symmetric spaces in general (see Ref. 45, p. 518) and with the various types of Grassmannians in particular: the subalgebras $\mathfrak{su}(p) \oplus \mathfrak{su}(q)$ of $\mathfrak{su}(p+q)$ with the complex Grassmannians $\mathrm{SU}(p+q)/S(\mathrm{U}(p) \times \mathrm{U}(q))$ (series $A\ III$), the subalgebras $\mathfrak{so}(p) \oplus \mathfrak{so}(q)$ of $\mathfrak{so}(p+q)$ with the real Grassmannians $\mathrm{SO}(p+q)/\mathrm{SO}(p) \times \mathrm{SO}(q)$ (series $BD\ I$) and the subalgebras $\mathfrak{sp}(p) \oplus \mathfrak{sp}(p+q)$ of $\mathfrak{sp}(p+q)$ with the quaternionic Grassmannians $\mathrm{Sp}(p+q)/\mathrm{Sp}(p) \times \mathrm{Sp}(q)$ (series $C\ II$). Other important series involve the subalgebras $\mathfrak{so}(n)$ of $\mathfrak{su}(n)$ corresponding to the symmetric spaces $\mathrm{SU}(n)/\mathrm{SO}(n)$ (series $A\ I$), the subalgebras $\mathfrak{sp}(2n)$ of $\mathfrak{su}(2n)$ corresponding to the symmetric spaces $\mathrm{SU}(2n)/\mathrm{Sp}(2n)$ (series $A\ II$), the subalgebras $\mathfrak{su}(n)$ of $\mathfrak{so}(2n)$ corresponding to the symmetric spaces $\mathrm{SO}(2n)/\mathrm{U}(n)$ (series $D\ III$) and the subalgebras $\mathfrak{su}(n)$ of $\mathfrak{sp}(2n)$ corresponding to the symmetric spaces $\mathrm{Sp}(2n)/\mathrm{U}(n)$ (series $C\ I$). Similarly, the branching rules for irreducible representations under such reductions are known: they have been the subject of long and intensive investigations by many authors, using a variety

Table 6. Maximal semisimple subalgebras of some simple Lie algebras ($\mathfrak{su}(2) \cong \mathfrak{so}(3) \cong \mathfrak{sp}(2)$, $\mathfrak{su}(2) \oplus \mathfrak{su}(2) \cong \mathfrak{so}(4)$, $\mathfrak{sp}(4) \cong \mathfrak{so}(5)$, $\mathfrak{su}(4) \cong \mathfrak{so}(6)$; the last column indicates the associated symmetric space [Ref. 45, p. 518]).

Cartan Label	Simple Lie Algebra	Maximal Semisimple Subalgebras	Cartan Label
A_1	$\mathfrak{su}(2)$	—	—
A_2	$\mathfrak{su}(3)$	$\mathfrak{su}(2)$	$A\ III$
		$\mathfrak{so}(3)$	$A\ I$
C_2	$\mathfrak{sp}(4)$	$\mathfrak{sp}(2) \oplus \mathfrak{sp}(2)$	$C\ II/BD\ I$
		$\mathfrak{su}(2)$	—
G_2		$\mathfrak{su}(2) \oplus \mathfrak{su}(2)$	G
		$\mathfrak{su}(3), \mathfrak{su}(2)$	—
A_3	$\mathfrak{su}(4)$	$\mathfrak{su}(3)$	$A\ III$
		$\mathfrak{su}(2) \times \mathfrak{su}(2)$	—
		$\mathfrak{sp}(4)$	$A\ II/BD\ I$
C_3	$\mathfrak{sp}(6)$	$\mathfrak{sp}(4) \oplus \mathfrak{sp}(2)$	$C\ II$
		$\mathfrak{su}(3)$	$C\ I$
		$\mathfrak{sp}(2) \times \mathfrak{so}(3)$	—
		$\mathfrak{su}(2)$	—
B_6	$\mathfrak{so}(13)$	$\mathfrak{so}(12), \mathfrak{so}(3) \oplus \mathfrak{so}(10),$ $\mathfrak{so}(4) \oplus \mathfrak{so}(9), \mathfrak{so}(5) \oplus \mathfrak{so}(8),$ $\mathfrak{so}(6) \oplus \mathfrak{so}(7)$	$BD\ I$
		$\mathfrak{su}(2)$	—
D_7	$\mathfrak{so}(14)$	$\mathfrak{so}(13), \mathfrak{so}(3) \oplus \mathfrak{so}(11),$ $\mathfrak{so}(4) \oplus \mathfrak{so}(10), \mathfrak{so}(5) \oplus \mathfrak{so}(9),$ $\mathfrak{so}(6) \oplus \mathfrak{so}(8), \mathfrak{so}(7) \oplus \mathfrak{so}(7)$	$BD\ I$
		$\mathfrak{su}(7)$	$D\ III$
		$\mathfrak{sp}(6), \mathfrak{sp}(4), G_2$	—

of techniques, and most of the results (for simple Lie algebras of rank ≤ 8) are summarized in the tables of McKay and Patera.[53]

This process of symmetry reduction to maximal subalgebras can be repeated and leads to descending chains of subalgebras, each of which is maximal in the previous one. (It may be convenient to point out that the restriction to maximal subalgebras is essentially just a matter of convenience: it means that there is no intermediate symmetry between two successive steps in the chain, for if there were, one could simply insert the missing step(s) to produce another, longer chain which does satisfy the maximality condition.) We also note that the problem of explicitly constructing all possible such chains requires classifying the maximal semisimple subalgebras of semisimple Lie algebras, not just of simple Lie algebras. Fortunately, this problem can be reduced to the previous one, due to a theorem of Dynkin (see

Ref. 51, Theorem 15.1) which states that the maximal semisimple subalgebras $\mathfrak{g}'$ of a semisimple Lie algebra $\mathfrak{g}$ are of two types:

(a) the *simple type*: up to an isomorphism, including an appropriate permutation of the simple ideals that constitute $\mathfrak{g}$, we have $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1$ and $\mathfrak{g}' = \mathfrak{g}'_0 \oplus \mathfrak{g}_1$, where $\mathfrak{g}_0$ is one of the simple ideals of $\mathfrak{g}$, $\mathfrak{g}_1$ is the direct sum of the other simple ideals of $\mathfrak{g}$ and $\mathfrak{g}'_0$ is a maximal semisimple subalgebra of $\mathfrak{g}_0$,
(b) the *diagonal type*: up an isomorphism, including an appropriate permutation of the simple ideals that constitute $\mathfrak{g}$, we have $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_1$ and $\mathfrak{g}' = \mathfrak{g}_0 \oplus \mathfrak{g}_1$, where $\mathfrak{g}_0$ is one of the simple ideals of $\mathfrak{g}$ that occurs (at least) twice in $\mathfrak{g}$, $\mathfrak{g}_1$ is the direct sum of the other simple ideals of $\mathfrak{g}$ and the embedding of $\mathfrak{g}_0$ into $\mathfrak{g}_0 \oplus \mathfrak{g}_0$ is the diagonal one, mapping $X_0 \in \mathfrak{g}_0$ to $(X_0, X_0) \in \mathfrak{g}_0 \oplus \mathfrak{g}_0$.

Correspondingly, the chains can be classified into simple chains, i.e. chains that do not involve breaking to a subalgebra of diagonal type at any step, and diagonal chains, i.e. chains that do involve breaking to a subalgebra of diagonal type in at least one step. In this review, we shall — following the original approach[38,39] — restrict ourselves to considering only simple chains, but we point out that the inclusion of diagonal chains, which has been completed recently,[54,55] does not change the final picture in any significant way: these results will be presented in a separate paper.[56]

For every chain, the resulting distribution of multiplets must be compared with that observed in the genetic code, summarized in Table 7. More precisely, the strategy is to proceed along each chain one step at a time and to analyze, after each step, whether the resulting pattern of degeneracies is still compatible with that of the genetic code. If it is not, the chain is *non-surviving* in the sense that it may simply be discarded without further analysis. If it is, the chain is *surviving* (up to the stage considered), which means that we must proceed to analyze the possible next steps of the symmetry breaking. Of course, the number of chains that must be considered is *a priori* very large, especially when the simple Lie algebras of medium rank, $\mathfrak{so}(13)$ and $\mathfrak{so}(14)$, are included. Therefore, it is important to

Table 7. Dimensions and multiplicities in the standard genetic code.

Dimension of Multiplet	Number of Multiplets	Amino Acids
6	3	Arg, Leu, Ser
4	5	Ala, Gly, Pro, Thr, Val
3	2	Ile, TERM
2	9	Asn, Asp, Cys, Gln, Glu, His, Lys, Phe, Tyr
1	2	Met, Trp

formulate criteria that allow to identify non-surviving chains and thereby to reduce drastically the number of possibilities which remain to be analyzed.

To begin with, observe that when any one of the following criteria is satisfied, the symmetry breaking must be stopped at (or even before) the stage considered.

- *More than* 21 *multiplets.*
- *More than* 2 *singlets.*
- *More than* 4 *odd-dimensional multiplets.*
- *Not enough multiplets of dimension* ≥ 6 *or* ≥ 4.

This results from the fact that, as one proceeds along any given chain, the total number of multiplets, the number of singlets and the number of odd-dimensional multiplets can never decrease. (An odd-dimensional multiplet will always break into at least one odd-dimensional multiplet, plus some others.) Similarly, the last criterion expresses the necessity of having sufficiently many subspaces of sufficiently high dimension in order to be able to generate, through posterior breaking, the sextets 3 and 5 quartets observed in the genetic code. On the other hand, it is equally clear that when one of the following criteria is satisfied, the symmetry breaking must not be stopped but has to proceed at least one stage further.

- *Less than* 21 *multiplets.*
- *Existence of multiplets of dimension* ≥ 7.
- *More than* 3 *multiplets of dimension* 6.
- *Existence of multiplets of dimension* 5.

In many cases, it is the clash between one criterion from the first list and one from the second list that enables us to discard a chain as non-surviving.

Another important criterion for a chain to be non-surviving is

- *Total pairing.*

This means that the reduction process has led to a situation in which all irreducible representations of the subalgebra considered come in complex conjugate pairs or, in the case of self-conjugate representations, with even multiplicity: since such a pairing of representations cannot be removed by any posterior symmetry breaking, this leads to schemes where all multiplicities are even and which are therefore unable to generate the 3 sextets, 5 quartets or 9 doublets observed in the genetic code.

The search for symmetries in the genetic code is thus performed according to the following strategy.

(1) Select one of the codon representations from (the upper part of) Table 5.
(2) Submit this representation to symmetry breaking through all possible chains of maximal subalgebras, beginning with one of the possibilities listed in Table 6.
(3) After each step, analyze the result and discard all chains that turn out to be non-surviving, according to the criteria stated above, before proceeding to the next stage.

This procedure is conveniently divided into two phases.

Phase 1: *Breaking the primordial symmetry to* $\mathfrak{su}(2)$*-symmetries.*
During the first phase, symmetry breaking proceeds through chains of maximal semisimple subalgebras. Every such chain will necessarily terminate in a direct sum of p copies of $\mathfrak{su}(2)$, the most elementary of all simple Lie algebras, where p may range from 1 up to the rank of the original simple Lie algebra.

Phase 2: *Breaking the* $\mathfrak{su}(2)$*-symmetries.*
The second phase consists in breaking (some of) the $\mathfrak{su}(2)$-subalgebras obtained after phase 1 has been completed.

7.2.1. *Breaking the primordial symmetry to* $\mathfrak{su}(2)$*-symmetries*

As an example, we shall in the following analyze all possible simple chains for the simple Lie algebras of low rank that admit a codon representation, together with the pertinent branching rules, in order to identify which of these are surviving after the first phase has been completed. For $\mathfrak{su}(2)$, this part of the chain is trivial, so we may pass on to consider the five simple Lie algebras $\mathfrak{su}(3)$, $\mathfrak{sp}(4)$, G_2, $\mathfrak{su}(4)$ and $\mathfrak{sp}(6)$. The irreducible (sub-)representations appearing in the course of the reduction process will be labeled by their highest weight, which in the case of $\mathfrak{su}(2)$ is equal to $2s$, s beign the spin.

(1) $\mathfrak{su}(3)$ *chains*: There are two possible chains:

Chain 1: $\mathfrak{su}(3) \supset \mathfrak{su}(2)$ (corresponding to *A III*)
Chain 2: $\mathfrak{su}(3) \supset \mathfrak{su}(2)$ (correspnding to *A I*).

The branching rules for an arbitrary irreducible representation of $\mathfrak{su}(3)$ under reduction to its maximal subalgebras can be obtained by the method of generating functions,[53] which for the case of the irreducible representation of highest weight (3, 3) leads to the two decompositions shown in Table 8. In both cases, we have less than 21 subspaces (chain 1: 16, chain 2: 8), so the symmetry breaking will have to proceed, but we already have 8 odd-dimensional subspaces, so neither of the two chains is surviving.

(2) $\mathfrak{sp}(4)$ *chains*: Again, there are two possible chains:

Chain 1: $\mathfrak{sp}(4) \supset \mathfrak{su}(2)$
Chain 1: $\mathfrak{sp}(4) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$.

The branching rules for the irreducible representation of highest weight (3, 1) of $\mathfrak{sp}(4)$ under reduction to its maximal subalgebras[53] give the decompositions shown in Table 9. In both cases, we have only 8 subspaces, so the symmetry breaking will have to proceed. On the other hand, there are no odd-dimensional subspaces and sufficiently many subspaces of sufficiently high dimension to generate 3 sextets and 5 quartets, so both chains are surviving.

Table 8. Branching of the codon representation of $\mathfrak{su}(3)$ in the non-surviving chains $\mathfrak{su}(3) \supset \mathfrak{su}(2)$ ($A\ III$, left) and $\mathfrak{su}(3) \supset \mathfrak{su}(2)$ ($A\ I$, right).

$\mathfrak{su}(2)$	$(A\ III)$	$\mathfrak{su}(2)$	$(A\ I)$
$2s$	d	$2s$	d
6	7	12	13
5	6	10	11
5	6	8	9
4	5	8	9
4	5	6	7
4	5	6	7
3	4	4	5
3	4	2	3
3	4		
3	4		
2	3		
2	3		
2	3		
1	2		
1	2		
0	1		
16 subspaces		8 subspaces	

Table 9. Branching of the codon representation of $\mathfrak{sp}(4)$ in the surviving chains $\mathfrak{sp}(4) \supset \mathfrak{su}(2)$ (left) and $\mathfrak{sp}(4) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$ (right).

$\mathfrak{su}(2)$		$\mathfrak{su}(2) \oplus \mathfrak{su}(2)$	
$2s$	d	$2s_1 - 2s_2$	d
13	14	$3-2$	12
11	12	$2-3$	12
9	10	$4-1$	10
7	8	$1-4$	10
7	8	$2-1$	6
5	6	$1-2$	6
3	4	$3-0$	4
1	2	$0-3$	4
8 subspaces		8 subspaces	

Table 10. Branching of the codon representation of G_2 in the chains $G_2 \supset \mathfrak{su}(2)\oplus\mathfrak{su}(2)$ (surviving, left), $G_2 \supset \mathfrak{su}(3)$ (non-surviving, middle) and $G_2 \supset \mathfrak{su}(2)$ (non-surviving, right).

$\mathfrak{su}(2) \oplus \mathfrak{su}(2)$		$\mathfrak{su}(3)$		$\mathfrak{su}(2)$	
$2s_1 - 2s_2$	d	Highest Weight	d	$2s$	d
$2-4$	15	(2, 1)	15	16	17
$1-5$	12	(1, 2)	15	14	15
$2-2$	9	(1, 1)	8	10	11
$3-1$	8	(1, 1)	8	8	9
$1-3$	8	(2, 0)	6	6	7
$0-4$	5	(0, 2)	6	4	5
$1-1$	4	(1, 0)	3		
$0-2$	3	(0, 1)	3		
8 subspaces		8 subspaces		6 subspaces	

(3) G_2 *chains*: There are now four possible chains:

Chain 1: $G_2 \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$
Chain 2: $G_2 \supset \mathfrak{su}(3) \supset \mathfrak{su}(2)$ (corresponding to *A III*)
Chain 3: $G_2 \supset \mathfrak{su}(3) \supset \mathfrak{su}(2)$ (corresponding to *A I*)
Chain 4: $G_2 \supset \mathfrak{su}(2)$

The branching rules for the irreducible representation of highest weight (1, 1) of G_2 under reduction to its maximal subalgebras[53] give the decompositions shown (partially) in Table 10. The second and third chain are eliminated due to the onset of total pairing at the $\mathfrak{su}(3)$ level. In the last case, we have only 6 subspaces, so the symetry breaking will have to proceed, but all of these are odd-dimensional. Thus only chain 1 is surviving.

(4) $\mathfrak{su}(4)$ *chains*: There are five chains to be considered:

Chain 1: $\mathfrak{su}(4) \supset \mathfrak{su}(3) \supset \mathfrak{su}(2)$ (corresponding to *A III*)
Chain 2: $\mathfrak{su}(4) \supset \mathfrak{su}(3) \supset \mathfrak{su}(2)$ (corresponding to *A I*)
Chain 3: $\mathfrak{su}(4) \supset \mathfrak{sp}(4) \supset \mathfrak{su}(2)$
Chain 4: $\mathfrak{su}(4) \supset \mathfrak{sp}(4) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$
Chain 5: $\mathfrak{su}(4) \supset \mathfrak{su}(2) \times \mathfrak{su}(2)$.

Using the branching rules for the irreducible representation of highest weight (1, 1, 1) of $\mathfrak{su}(4)$,[53] all five chains are eliminated. The first two chains are eliminated due to the onset of total pairing at the $\mathfrak{su}(3)$ level. In the remaining cases, we have less than 21 subspaces (chain 3: 10, chain 4: 14, chain 5: 8), so the symmetry breaking will have to proceed, but we already have more than 4 odd-dimensional subspaces (chain 3: 10, chain 4: 8, chain 5: 8).

(5) $\mathfrak{sp}(6)$ *chains*: In this case, we have six chains:

Chain 1: $\mathfrak{sp}(6) \supset \mathfrak{su}(3) \supset \mathfrak{su}(2)$ (corresponding to *A III*)
Chain 2: $\mathfrak{sp}(6) \supset \mathfrak{su}(3) \supset \mathfrak{su}(2)$ (corresponding to *A I*)

Chain 3: $\mathfrak{sp}(6) \supset \mathfrak{sp}(4) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$
Chain 4: $\mathfrak{sp}(6) \supset \mathfrak{sp}(4) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$.
Chain 5: $\mathfrak{sp}(6) \supset \mathfrak{su}(2)$
Chain 6: $\mathfrak{sp}(6) \supset \mathfrak{su}(2) \times \mathfrak{su}(2)$.

Using the branching rules for the irreducible representation of highest weight (1, 1, 0) of $\mathfrak{sp}(6)$,[53] the first two chains are eliminated due to the onset of total pairing at the $\mathfrak{su}(3)$ level. The remaining ones, shown in Tables 11a and 11b, are surviving: they meet all the requirements to reproduce the degeneracies of the genetic code, provided the symmetry breaking process proceeds.

7.2.2. *Breaking the* $\mathfrak{su}(2)$*-summetries: Part 1*

Let us begin by collecting the surviving chains from the first phase of the symmetry breaking process: they are

Chain 1: $\mathfrak{su}(2)$
Chain 2: $\mathfrak{sp}(4) \supset \mathfrak{su}(2)$
Chain 3: $\mathfrak{sp}(4) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$
Chain 4: $G_2 \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$
Chain 5: $\mathfrak{sp}(6) \supset \mathfrak{su}(2)$
Chain 6: $\mathfrak{sp}(6) \supset \mathfrak{su}(2) \times \mathfrak{su}(2)$
Chain 7: $\mathfrak{sp}(6) \supset \mathfrak{sp}(4) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$
Chain 8: $\mathfrak{sp}(6) \supset \mathfrak{sp}(4) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$.

It is now a simple exercise to see what happens when the symmetry breaking process continues beyond $\mathfrak{su}(2)$ in the sense of breaking one or several of the

Table 11a. Branching of the codon representation of $\mathfrak{sp}(6)$ in the surviving chains $\mathfrak{sp}(6) \supset \mathfrak{su}(2)$ (left), $\mathfrak{sp}(6) \supset \mathfrak{su}(2) \times \mathfrak{su}(2)$ (middle), $\mathfrak{sp}(6) \supset \mathfrak{sp}(4) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$ (right).

$\mathfrak{su}(2)$		$\mathfrak{su}(2) \times \mathfrak{su}(2)$		$\mathfrak{sp}(4) \oplus \mathfrak{su}(2)$		$\mathfrak{su}(2) \oplus \mathfrak{su}(2)$	
$2s$	d	$2s_1 - 2s_2$	d	Highest Weight	d	$2s_1 - 2s_2$	d
13	14	$3-4$	20	$(2,0)-1$	20	$6-1$	14
11	12	$1-6$	14			$2-1$	6
9	10	$3-2$	12	$(1,1)-0$	16	$7-0$	8
7	8	$1-4$	10			$5-0$	6
7	8	$1-2$	6			$1-0$	2
5	6	$1-0$	2	$(1,0)-2$	12	$3-2$	12
3	4			$(0,1)-1$	10	$4-1$	10
1	2			$(1,0)-0$	4	$3-0$	4
				$(0,0)-1$	2	$0-1$	2
8 subspaces		6 subspaces		6 subspaces		9 subspaces	

Table 11b. Branching of the codon representation of $\mathfrak{sp}(6)$ in the surviving chain $\mathfrak{sp}(6) \supset \mathfrak{sp}(4) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$.

$\mathfrak{sp}(4) \oplus \mathfrak{su}(2)$		$\mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$	
Highest Weight	d	$2s_1 - 2s_2 - 2s_3$	d
$(2,0) - 1$	20	$1 - 1 - 1$	8
		$2 - 0 - 1$	6
		$0 - 2 - 1$	6
$(1,1) - 0$	16	$2 - 1 - 0$	6
		$1 - 2 - 0$	6
		$1 - 0 - 0$	2
		$0 - 1 - 0$	2
$(1,0) - 2$	12	$1 - 0 - 2$	6
		$0 - 1 - 2$	6
$(0,1) - 1$	10	$1 - 1 - 1$	8
		$0 - 0 - 1$	2
$(1,0) - 0$	4	$1 - 0 - 0$	2
		$0 - 1 - 0$	2
$(0,0) - 1$	2	$0 - 0 - 1$	2
6 subspaces		14 subspaces	

$\mathfrak{su}(2)$-subalgebras down to Abelian $\mathfrak{u}(1)$-subalgebras. First of all, at least one of the $\mathfrak{su}(2)$-subalgebras must remain unbroken since otherwise, we just obtain 64 singlets: this immediately excludes chains 1, 2 and 5. Chain 3 is excluded since breaking one of the two $\mathfrak{su}(2)$-subalgebras without breaking the other produces 24 multiplets: 2 quintets, 4 quartets, 6 triplets, 8 doublets and 4 singlets. Chains 4, 6 and 7 are excluded since

Chain 4: breaking the first $\mathfrak{su}(2)$-subalgebra without breaking the second produces 18 multiplets: 2 sextets, 4 quintets, 2 quartets, 4 triplets and 6 doublets,
breaking the second $\mathfrak{su}(2)$-subalgebra without breaking the first produces 30 multiplets: 2 quartets, 8 triplets, 12 doublets and 8 singlets,

Chain 6: breaking the first $\mathfrak{su}(2)$-subalgebra without breaking the second produces 16 multiplets: 2 septets, 6 quintets, 6 triplets and 2 singlets,
breaking the second $\mathfrak{su}(2)$-subalgebra without breaking the first produces 24 multiplets: 8 quartets and 16 doublets,

Chain 7: breaking the first $\mathfrak{su}(2)$-subalgebra without breaking the second produces 40 multiplets: 4 triplets, 16 doublets and 20 singlets,
breaking the second $\mathfrak{su}(2)$-subalgebra without breaking the first produces 15 multiplets: 1 octet, 2 septets, 1 sextet, 2 quintets, 4 quartets, 2 triplets, 1 doublet and 2 singlets;

all these schemes are remote from the genetic code. Finally, chain 8 is excluded since breaking one of the three $\mathfrak{su}(2)$-subalgebras without breaking the other two produces 24 multiplets: 2 sextets, 4 quartets, 4 triplets, 10 doublets and 4 singlets.

Summarizing, we arrive at the first fundamental result of our investigation:

> *There is no symmetry breaking pattern through chains of subalgebras capable of reproducing exactly the degeneracies of the genetic code.*

We already note at this point that this statement continues to hold true even when the simple Lie algebras of medium rank, $\mathfrak{so}(13)$ and $\mathfrak{so}(14)$, are included and even when diagonal chains are allowed.

At first sight, this negative result seems to be a fatal strike against the programmatic claim of the algebraic approach to the genetic code. However, as will be shown next, there is a generalization of the symmetry breaking procedure discussed above which does provide a positive answer: it is based on the introduction of certain Casimir-like operators associated with chains of subalgebras and allows to incorporate, in rigorous mathematical terms, the phenomenon of a (partial) *freezing* of the symmetry breaking process during the last step, in agreement with the freezing in the evolution of the genetic code postulated by biologists and geneticists (see Sec. 4).

7.3. *Symmetry breaking, Hamiltonians and freezing*

Given a semisimple Lie algebra $\mathfrak{g}$ together with a descending chain

$$\mathfrak{g} \supset \mathfrak{g}_1 \supset \mathfrak{g}_2 \supset \cdots \tag{87}$$

of semisimple subalgebras $\mathfrak{g}_1, \mathfrak{g}_2, \ldots$, the distribution of multiplets obtained by successively decomposing a given irreducible representation of $\mathfrak{g}$ can be encoded into the spectrum of a single operator H: it can be defined as a generic linear combination of the Casimir operators C_j of the simple subalgebras of $\mathfrak{g}$ which constitute the semisimple subalgebras $\mathfrak{g}_1, \mathfrak{g}_2, \ldots$ appearing in the chain:

$$H = \sum_j \lambda_j C_j \,. \tag{88}$$

Indeed, due to the inclusion relations between the $\mathfrak{g}_j$, the Casimir operators C_j commute among themselves, and for a generic choice of the coefficients λ_j, the eigenspaces of the operator H coincide with the joint eigenspaces of the set of Casimir operators C_j, which in turn are just the irreducible subspaces for the smallest (last) subalgebra in the chain. Note that in applications to physics such as, e.g. in the vibron model, H is nothing but the Hamiltonian of the system, and the coefficients λ_j are determined by fitting to the experimentally observed energy spectrum. We shall therefore continue to call this operator the Hamiltonian, even though in applications to biology, it is not necessarily associated with the notion of energy.

To be somewhat more specific, recall that after the first phase of the symmetry breaking process has been completed, the last subalgebra in the chain is a direct sum of $\mathfrak{su}(2)$-subalgebras, so that the Hamiltonian H associated with this stage can be written in the form

$$H = \sum_j \lambda_j C_j + \sum_{k=1}^{p} \alpha_k \boldsymbol{L}_k^2 \,, \tag{89}$$

where p, the total number of $\mathfrak{su}(2)$-subalgebras appearing at the end of the chain, lies between 1 and the rank of $\mathfrak{g}$, depending on the chain. The C_j are now the Casimir operators associated with the simple subalgebras $\neq \mathfrak{su}(2)$ that constitute the semisimple subalgebras $\mathfrak{g}_1, \mathfrak{g}_2, \ldots$ appearing in the chain, whereas $\boldsymbol{L}_k^2 = L_{k,x}^2 + L_{k,y}^2 + L_{k,z}^2$ is the standard Casimir operator of the kth $\mathfrak{su}(2)$-subalgebra ($1 \leq k \leq p$).

The second phase, which involves breaking one or several of the $\mathfrak{su}(2)$-subalgebras, will be implemented by a Hamiltonian of the form

$$H = \sum_j \lambda_j C_j + \sum_{k=1}^{p} \alpha_k \boldsymbol{L}_k^2 + \sum_{k=1}^{p} \beta_k L_{k,z}^2 + \sum_{k=1}^{p} \gamma_k L_{k,z} \,. \tag{90}$$

To explain the effect of the new terms, we consider the simplest case of a single copy of $\mathfrak{su}(2)$. Taking into account the fact that the irreducible representations of $\mathfrak{su}(2)$, characterized by their spin s (corresponding to the highest weight $2s$, which may take any non-negative integer value), form $(2s+1)$-dimensional spaces on which the standard Casimir operator $\boldsymbol{L}^2$ of $\mathfrak{su}(2)$ takes the value $s(s+1)$, whereas the operator L_z has $2s+1$ distinct eigenvalues $m = -s, \ldots, s$, we see that

(a) the operator L_z provides a complete splitting of a $(2s+1)$-dimensional multiplet into $2s+1$ singlets, whereas
(b) the operator L_z^2 provides a softer splitting of a $(2s+1)$-dimensional multiplet into s doublets and one singlet if s is integer, or s doublets if s is half-integer.

Only the first possibility corresponds to a genuine symmetry breaking at the level of Lie algebras: from the Lie algebra $\mathfrak{su}(2)$ to its maximal subalgebra $\mathfrak{u}(1)$. It has however been observed[57] that both possibilities allow for a natural interpretation in terms of a genuine symmetry breaking at the level of Lie groups, namely from the (connected) group SU(2) down to (a) its maximal connected subgroup $\mathrm{U}(1) \cong \mathrm{SO}(2)$ or (b) down to its maximal (non-connected) subgroup $\mathbb{Z}_2 \times \mathrm{U}(1) \cong \mathrm{O}(2)$, a subgroup of SU(2) formed by two circles:

$$\mathbb{Z}_2 \times \mathrm{U}(1) = \left\{ \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} \Bigg/ \alpha \in \mathbb{R} \right\} \cup \left\{ \begin{pmatrix} 0 & e^{i\beta} \\ e^{-i\beta} & 0 \end{pmatrix} \Bigg/ \beta \in \mathbb{R} \right\} . \tag{91}$$

Note that $\mathbb{Z}_2 \times \mathrm{U}(1)$ is generated by U(1) together with the single matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

which is nothing but the generator of the Weyl group of SU(2) and which in the spin s representation connects the two states of magnetic quantum number m and $-m$. By abuse of notation, we shall in the following refer to these two reductions of $\mathfrak{su}(2)$-symmetry, in abbreviated form, as $\mathfrak{so}(2)$-symmetry and $\mathfrak{o}(2)$-symmetry, respectively (even though "SO(2)-symmetry" and "O(2)-symmetry" would be more appropriate).

The last important ingredient in the symmetry breaking process is that we allow the coefficients γ_k in the Hamiltonian H given above to be polynomials in the standard Casimir operators $\boldsymbol{L}_k^2$ of the $\mathfrak{su}(2)$-subalgebras, rather than just constants. This allows for a very specific form of interrupting the symmetry breaking process in the last step, because multiplets of the penultimate stage that would normally subdivide in the last step will remain unbroken if (and only if) their $\mathfrak{su}(2)$ labels are such that the corresponding γ-coefficient vanishes. Such a partial interruption in the differentiation of the genetic code is in accordance with its *freezing* into the presently observed form as proposed by biologists and geneticists (see the discussion in Sec. 4.2), provided the phenomenon is assumed to have occurred exclusively during the last step of the symmetry breaking. Concrete examples will be given below.

7.3.1. *Breaking the* $\mathfrak{su}(2)$*-symmetries: Part 2*

With these remarks and observations out of the way, we proceed to reconsider the second phase of the symmetry breaking process, during which one or several of the $\mathfrak{su}(2)$-subalgebras are broken. The main novelty is that this can now be done in two ways: either

- down to $\mathfrak{o}(2)$, using the operator L_z^2, or
- down to $\mathfrak{so}(2)$, using the operator L_z.

In the second case, the degeneracy is completely lifted, whereas the first case yields only doublets in the even-dimensional (half-integer spin) representations of $\mathfrak{su}(2)$ and a collection of doublets plus one singlet in the odd-dimensional (integer spin) representations of $\mathfrak{su}(2)$: we shall in the following label the singlets by $2m$ ($m = -s, \ldots, s$) and the doublets by $\pm 2m$ ($m = 0, \ldots, s$, with $m > 0$), where m is the magnetic quantum number. Moreover, we note that the option of breaking down to $\mathfrak{so}(2)$ can be performed in two different ways: either

- by *direct breaking*, i.e. directly from $\mathfrak{su}(2)$ down to $\mathfrak{so}(2)$, or
- by *indirect breaking*, i.e. from $\mathfrak{su}(2)$ down to $\mathfrak{o}(2)$ in a first step and then from $\mathfrak{o}(2)$ down to $\mathfrak{so}(2)$ in a second step.

As well shall see, these two possibilities can lead to different results, due to the possibility of freezing in the last step. In the original approach,[38,39] the possibility of indirect breaking was disregarded since the intermediate step has no natural

interpretation at the level of Lie algebras: such an interpretation is only possible when global aspects are taken into account, as has first been done in Ref. 57.

As a preliminary remark, we note that the first phase of the symmetry breaking process must have produced multiplets which transform according to the spin 1 (vector) representation of at least one of the $\mathfrak{su}(2)$-subalgebras. Indeed, if this is not the case, we shall not be able to generate the sextets and the triplets of the genetic code. Note also that up to the penultimate step of the process, at least one $\mathfrak{su}(2)$-subalgebra must remain unbroken, because if all $\mathfrak{su}(2)$-subalgebras are broken down to $\mathfrak{o}(2)$ or $\mathfrak{so}(2)$, the multiplicity of all multiplets will be a power of 2. (In the last step, freezing may prevent the disappearance of the multiplets which are not of this type, i.e. the sextets and the triplets.) In fact, the argument concerning the impossibility to generate the sextets and the triplets of the genetic code can be refined by introducing an additional criterion that allows to perform a further significant reduction in the number of chains that have to be analyzed. It is based on the observation that during its second phase, the symmetry breaking process cannot generate multiplets whose dimension is a multiple of 3 out of multiplets whose dimension is not. Therefore, it is convenient to define a *triality number* d_3 as follows:

$$d_3 = \begin{array}{c}\text{sum of the dimensions of all multiplets} \\ \text{whose dimension is a multiple of 3}\end{array} . \tag{92}$$

With this notation, the statement is that as one proceeds along any given chain, the triality number d_3 cannot increase during the second phase. Observing that the value of d_3 in the final distribution of multiplets in the genetic code is 24, we arrive at the conclusion that all surviving chains which, at the beginning or at some other point of the second phase of the symmetry breaking process up to the penultimate step, violate the estimate

$$d_3 \geq 24 \tag{93}$$

will not be able to generate the sextets and the triplets of the genetic code and may therefore be discarded.

As an example, consider the chains 1–8 listed at the beginning of Sec. 7.2.2. All of them have, up to the end of the first phase, produced less than 21 multiplets, so the symmetry breaking process must proceed into phase 2 and no freezing has occurred so far. Therefore, chain 1 ($d_3 = 0$) and chains 2, 5 and 6 ($d_3 = 18$) are immediately eliminated. For the remaining chains, we argue as follows, discarding without further notice all chains with less than 21 multiplets and $d_3 < 24$.

- *Chain 3*: $\mathfrak{sp}(4) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$, 8 multiplets, $d_3 = 36$; see Table 9.
  Note first the symmetry of the distribution of multiplets under exchange of the two $\mathfrak{su}(2)$-subalgebras. Now observe that

  (1) breaking, say, the second $\mathfrak{su}(2)$ to $\mathfrak{o}(2)$ generates 14 multiplets with $d_3 = 18$,
  (2) breaking, say, the second $\mathfrak{su}(2)$ to $\mathfrak{so}(2)$ generates 24 multiplets with $d_3 = 18$.

In the last case, the symmetry breaking process must terminate, and we must take into account the possibility of freezing. However, the multiplets of dimension > 6 must not be frozen, so we get at least 2 quintets and 4 triplets. Thus both possibilities are eliminated.

- *Chain 4*: $G_2 \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$, 8 multiplets, $d_3 = 39$; see Table 10.
  In this chain, the second $\mathfrak{su}(2)$-subalgebra must be broken, in order to eliminate all multiplets whose dimension is a multiple of 5. Now observe that, in a first step,
  (1) breaking the first $\mathfrak{su}(2)$ to $\mathfrak{o}(2)$ generates 11 multiplets with $d_3 = 24$,
  (2) breaking the first $\mathfrak{su}(2)$ to $\mathfrak{so}(2)$ generates 18 multiplets with $d_3 = 24$, among which there are already 8 odd-dimensional ones (4 quintets and 4 triplets),
  (3) breaking the second $\mathfrak{su}(2)$ to $\mathfrak{o}(2)$ generates 17 multiplets with $d_3 = 24$,
  (4) breaking the second $\mathfrak{su}(2)$ to $\mathfrak{so}(2)$ generates 30 multiplets with $d_3 = 24$.

  In the last case, the symmetry breaking process must terminate, and we must take into account the possibility of freezing. However, the multiplets of dimension > 6 and of dimension 5 must not be frozen, so we get at least 8 triplets, 10 doublets and 5 singlets. Therefore, the only surviving options for continuing the symmetry breaking process are 1 and 3. Hence in a second step,
  (1.1) breaking the first $\mathfrak{o}(2)$ further to $\mathfrak{so}(2)$ generates 18 multiplets with $d_3 = 24$; this gives the same distribution of multiplets as option 2 above,
  (1.2) breaking the second $\mathfrak{su}(2)$ to $\mathfrak{o}(2)$ generates 23 multiplets with $d_3 = 0$,
  (1.3) breaking the second $\mathfrak{su}(2)$ to $\mathfrak{so}(2)$ generates 40 multiplets with $d_3 = 0$,
  (3.1) breaking the first $\mathfrak{su}(2)$ to $\mathfrak{o}(2)$ generates 23 multiplets with $d_3 = 0$,
  (3.2) breaking the first $\mathfrak{su}(2)$ to $\mathfrak{so}(2)$ generates 36 multiplets with $d_3 = 0$,
  (3.3) breaking the second $\mathfrak{o}(2)$ further to $\mathfrak{so}(2)$ generates 30 multiplets with $d_3 = 24$.

  In all cases, the symmetry breaking process must terminate, and we must take into account the possibility of freezing. However, the multiplets of dimension > 6 and of dimension 5 must not be frozen. Therefore, in the cases of options 1.2 and 1.3, we get at must 1 sextet, whereas in the cases of options 3.1, 3.2 and 3.3, we can and must freeze the 3 sextets and 2 triplets coming from the multiplets of dimension 15 and 9 under $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$. In the case of option 3.1, we get 3 sextets, 8 quartets, 2 triplets, 3 doublets and 2 singlets, independently of what other multiplets are frozen. In the case of option 3.2, it suffices to freeze, in addition, the 3 quartets coming from the multiplet of dimension 12 under $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$, as well as the 2 quartets coming from one of the two octets under $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$, whereas in the case of option 3.3, it suffices to freeze, in addition, the 3 quartets coming from the multiplet of dimension 12 under $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$, as well as the doublets coming from the quintet and the triplet under $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$: this will reproduce the distribution of multiplets found in the genetic code, as shown in Tables 13 and 14, with the freezing indicated by a thick line at the point where the interruption is to occur.

- *Chain 7*: $\mathfrak{sp}(6) \supset \mathfrak{sp}(4) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2)$,
  9 multiplets, $d_3 = 24$; see Table 11a.
  In this chain, the first $\mathfrak{su}(2)$-subalgebra must be broken, in order to eliminate all multiplets whose dimension is a multiple of 7 or of 5. Now observe that

  (1) breaking the first $\mathfrak{su}(2)$ to $\mathfrak{o}(2)$ generates 22 multiplets with $d_3 = 12$,
  (2) breaking the first $\mathfrak{su}(2)$ to $\mathfrak{so}(2)$ generates 40 multiplets with $d_3 = 12$,
  (3) breaking the second $\mathfrak{su}(2)$ to $\mathfrak{o}(2)$ generates 10 multiplets with $d_3 = 12$,
  (4) breaking the second $\mathfrak{su}(2)$ to $\mathfrak{so}(2)$ generates 15 multiplets with $d_3 = 12$.

  In the first two cases, the symmetry breaking process must terminate, and we must take into account the possibility of freezing. However, the multiplets of dimension > 6 must not be frozen. In the first case, we get no triplets, whereas in the second case, we get at least 4 triplets, 12 doublets and 8 singlets. Thus all four possibilities are eliminated.
- *Chain 8*: $\mathfrak{sp}(6) \supset \mathfrak{sp}(4) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$,
  14 multiplets, $d_3 = 36$; see Table 11b.
  Note first the symmetry of the distribution of multiplets under arbitrary permutations of the three $\mathfrak{su}(2)$-subalgebras. Now observe that, in a first step,

  (1) breaking, say, the second $\mathfrak{su}(2)$ to $\mathfrak{o}(2)$ generates 16 multiplets with $d_3 = 24$,
  (2) breaking, say, the second $\mathfrak{su}(2)$ to $\mathfrak{so}(2)$ generates 24 multiplets with $d_3 = 24$.

  In the last case, the symmetry breaking process must terminate, and we must take into account the possibility of freezing. But whatever we do, we get at least 4 sextets and no triplets. Therefore, the only surviving option for continuing the symmetry breaking process is 1. Hence in a second step,

  (1.1) breaking the second $\mathfrak{o}(2)$ further to $\mathfrak{so}(2)$ generates 24 multiplets with $d_3 = 24$,
  (1.2) breaking, say, the third $\mathfrak{su}(2)$ to $\mathfrak{o}(2)$ generates 18 multiplets with $d_3 = 12$,
  (1.3) breaking, say, the third $\mathfrak{su}(2)$ to $\mathfrak{so}(2)$ generates 27 multiplets with $d_3 = 12$.

  In the first and in the last case, the symmetry breaking process must terminate, and we must take into account the possibility of freezing. However, the multiplets of dimension > 6 must not be frozen. In the case of option 1.1, we again get at least 4 sextets and no triplets. In the case of option 1.3, it suffices to freeze the two sextets under $\mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$ that would otherwise break into doublets and the two doublets under $\mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$ that would otherwise break into singlets: this will reproduce the distribution of multiplets found in the genetic code, as shown in Table 15, with the freezing indicated by a thick line at the point where the interruption is to occur.

With this result, our search for symmetries using the simple Lie algebras of low rank is complete. We have identified three chains which give rise to the degeneracies of the genetic code, namely the $\mathfrak{sp}(6)$-chain first found by the authors of Refs. 38 and 39,

$$\mathfrak{sp}(6) \supset \mathfrak{sp}(4) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$$
$$\supset \mathfrak{su}(2) \oplus \mathfrak{o}(2) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{o}(2) \oplus \mathfrak{so}(2)^{\mathrm{f}}\,, \quad (94)$$

the G_2-chain

$$G_2 \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{o}(2) \supset \mathfrak{so}(2)^{\mathrm{f}} \oplus \mathfrak{o}(2)\,, \quad (95)$$

and finally the G_2-chain first found by the authors of Ref. 57,

$$G_2 \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{o}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{so}(2)^{\mathrm{f}}\,. \quad (96)$$

Here, the superscript "f" indicates the fact that the last breaking is only partial, due to the freezing phenomenon, which can be implemented by the Hamiltonians

$$H = H_0 + \lambda C_2(\mathfrak{sp}(4)) + \alpha_1 \boldsymbol{L}_1^2 + \alpha_2 \boldsymbol{L}_2^2 + \alpha_3 \boldsymbol{L}_3^2$$
$$+ \beta_2 L_{2,z}^2 + \gamma_3 (\boldsymbol{L}_1^2 + \boldsymbol{L}_2^2)(\boldsymbol{L}_3^2 - 2) L_{3,z} \quad (97)$$

for the $\mathfrak{sp}(6)$-chain,

$$H = H_0 + \alpha_1 \boldsymbol{L}_1^2 + \alpha_2 \boldsymbol{L}_2^2 + \beta_2 L_{2,z}^2$$
$$+ \gamma_1 (\boldsymbol{L}_2^2 - 2) \left(\boldsymbol{L}_2^2 - \frac{15}{4} \right) (\boldsymbol{L}_2^2 - 6) \left(\boldsymbol{L}_2^2 - \frac{35}{4} \right) L_{1,z} \quad (98)$$

for the first G_2-chain and

$$H = H_0 + \alpha_1 \boldsymbol{L}_1^2 + \alpha_2 \boldsymbol{L}_2^2 + \beta_2 L_{2,z}^2$$
$$+ \gamma_2 (\boldsymbol{L}_2^2 - 2)(\boldsymbol{L}_2^2 - 6) \left(\boldsymbol{L}_2^2 - \frac{35}{4} \right) L_{2,z} \quad (99)$$

for the second G_2-chain. Note that according to the terminology introduced above, the first two chains correspond to "direct breaking" and the third chain to "indirect breaking".

Finally, let us comment on the problem of including the simple Lie algebras of medium rank, $\mathfrak{so}(13)$ and $\mathfrak{so}(14)$, in the search. The strategy is the one outlined above, but the details of this tedious investigation, which requires analyzing hundreds of chains, are too long to be presented here and will be reported elsewhere.[56] The first result is that no new possibilities arise from $\mathfrak{so}(13)$. The second result is that there emerges a unique simple chain starting out from $\mathfrak{so}(14)$ that reproduces the distribution of multiplets found in the genetic code, which is the following:

$$\mathfrak{so}(14) \supset \mathfrak{so}(7) \oplus \mathfrak{so}(7) \supset \mathfrak{so}(7) \oplus G_2$$
$$\supset G_2 \oplus G_2 \supset G_2 \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$$
$$\supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$$
$$\supset \mathfrak{so}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{o}(2) \supset \mathfrak{so}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{so}(2)^{\mathrm{f}}\,. \quad (100)$$

The corresponding branching rules are shown in Table 12 (for phase 1) and Table 16 (for phase 2). Note that according to the terminology introduced above, this chain

Table 12. Branching of the codon representation of $\mathfrak{so}(14)$ in the surviving chain $\mathfrak{so}(14) \supset \mathfrak{so}(7) \oplus \mathfrak{so}(7) \supset G_2 \oplus G_2 \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$ (abbreviated).

$\mathfrak{so}(7) \oplus \mathfrak{so}(7)$		$G_2 \oplus G_2$		$\mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$	
Highest Weight	d	Highest Weight	d	$2s_1 - 2s_2 - 2s_3 - 2s_4$	d
$(0,0,1) - (0,0,1)$	64	$(0,1) - (0,1)$	49	$1-1-1-1$	16
				$0-2-1-1$	12
				$1-1-0-2$	12
				$0-2-0-2$	9
		$(0,0) - (0,1)$	7	$0-0-1-1$	4
				$0-0-0-2$	3
		$(0,1) - (0,0)$	7	$1-1-0-0$	4
				$0-2-0-0$	3
		$(0,0) - (0,0)$	1	$0-0-0-0$	1
1 subspace		4 subspaces		9 subspaces	

corresponds to "indirect breaking". The third result is that when diagonal chains are included, there emerges a handful of additional chains starting out from $\mathfrak{so}(14)$ that also reproduce the distribution of multiplets found in the genetic code. For more details, we refer to Ref. 56.

At any rate, the $\mathfrak{sp}(6)$-model has been the first model for the genetic code to be discovered and in our view it continues to be the most natural candidate for a biological interpretation. This is the reason why we shall conclude this review with a more detailed exposition of that model.

8. The $\mathfrak{sp}(6)$ Model

8.1. *The codon representation of* $\mathfrak{sp}(6)$

The codon representation of $\mathfrak{sp}(6)$ is, by definition, the 64-dimensional irreducible representation of $\mathfrak{sp}(6)$ of highest weight (1, 1, 0) (in the basis formed by the fundamental weights). In terms of the orthogonal basis $\{e_1, e_2, e_3\}$ introduced above (which is the one we shall use in the following to describe the coordinates of roots and weights of $\mathfrak{sp}(6)$, except when explicitly mentioned otherwise), the highest weight is (2, 1, 0) and the nine positive roots of $\mathfrak{sp}(6)$ are

$$(2,0,0), (0,2,0), (0,0,2) \quad \text{(long roots)}\,, \tag{101}$$

$$(1,\pm 1,0), (1,0,\pm 1), (0,1,\pm 1) \quad \text{(short roots)}\,. \tag{102}$$

Calculating their scalar product with the highest weight (2, 1, 0) gives the length $m(\alpha)$ of the corresponding α-string of weights through (2, 1, 0), and subtraction

Table 13. Branching of the codon representation of G_2 in the chain $G_2 \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{o}(2) \supset \mathfrak{so}(2) \oplus \mathfrak{o}(2)$.

$\mathfrak{su}(2) \oplus \mathfrak{su}(2)$		$\mathfrak{su}(2) \oplus \mathfrak{o}(2)$		$\mathfrak{so}(2) \oplus \mathfrak{o}(2)$	
$2s_1 - 2s_2$	d	$2s_1 - 2m_2$	d	$2m_1 - 2m_2$	d
$2 - 4$	15	$2 - (\pm 4)$	6	$(+2) - (\pm 4)$	2
				$(-2) - (\pm 4)$	2
				$0 - (\pm 4)$	2
		$2 - (\pm 2)$	6	$(+2) - (\pm 2)$	2
				$(-2) - (\pm 2)$	2
				$0 - (\pm 2)$	2
		$2 - 0$	3	$(+2) - 0$	1
				$(-2) - 0$	1
				$0 - 0$	1
$1 - 5$	12	$1 - (\pm 5)$	4	$(+1) - (\pm 5)$	2
				$(-1) - (\pm 5)$	2
		$1 - (\pm 3)$	4	$(+1) - (\pm 3)$	2
				$(-1) - (\pm 3)$	2
		$1 - (\pm 1)$	4	$(+1) - (\pm 1)$	2
				$(-1) - (\pm 1)$	2
$2 - 2$	9	$2 - (\pm 2)$	6	$(+2) - (\pm 2)$	2
				$(-2) - (\pm 2)$	2
				$0 - (\pm 2)$	2
		$2 - 0$	3	$(+2) - 0$	1
				$(-2) - 0$	1
				$0 - 0$	1
$3 - 1$	8	$3 - (\pm 1)$	8	$(+3) - (\pm 1)$	2
				$(-3) - (\pm 1)$	2
				$(+1) - (\pm 1)$	2
				$(-1) - (\pm 1)$	2
$1 - 3$	8	$1 - (\pm 3)$	4	$(+1) - (\pm 3)$	2
				$(-1) - (\pm 3)$	2
		$1 - (\pm 1)$	4	$(+1) - (\pm 1)$	2
				$(-1) - (\pm 1)$	2
$0 - 4$	5	$0 - (\pm 4)$	2	$0 - (\pm 4)$	2
		$0 - (\pm 2)$	2	$0 - (\pm 2)$	2
		$0 - 0$	1	$0 - 0$	1
$1 - 1$	4	$1 - (\pm 1)$	4	$(+1) - (\pm 1)$	2
				$(-1) - (\pm 1)$	2
$0 - 2$	3	$0 - (\pm 2)$	2	$0 - (\pm 2)$	2
		$0 - 0$	1	$0 - 0$	1
8 subspaces		17 subspaces		36 subspaces	

Table 14. Branching of the codon representation of G_2 in the chain $G_2 \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{o}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{so}(2)$.

$\mathfrak{su}(2) \oplus \mathfrak{su}(2)$		$\mathfrak{su}(2) \oplus \mathfrak{o}(2)$		$\mathfrak{su}(2) \oplus \mathfrak{so}(2)$	
$2s_1 - 2s_2$	d	$2s_1 - 2m_2$	d	$2s_1 - 2m_2$	d
$2-4$	15	$2-(\pm4)$	6	$2-(+4)$	3
				$2-(-4)$	3
		$2-(\pm2)$	6	$2-(+2)$	3
				$2-(-2)$	3
		$2-0$	3	$2-0$	3
$1-5$	12	$1-(\pm5)$	4	$1-(+5)$	2
				$1-(-5)$	2
		$1-(\pm3)$	4	$1-(+3)$	2
				$1-(-3)$	2
		$1-(\pm1)$	4	$1-(+1)$	2
				$1-(-1)$	2
$2-2$	9	$2-(\pm2)$	6	$2-(+2)$	3
				$2-(-2)$	3
		$2-0$	3	$2-0$	3
$3-1$	8	$3-(\pm1)$	8	$3-(+1)$	4
				$3-(-1)$	4
$1-3$	8	$1-(\pm3)$	4	$1-(+3)$	2
				$1-(-3)$	2
		$1-(\pm1)$	4	$1-(+1)$	2
				$1-(-1)$	2
$0-4$	5	$0-(\pm4)$	2	$0-(+4)$	1
				$0-(-4)$	1
		$0-(\pm2)$	2	$0-(+2)$	1
				$0-(-2)$	1
		$0-0$	1	$0-0$	1
$1-1$	4	$1-(\pm1)$	4	$1-(+1)$	2
				$1-(-1)$	2
$0-2$	3	$0-(\pm2)$	2	$0-(+2)$	1
				$0-(-2)$	1
		$0-0$	1	$0-0$	1
8 subspaces		17 subspaces		30 subspaces	

Table 15. Branching of the codon representation of $\mathfrak{su}(6)$ in the chain $\mathfrak{sp}(6) \supset \mathfrak{sp}(4) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{o}(2) \oplus \mathfrak{su}(2) \supset \mathfrak{su}(2) \oplus \mathfrak{o}(2) \oplus \mathfrak{so}(2)$.

$\mathfrak{sp}(4) \oplus \mathfrak{su}(2)$		$\mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$		$\mathfrak{su}(2) \oplus \mathfrak{o}(2) \oplus \mathfrak{su}(2)$		$\mathfrak{su}(2) \oplus \mathfrak{o}(2) \oplus \mathfrak{so}(2)$	
Highest Weight	d	$2s_1 - 2s_2 - 2s_3$	d	$2s_1 - 2m_2 - 2s_3$	d	$2s_1 - 2m_2 - 2m_3$	d
$(2,0) - 1$	20	$1 - 1 - 1$	8	$1 - (\pm 1) - 1$	8	$1 - (\pm 1) - (+1)$	4
						$1 - (\pm 1) - (-1)$	4
		$2 - 0 - 1$	6	$2 - 0 - 1$	6	$2 - 0 - (+1)$	3
						$2 - 0 - (-1)$	3
		$0 - 2 - 1$	6	$0 - (\pm 2) - 1$	4	$0 - (\pm 2) - (+1)$	2
						$0 - (\pm 2) - (-1)$	2
				$0 - 0 - 1$	2	$0 - 0 - (+1)$	1
						$0 - 0 - (-1)$	1
$(1,1) - 0$	16	$2 - 1 - 0$	6	$2 - (\pm 1) - 0$	6	$2 - (\pm 1) - 0$	6
		$1 - 2 - 0$	6	$1 - (\pm 2) - 0$	4	$1 - (\pm 2) - 0$	4
				$1 - 0 - 0$	2	$1 - 0 - 0$	2
		$1 - 0 - 0$	2	$1 - 0 - 0$	2	$1 - 0 - 0$	2
		$0 - 1 - 0$	2	$0 - (\pm 1) - 0$	2	$0 - (\pm 1) - 0$	2
$(1,0) - 2$	12	$1 - 0 - 2$	6	$1 - 0 - 2$	6	$1 - 0 - (+1)$	2
						$1 - 0 - (-1)$	2
						$1 - 0 - 0$	2
		$0 - 1 - 2$	6	$0 - (\pm 1) - 2$	6	$0 - (\pm 1) - (+2)$	2
						$0 - (\pm 1) - (-2)$	2
						$0 - (\pm 1) - 0$	2
$(0,1) - 1$	10	$1 - 1 - 1$	8	$1 - (\pm 1) - 1$	8	$1 - (\pm 1) - (+2)$	4
						$1 - (\pm 1) - (-2)$	4
		$0 - 0 - 1$	2	$0 - 0 - 1$	2	$0 - 0 - (+1)$	1
						$0 - 0 - (-1)$	1
$(1,0) - 0$	4	$1 - 0 - 0$	2	$1 - 0 - 0$	2	$1 - 0 - 0$	2
		$0 - 1 - 0$	2	$0 - (\pm 1) - 0$	2	$0 - (\pm 1) - 0$	2
$(0,0) - 1$	2	$0 - 0 - 1$	2	$0 - 0 - 1$	2	$0 - 0 - (+1)$	1
						$0 - 0 - (-1)$	1
6 subspaces		14 subspaces		16 subspaces		27 subspaces	

Table 16. Branching of the codon representation of $\mathfrak{so}(14)$ in the chain $\mathfrak{so}(14) \supset \mathfrak{so}(7) \oplus \mathfrak{so}(7) \supset G_2 \oplus G_2 \supset \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \supset \mathfrak{so}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{o}(2) \supset \mathfrak{so}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{so}(2)$ (abbreviated).

$\mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$		$\mathfrak{so}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{o}(2)$		$\mathfrak{so}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{so}(2)$	
$2s_1 - 2s_2 - 2s_3 - 2s_4$	d	$2m_1 - 2s_2 - 2s_3 - 2m_4$	d	$2m_1 - 2s_2 - 2s_3 - 2m_4$	d
$1-1-1-1$	16	$(+1)-1-1-(\pm 1)$	8	$(+1)-1-1-(+1)$	4
				$(+1)-1-1-(-1)$	4
		$(-1)-1-1-(\pm 1)$	8	$(-1)-1-1-(+1)$	4
				$(-1)-1-1-(-1)$	4
$0-2-1-1$	12	$0-2-1-(\pm 1)$	12	$0-2-1-(+1)$	6
				$0-2-1-(-1)$	6
$1-1-0-2$	12	$(+1)-1-0-(\pm 2)$	4	$(+1)-1-0-(+2)$	2
				$(+1)-1-0-(-2)$	2
		$(+1)-1-0-0$	2	$(+1)-1-0-0$	2
		$(-1)-1-0-(\pm 2)$	4	$(-1)-1-0-(+2)$	2
				$(-1)-1-0-(-2)$	2
		$(-1)-1-0-0$	2	$(-1)-1-0-0$	2
$0-2-0-2$	9	$0-2-0-(\pm 2)$	6	$0-2-0-(+2)$	3
				$0-2-0-(-2)$	3
		$0-2-0-0$	3	$0-2-0-0$	3
$0-0-1-1$	4	$0-0-1-(\pm 1)$	4	$0-0-1-(+1)$	2
				$0-0-1-(-1)$	2
$0-0-0-2$	3	$0-0-0-(\pm 2)$	2	$0-0-0-(+2)$	1
				$0-0-0-(-2)$	1
		$0-0-0-0$	1	$0-0-0-0$	1
$1-1-0-0$	4	$(+1)-1-0-0$	2	$(+1)-1-0-0$	2
		$(-1)-1-0-0$	2	$(-1)-1-0-0$	2
$0-2-0-0$	3	$0-2-0-0$	3	$0-2-0-0$	3
$0-0-0-0$	1	$0-0-0-0$	1	$0-0-0-0$	1
9 subspaces		16 subspaces		24 subspaces	

produces the following weights:

α	$m(\alpha)$	$\Lambda-\alpha$	$\Lambda-2\alpha$	$\Lambda-3\alpha$
(2, 0, 0)	2	(0, 1, 0)	$(-2,1,0)$	
(0, 2, 0)	1	$(2,-1,0)$		
(0, 0, 2)	0			
(1, +1, 0)	3	(1, 0, 0)	$(0,-1,0)$	$(-1,-2,0)$
(1, −1, 0)	1	(1, 2, 0)		
(1, 0, +1)	2	(1, 1, −1)	(0, 1, −2)	
(1, 0, −1)	2	(1, 1, +1)	(0, 1, +2)	
(0, 1, +1)	1	(2, 0, −1)		
(0, 1, −1)	1	(2, 0, +1)		

Applying Weyl group transformations produces the full weight diagram for the codon representation of $\mathfrak{sp}(6)$, consisting of 3 Weyl group orbits: the outer shell formed by the 24 weights

$$(\pm 2, \pm 1, 0), (\pm 1, \pm 2, 0), (\pm 2, 0, \pm 1), (\pm 1, 0, \pm 2), (0, \pm 2, \pm 1), (0, \pm 1, \pm 2)\,,$$

of length square 5/2, the middle shell formed by the 8 weights

$$(\pm 1, \pm 1, \pm 1)\,,$$

of length square 3/2 and finally the inner shell formed by the 6 weights

$$(\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1)\,,$$

of length square 1/2. (All signs are to be read independently.)

This diagram is shown geometrically in Fig. 15. Its outer shell forms a truncated octrahedron, which incidentally is one of the few regular polyhedra which can be joined periodically in such a way as to completely fill the three-dimensional space. It has 8 hexagonal faces and 6 square faces. The 24 weights in its vertices are the longest weights in the diagram and have multiplicity 1. In the centers of the hexagonal faces we have the 8 weights of medium length; they have multiplicity 2. Within the truncated octahedron there is a smaller regular octahedron whose 6 vertices are formed by the shortest weights in the diagram; they have multiplicity 4.

The decomposition of the codon representation of $\mathfrak{sp}(6)$ along the canonical chain

$$\mathfrak{sp}(6) \supset \mathfrak{sp}(4) \oplus \mathfrak{su}(2)_3 \supset \mathfrak{su}(2)_1 \oplus \mathfrak{su}(2)_2 \oplus \mathfrak{su}(2)_3 \tag{103}$$

can be understood geometrically by remarking that the root generators of the $\mathfrak{sp}(4)$ subalgebra act in the horizontal directions (parallel to the 1–2-plane) whereas the

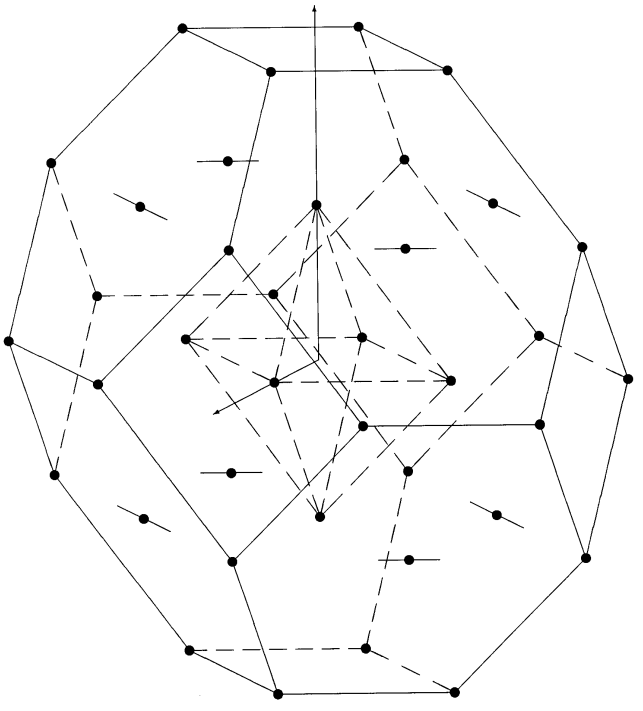


Fig. 15. Weight diagram of the codon representation of $\mathfrak{sp}(6)$.

root generators of the $\mathfrak{su}(2)_k$ subalgebra act along the kth coordinate axis, for $k = 1$, 2, 3. It is therefore convenient to divide the diagram into the five horizontal planes located at $z = +2$, $z = +1$, $z = 0$, $z = -1$, $z = -2$ which are invariant under the $\mathfrak{sp}(4)$ subalgebra but are related by the action of the lowering and raising operators of the $\mathfrak{su}(2)_3$ subalgebra: the top plane at $z + 2$, the bottom plane at $z = -2$ and part of the central plane at $z = 0$ belong to the vector representation, the planes at $z = +1$ and at $z = -1$ to the spinor representation and the rest of the $z = 0$ plane to the scalar representation of $\mathfrak{su}(2)_3$. Following common usage in physics, we shall find it convenient to call representations of integer spin with respect to $\mathfrak{su}(2)_3$ *bosonic* or *tensorial* representations, distinguishing when necessary between *scalar bosons* (spin 0 under $\mathfrak{su}(2)_3$) and *vector bosons* (spin 1 under $\mathfrak{su}(2)_3$), whereas representations of half-integer spin with respect to $\mathfrak{su}(2)_3$ will be referred to as *fermionic* or *spinorial* representations.

With this terminology at our disposal, we can describe the first phase of the symmetry breaking process in the $\mathfrak{sp}(6)$ model, that is, the decomposition of the codon representation of $\mathfrak{sp}(6)$ along the canonical chain (103), as follows: In the first step, the codon representation of $\mathfrak{sp}(6)$ breaks into six irreducible representations of $\mathfrak{sp}(4) \oplus \mathfrak{su}(2)_3$: First, there are two scalar boson representations of $\mathfrak{sp}(4) \oplus \mathfrak{su}(2)_3$, both located in the $z = 0$ plane:

(1) one scalar boson representation of dimension 16, generated by the representation of $\mathfrak{sp}(4)$ with highest weight (1, 1) (it will be called the principal one), and
(2) one scalar boson representation of dimension 4, generated by the first fundamental (or defining) representation of $\mathfrak{sp}(4)$, with highest weight (1, 0).

Next, we have one vector boson representation of $\mathfrak{sp}(4) \oplus \mathfrak{su}(2)_3$, filling up the $z = +2$ and $z = -2$ planes, together with the rest of the $z = 0$ plane, namely

(3) one vector boson representation of dimension $4 \times 3 = 12$, generated by the first fundamental (or defining) representation of $\mathfrak{sp}(4)$, with highest weight (1, 0).

Finally, we must distinguish three fermionic representations of $\mathfrak{sp}(4) \oplus \mathfrak{su}(2)_3$ which, taken together, fill up the $z = +1$ and $z = -1$ planes:

(4) one fermionic representation of dimension $10 \times 2 = 20$, generated by the adjoint representation of $\mathfrak{sp}(4)$, with highest weight (2, 0),
(5) one fermionic representation of dimension $5 \times 2 = 10$, generated by the second fundamental representation of $\mathfrak{sp}(4)$, with highest weight (0, 1), and
(6) one fermionic representation of dimension $1 \times 2 = 2$, generated by the trivial representation of $\mathfrak{sp}(4)$, with highest weight (0, 0).

In Fig. 16 we show the weight diagrams of the irreducible representations of $\mathfrak{sp}(4)$ that appear in this context. (Contrary to the convention adopted in this subsection for roots and weights of $\mathfrak{sp}(6)$, the component expression of highest weights of irreducible representations of $\mathfrak{sp}(4)$ continues to refer to the standard (non-orthonormal) basis formed by the fundamental weights.) Passing to the second step, where the invariance under the generators of $\mathfrak{sp}(4)$ that act diagonally in the 1-2-plane is lost, we see by inspection of Fig. 16 that

(1) the scalar boson representation of dimension 16 breaks into
    (a) one "horizontal" sextet (spin 1 under $\mathfrak{su}(2)_1$ and spin 1/2 under $\mathfrak{su}(2)_2$),
    (b) one "vertical" sextet (spin 1/2 under $\mathfrak{su}(2)_1$ and spin 1 under $\mathfrak{su}(2)_2$),
    (c) one "horizontal" doublet (spin 1/2 under $\mathfrak{su}(2)_1$ and spin 0 under $\mathfrak{su}(2)_2$),
    (d) one "vertical" doublet (spin 0 under $\mathfrak{su}(2)_1$ and spin 1/2 under $\mathfrak{su}(2)_2$),
(2) the scalar boson representation of dimension 4 breaks into
    (a) one "horizontal" doublet (spin 1/2 under $\mathfrak{su}(2)_1$ and spin 0 under $\mathfrak{su}(2)_2$),
    (b) one "vertical" doublet (spin 0 under $\mathfrak{su}(2)_1$ and spin 1/2 under $\mathfrak{su}(2)_2$),
(3) the vector boson representation of dimension 12 breaks into
    (a) one "horizontal" sextet (spin 1/2 under $\mathfrak{su}(2)_1$ and spin 0 under $\mathfrak{su}(2)_2$),
    (b) one "vertical" sextet (spin 0 under $\mathfrak{su}(2)_1$ and spin 1/2 under $\mathfrak{su}(2)_2$),

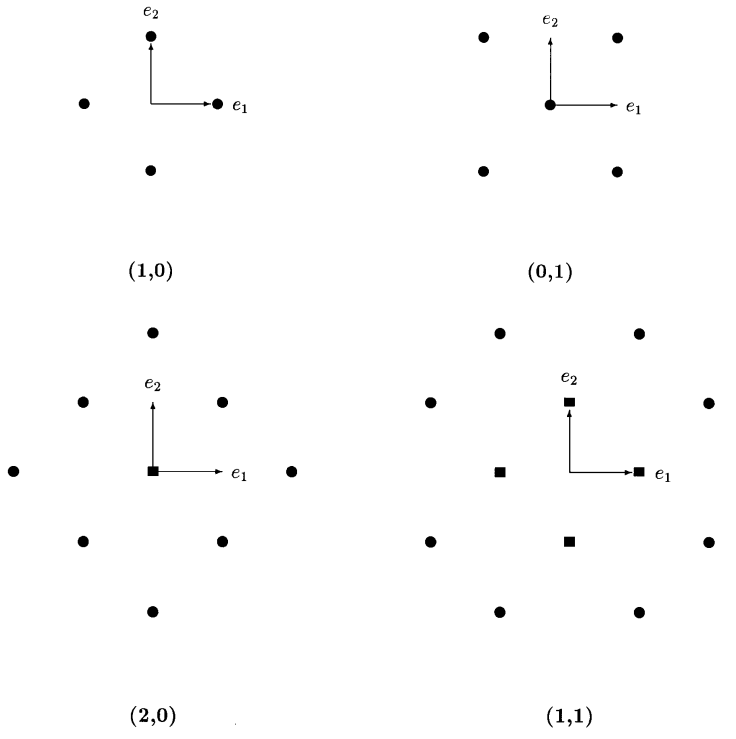


Fig. 16. Weight diagrams of the irreducible representations of $\mathfrak{sp}(4)$ in the decomposition of the codon representation of $\mathfrak{sp}(6)$ along the canonical chain (black circles and black squares represent weights of multiplicity 1 and 2, respectively).

(4) the fermionic representation of dimension 20 breaks into

   (a) one "horizontal" sextet (spin 1 under $\mathfrak{su}(2)_1$ and spin 0 under $\mathfrak{su}(2)_2$),
   (b) one "vertical" sextet (spin 0 under $\mathfrak{su}(2)_1$ and spin 1 under $\mathfrak{su}(2)_2$),
   (c) one octet (spin 1/2 under $\mathfrak{su}(2)_1$ and under $\mathfrak{su}(2)_2$),

(5) the fermionic representation of dimension 10 breaks into

   (a) one octet (spin 1/2 under $\mathfrak{su}(2)_1$ and under $\mathfrak{su}(2)_2$),
   (b) one doublet (spin 0 under $\mathfrak{su}(2)_1$ and under $\mathfrak{su}(2)_2$),

(6) the fermionic representation of dimension 2 remains unbroken,

where the adjectives "horizontal" and "vertical" refer to directions in the 1–2-plane. These branching rules, here derived by pure geometric intuition, are identical with those presented in Table 11b and Table 15; they also appear as part of the information contained in Figs. 17–21.

The second phase of the symmetry breaking process in the $\mathfrak{sp}(6)$ model, which consists in breaking the distribution of multiplets obtained as the result of the first

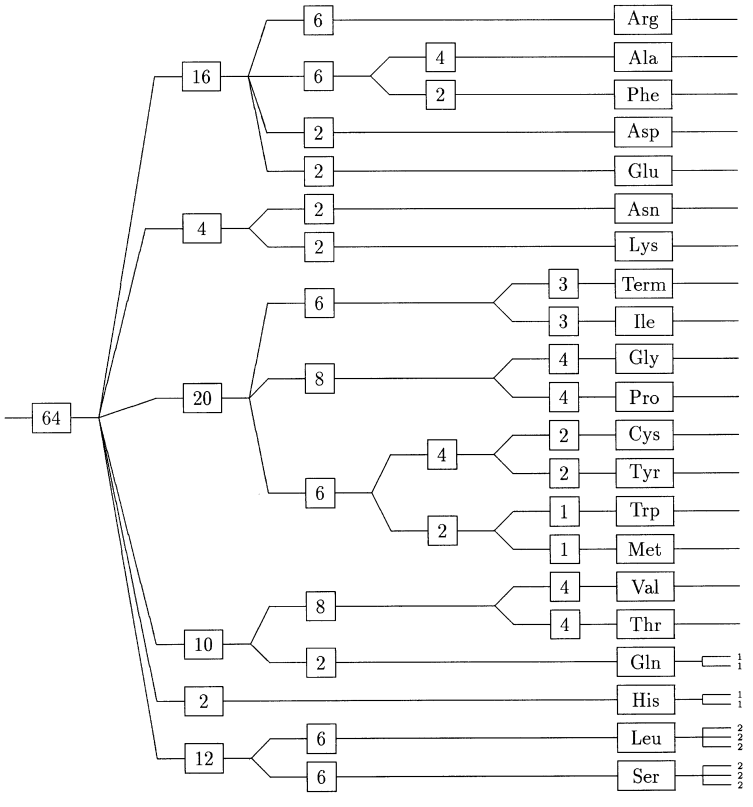


Fig. 17. Evolutionary tree for the genetic code in the $\mathfrak{sp}(6)$ model.

phase along the chain

$$\mathfrak{su}(2)_1 \oplus \mathfrak{su}(2)_2 \oplus \mathfrak{su}(2)_3 \supset \mathfrak{su}(2)_1 \oplus \mathfrak{o}(2)_2 \oplus \mathfrak{su}(2)_3 \supset \mathfrak{su}(2)_1 \oplus \mathfrak{o}(2)_2 \oplus \mathfrak{so}(2)_3^{\mathrm{f}}\,, \tag{104}$$

can be analyzed in exactly the same way. In the first step, only two multiplets are broken, namely the "vertical" sextets coming from the scalar boson representation of dimension 16 and the fermionic representation of dimension 20, since these are the only multiplets that have spin 1 under $\mathfrak{su}(2)_2$: each of them breaks into a quartet and a doublet. In the second step, the multiplets coming from the scalar boson representations remain intact, whereas all others are broken, with the exception of the two sextets coming from the vector boson representation of dimension 12, the doublet coming from the fermionic representation of dimension 10 and the fermionic representation of dimension 2: their breaking is prevented by the freezing phenomenon. Once again, the resulting branching rules are identical with those presented in Table 15; they also appear as part of the information contained in Figs. 17–21.

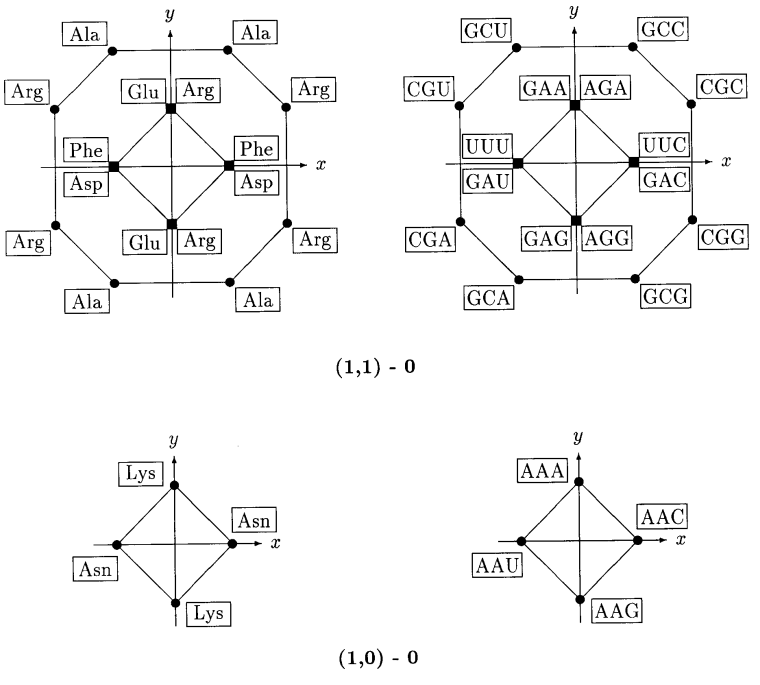
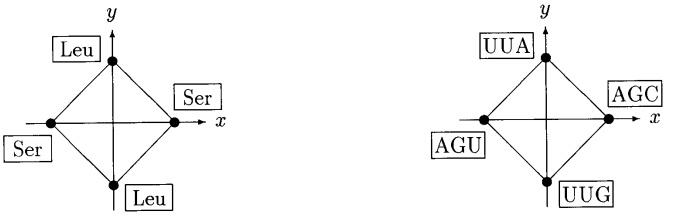


Fig. 18. Amino acid and codon assignments in the codon representation of $\mathfrak{sp}(6)$: the scalar boson sector of $\mathfrak{sp}(4) \oplus \mathfrak{su}(2)$, which lies entirely in the $z = 0$ plane.

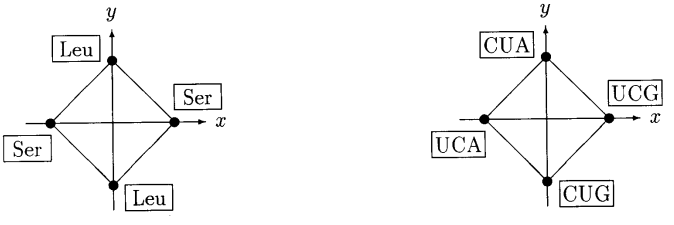
A further prerequisite for a biological interpretation of the model is an appropriate choice of a vector basis in the 64-dimensional codon space, which will be needed whenever it comes to performing explicit calculations. Of course, such a basis should consist of weight vectors, but as mentioned before, this condition is not sufficient, due to the existence of weights of multiplicity > 1. According to the general strategy outlined at the end of Sec. 6.1, an additional requirement that can be imposed is that each basis vector should also belong to a definite irreducible representation of each of the subalgebras in the canonical chain (103), which means that it can be characterized by its highest weight (k_1, k_2) under the $\mathfrak{sp}(4)$ subalgebra and by its highest weight $2s_1$, $2s_2$, $2s_3$ or spin s_1, s_2, s_3 under each of the three $\mathfrak{su}(2)$ subalgebras, apart from and in addition to the components $2m_1$, $2m_2$, $2m_3$ of its usual weight or magnetic quantum numbers m_1, m_2, m_3 with respect to the standard Cartan subalgebra. Despite the missing label problem for the canonical chain in the C-series, mentioned at the end of Sec. 6.1, it turns out that for the codon representation of $\mathfrak{sp}(6)$, this classification is complete: a vector in codon space is fully determined, up to a constant multiple, by the value of these numbers. Thus in the notation familiar from quantum mechanics, these basis vectors can be written as

$$|k_1, k_2; s_1, s_2, s_3; m_1, m_2, m_3\rangle \,. \tag{105}$$



(1,0) - 2: the $z = 0$ plane

(1,0) - 2: the $z = +2$ plane



(1,0) - 2: the $z = -2$ plane

Fig. 19. Amino acid and codon assignments in the codon representation of $\mathfrak{sp}(6)$: the vector boson sector of $\mathfrak{sp}(4) \oplus \mathfrak{su}(2)$.

Loosely speaking, this means that an efficient labelling of the states is obtained if we require each state to be characterized by (a) its position in the weight diagram and (b) its history during the first phase of the symmetry breaking process: each state remembers where it came from.

In passing, we note that the labelling advocated here is by no means the only possible one: at least two other schemes have been proposed and to some extent elaborated in the literature.

One of these methods, due to Weyl,[24] is based on the decomposition of tensor powers of the first fundamental (or defining) representation into their irreducible constituents. It is to be noted that this decomposition uses the representation theory of the permutation group (which acts by permuting the factors in the tensor power), together with the elimination of partial traces; see for example Ref. 44. The central theorem of Weyl states that *all* irreducible representations of the (special) unitary

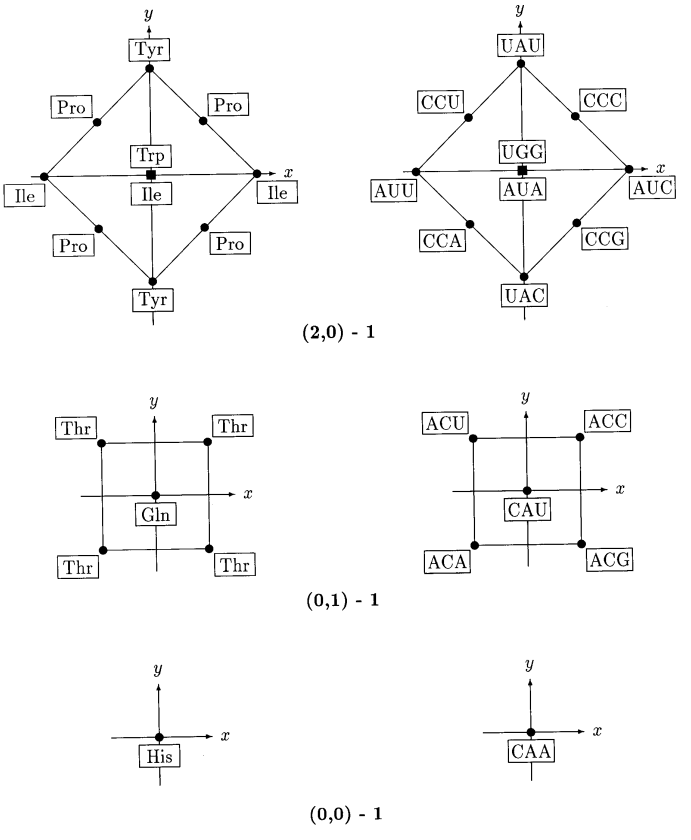


Fig. 20. Amino acid and codon assignments in the codon representation of $\mathfrak{sp}(6)$: the $z = +1$ plane in the fermion sector of $\mathfrak{sp}(4) \oplus \mathfrak{su}(2)$.

algebras and of the symplectic algebras, as well as all "non-spinorial" irreducible representations of the orthogonal algebras, can be obtained through this "tensorial construction". In particular, the codon representation of $\mathfrak{sp}(6)$ is obtained as a piece of the third tensor power of the first fundamental (or defining) representation of $\mathfrak{sp}(6)$, where it appears with multiplicity 2 in the mixed symmetry sector. Indeed, a careful analysis of the corresponding weight diagrams (including the multiplicities) gives the direct decomposition

$$\begin{aligned}(1,0,0) \otimes (1,0,0) \otimes (1,0,0) \\ = ((1,0,0) \oplus (0,0,1)) \oplus (3,0,0) \oplus 2 \times ((1,0,0) \oplus (1,1,0))\,, \end{aligned} \tag{106}$$

where $(1,0,0) \oplus (0,0,1)$ is the totally antisymmetric part, of dimension $6+14 = 20$, $(3,\,0,\,0)$ is the totally symmetric part, of dimension 56, and $2 \times ((1,0,0) \oplus (1,1,0))$ is the mixed symmetry part, of dimension $2(6+64) = 140$. (Note that, contrary to the convention adopted elesewhere in this subsection for roots and weights of $\mathfrak{sp}(6)$, the component expression of these highest weights refers to the standard

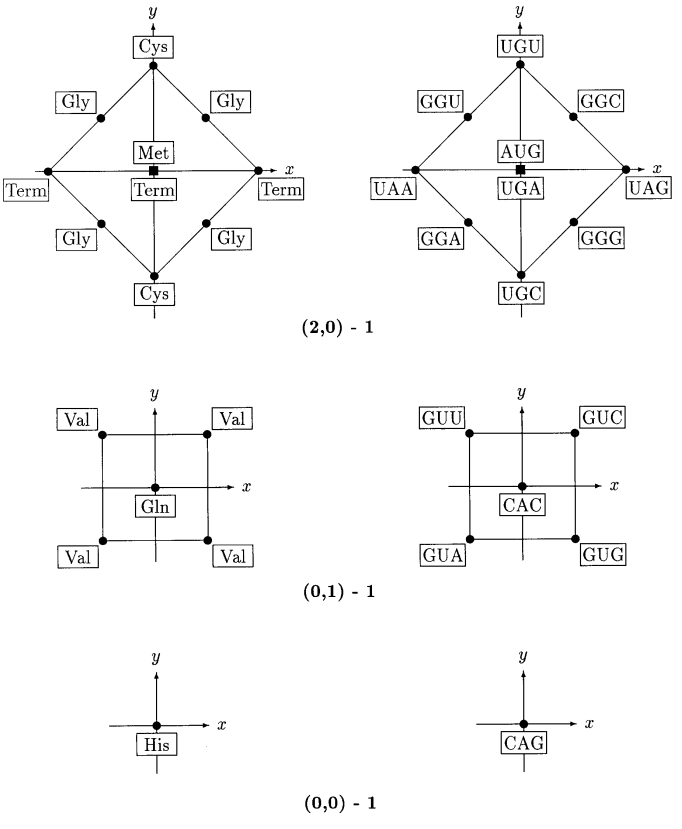


Fig. 21. Amino acid and codon assignments in the codon representation of $\mathfrak{sp}(6)$: the $z = -1$ plane in the fermion sector of $\mathfrak{sp}(4) \oplus \mathfrak{su}(2)$.

(non-orthonormal) basis formed by the fundamental weights.) It can then be shown that the codon representation can be built from the defining representation by forming rank 3 tensors which are

(a) antisymmetric in one pair of indices,
(b) traceless (with respect to the symplectic form) in that same pair of indices and
(c) cyclic.

However, it is not clear what would be the biological interpretation of such a representation of the codon space, in terms of an adequate codon assignment (see below).

The other technique, recently applied to the case of interest by Chacon and Moshinsky,[58] is the boson operator method, where one introduces creation and annihilation operators satisfying canonical commutation relations such that the generators of the symmetry algebra can be written as quadratic polynomials in these operators, whereas the state vectors in any irreducible representation can

be obtained by applying certain polynomials in the creation operators to a unique state, called the vacuum. This elegant and powerful technique, worked out in more detail in Ref. 59, offers as its main advantage the possibility to combine generators of the symmetry algebra and state vectors in irreducible representations into a common algebraic structure. However, the biological interpretation of such a unification has so far not been explored.

8.2. *Amino acid and codon assignment*

The next fundamental problem to be addressed is the assignment of the amino acids to the multiplets resulting from the symmetry breaking scheme and, in a further step, the assignment of the codons to the basis vectors labelled as in (105). *A priori*, the ambiguity in this respect is enormous, because there are no less than 3!5!2!9!2! possibilities to perform rearrangements of amino acids with the same degree of degeneracy, plus another $(6!)^3(4!)^5(3!)^2(2!)^9$ possibilities to perform rearrangements of codons that code for the same amino acid. A first amino acid assignment has already been presented in Ref. 38, based on a best fit between the eigenvalues of the Hamiltonian (97) for each multiplet of states and the Grantham polarity of the corresponding amino acid, but the method is too unspecific to rule out other assignments and the preference for this particular chemical parameter seems arguable. Subsequently, a new amino acid and codon assignment was proposed in Ref. 57, based on a combination of well-established biochemical facts with further global symmetry considerations. The biochemical argument makes use of the WC (Watson–Crick) duality mentioned at the end of Sec. 3.2: given the fact that in codon-anticodon recognition between mRNA and tRNA, C pairs with G and U pairs with A, at least in the first two bases, this version of the WC (Watson–Crick) transformation assigns to every nucleic base X its WC dual nucleic base $X^\dagger$, defined as follows:

$$A^\dagger = U\,, \qquad C^\dagger = G\,, \qquad G^\dagger = C\,, \qquad U^\dagger = A\,. \tag{107}$$

Similarly, every codon XYZ has a WC dual codon $(XYZ)^\dagger = Z^\dagger Y^\dagger X^\dagger$: note the inversion of order which is mathematically compelling and corresponds to the biological fact that the strands in codon-anticodon recognition between mRNA and tRNA, just like the two helices in a DNA molecule, run in opposite directions. Another obvious and useful fact is the weak dependence of the meaning of a codon on the third base, which has led molecular biologists to organize the rules of the genetic code in the (by now standard) form of Table 1, where codons are assembled in family boxes: all codons starting with the same two bases form a family box, and in 8 of the 16 family boxes, they all code for the same amino acid. This has led the authors to introduce a notion of partial WC (Watson–Crick) duality which refers only to the first two bases: thus every codon XYN has a partial WC dual codon defined as $Y^\dagger X^\dagger N$. The codon assignment proposed in Ref. 57 is then obtained by imposing the following two invariance principles:

(a) Principle of family box completeness:
Codons in the same family box (XYN, with $N = U, C, A, G$) are either all bosonic or all fermionic.
(b) Principle of partial WC (Watson–Crick) dual completeness:
For any codon that is bosonic or fermionic, the corresponding partial WC dual codon must also be bosonic or fermionic, respectively.

A simple inspection of the dimensions of the multiplets in the $\mathfrak{sp}(6)$ model (see Table 15) shows that the sextets are bosonic whereas the triplets and singlets are fermionic. Family box completeness then requires the phenylalanine codons to lie in the bosonic sector and the cysteine and tyrosine codons to lie in the fermionic sector, in order to complete the family boxes CGN (arginine), AGN (serine, arginine), UCN (serine), CUN (leucine), UUN (phenylalanine, leucine) in the bosonic sector and UAN (tyrosine, termination), UGN (cysteine, termination, tryptophan), AUN (isoleucine, methionine) in the fermionic sector. Next, partial WC dual completeness forces the codons in the family boxes GAN (the dual of UCN), for aspartic and glutamic acid, and AAN (the dual of UUN), for aspartine and lysine, to belong to the bosonic sector and the codons in the family box CAN (the dual of UGN), for histamine and glutamine, to belong to the fermionic sector. All that remains to be done is the allocation of the five quartets: one of them will be bosonic and the other four will be fermionic. Family box completeness will be fulfilled by any assignment, but partial WC dual completeness requires the bosonic quartet to be self-dual. The quartets are proline (CCN), glycine (GGN), valine (GUN), threonine (ACN) and alanine (GCN), and since only the alanine codons are self-dual, we arrive at the following list of family boxes:

- bosonic: UUN, AAN, CGN, AGN, CUN, UCN, GAN, GCN,
- fermionic: UAN, AUN, UGN, CAN, CCN, GGN, GUN, ACN.

The resulting codon and amino acid assignments are uniquely determined up to a few simple permutations: one choice is shown in Fig. 18–21. Interestingly, one arrives at the same result if, maintaining the requirement of family box completeness, one replaces the condition of partial WC dual completeness by the condition of invariance under permutation of the first two bases.

9. Conclusions and Outlook

In the present paper, we have attempted to explain the biological background and the mathematical basis of the algebraic approach to the genetic code, which aims at explaining its degeneracies as the result of a process of symmetry breaking. This process is supposed to have taken place during the early evolution of the genetic code, before its freezing into the presently observed form. The main result is that when the symmetry concept is based on the mathematical notion of compact Lie groups or (equivalently) of complex semisimple Lie algebras and their representations, it becomes possible to give a complete classification of all possible

schemes capable of reproducing the experimentally observed degeneracies (provided one excludes the uninteresting case of the high rank groups SU(64), SO(64) and Sp(64), from which one can reproduce any pattern of multiplets whatsoever). More specifically, the results of this exhaustive search for symmetries can be summarized as follows.

(1) There is no symmetry breaking pattern through chains of subalgebras capable of reproducing exactly the degeneracies of the genetic code. In other words, the phenomenon of freezing is an essential ingredient, without which there would be no solution.
(2) There are precisely two symmetry breaking patterns through chains of subalgebras which are particularly simple in the sense that they do not involve "indirect breaking" or "diagonal breaking": the $\mathfrak{sp}(6)$ chain (94) and the first G_2 chain (95). Of these, only the first has been considered in Ref. 38 because it requires only a small amount of freezing (out of 16 multiplets present in the penultimate stage, only 4 must be frozen to reduce the 27 multiplets of the final stage to 21), whereas the other requires a large amount of freezing (out of 17 multiplets present in the penultimate stage, no less than 10 must be frozen to reduce the 36 multiplets of the final stage to 21), so one might say that in this case, freezing is more the rule than an accident. Of course, this does not mean that one could not imagine a biological scenario in which the last step of the symmetry breaking could be characterized as a slight breaking (with strong freezing) rather than a strong breaking (with slight freezing), and the question certainly deserves further investigation.
(3) Relaxing the conditions used in Ref. 38, a couple of new possibilities arise. One generalization, first considered in Ref. 57, is to allow what we have here called "indirect breaking". Mathematically, this can be rigorously implemented only at the level of Lie groups and not of Lie algebras: it involves breaking SU(2) to its maximal connected subgroup SO(2) in a two-step procedure, with the maximal subgroup O(2) (which has two connected components) appearing in an intermediate step. Due to the rule that freezing may occur only in the last step, this may give a different result than the direct breaking from SU(2) to SO(2). In this way, two new symmetry breaking patterns appear: the second G_2 chain (96) first mentioned in Ref. 57 and the $\mathfrak{so}(14)$ chain (100). Once again, there is a marked difference in the amount of freezing between the two cases: the G_2 chain (96) requires strong freezing whereas the $\mathfrak{so}(14)$ chain (100) requires weak freezing.
(4) Another generalization that should be included in any classification that claims to be complete is to allow what we have here called "diagonal breaking": this possibility is perfectly legitimate even at the pure Lie algebra level. This part of the search for symmetries has been concluded only recently because it is the most extensive one, requiring the development of new techniques (for the systematic exclusion of entire branches of chains of subalgebras) that have been

presented in Sec. 7 of this paper. The final result is that there appear another seven $\mathfrak{so}(14)$ chains: three of them involve only "direct breaking" whereas the other four also require "indirect breaking". One of the latter is very similar to the $\mathfrak{so}(14)$ chain (100), whereas all others require strong freezing. The details will be presented elsewhere.[56]

Summarizing, we may say that although there are a few other possibilities, we continue to favor the $\mathfrak{sp}(6)$-model, at least for the time being.

Of course, compact Lie groups are not the only class of mathematical objects that can be used to study the concepts of symmetry and symmetry breaking. In fact, there is at least one other very natural setting for the algebraic approach to the genetic code, namely that of finite groups, which describe discrete symmetries, rather than continuous ones. Carrying out the search for symmetries within this context is definitely a highly complex and mathematically ambitious project, and so far there are only preliminary results which rule out the possibility of deriving the degeneracies of the genetic code using the two smallest simple finite groups that admit a codon representation. On the other hand, there are other notions of symmetry that have arisen in mathematical physics during the last three decades and that do not fit into the framework of group theory as we know it from Galois and Lie, since they are described in terms of algebras which generalize the concept of a Lie algebra or of its universal enveloping algebra: these are the supersymmetries and the q-deformed or quantum symmetries, respectively. The representation theory of these algebras is only partially similar to that of ordinary Lie algebras and is considerably less developed, mainly because of the existence of indecomposable representations, i.e. representations that are reducible but are not fully reducible, or in other words, contain irreducible subrepresentations without decomposing into the direct sum of irreducible subrepresentations. (The appearance of such representations for semisimple Lie algebras is excluded by a celebrated theorem of Weyl.) In particular, the dimensions and branching rules can be different from what would be expected. In the case of q-deformed symmetries, such differences appear when the deformation parameter q is a complex root of unity, and not much is presently known about the general representation theory in this case.

For Lie superalgebras, the situation is considerably better if one is willing to accept a couple of technical restrictions. For example, the correct analogue of the class of ordinary simple Lie algebras is a subclass of the class of simple Lie superalgebras called basic classical Lie superalgebras. Similarly, the correct analogue of the class of irreducible representations of ordinary Lie algebras is a subclass of the class of irreducible representations of Lie superalgebras called typical representations. Using the representation theory developed by Kac,[60,61] we have recently derived a complete classification of all typical codon representations of basic classical Lie superalgebras,[62] toghether with a systematic investigation of their symmetry breaking patterns through chains of subalgebras.[63] The subject has also been adressed by various other authors.[64,65]

Another set of questions arises when one wonders what might be the origin of symmetry in the genetic code and of the mechanism that triggers its breakdown. Generally speaking, we believe that the answer must be sought in the theory of dynamical systems with external parameters and of their bifurcations. Such a bifurcation occurs when (at least) one of these parameters is varied so as to cross a critical point, and in dynamical systems with inherent symmetries it is typically accompanied by symmetry breaking.[66] It should also be noted that the symplectic groups are naturally associated with Hamiltonian systems. Therefore, one of the possibilities that comes to mind is to look for a (Hamiltonian) dynamical system with 64 degrees of freedom and with external parameters where the symplectic group Sp(6) appears naturally and the chain of symmetry breakings on which the $\mathfrak{sp}(6)$ model is based results from a sequence of generic bifurcations. As a first step towards this goal, one of the authors has recently calculated the number of external parameters that can be present if the Hamiltonian function is supposed to be an Sp(6)-invariant polynomial of degree ≤ 4 on codon space.[67]

Apart from these open problems of more technical and mathematical nature, there remains as the main challenge the problem of providing a convincing biological interpretation for the symmetry in the genetic code. One important criterion in this respect is whether the model is able to accomodate, in a natural way, the non-standard codes found mainly in organelles such as mitochondria and chloroplasts. In the case of the $\mathfrak{sp}(6)$ model, this does indeed seem to be the case.[68] Another line of thought that should be pursued is to investigate how (if at all) the structure of anticodons and of the biological mediators of transcription, such as the different tRNAs and aminoacyl synthetases, fit into the picture suggested by the symmetry considerations — a difficult task due to the strong species dependence of these entities.

Finally, a word seems in order on the question whether it is or is not sensible to apply sophisticated mathematical techniques to problems in biology in the way we have done this here. It is often criticized that models derived by purely mathematical arguments are too far from the biological reality and that living systems are too complex to allow for precise mathematical modelling. The authors feel that statements of this kind are more the expression of a personal point of view than an academic argument and that they cannot be proved or disproved in such generality. A reliable answer can only come from down-to-earth investigation of concrete models. On the other hand, there have recently been various attempts in the literature to build mathematical models for the evolution of the genetic code which are also based on symmetry considerations but claim to stay closer to the biological facts. In this respect, we would like to stress that the approach advocated in Ref. 38 does not at all remain far from the biological facts. Quite to the contrary, it is fully compatible not only with the obvious properties of the genetic code table but also with the fine structure that resides in the deviations. In fact, we feel that the standard biological facts such as the Watson–Crick pairing rules and the Crick wobbling rules, leading to the reduced importance of the third nucleic base in codon-anticodon pairing and,

as a result, to the predominant organization of the genetic code table into family boxes, are fully understood and hardly need a sophisticated mathematical model for their explanation. The whole point of our approach is to concentrate on the deviations and, rather than looking at them as purely accidental, take them as serious hints towards an underlying fine structure that waits to be unravelled. History will show whether such a fine structure exists or not, but in our view it does not make sense to combine the algebraic approach with allowing, e.g. the sextets to result from the combination of a quartet and a doublet or the triplets to result from the combination of a doublet and a singlet. Recombination is about the only thing that is rigorously forbidden in the process of symmetry breaking, so biological processes where such a recombination occurs (if they exist) are strictly outside the scope of our approach. Moreover, if one allows for changes in the number of sextets, triplets or singlets that appear in the genetic code, the entire picture changes drastically. For example, a primordial breaking of the 64-dimensional codon representation into 16 quartets can be achieved with many different types of symmetry and along many different chains: such an approach would be practically devoid of predictive power.

References

1. J. D. Watson and F. H. C. Crick, *Nature* **171**, 737 (1953).
2. J. D. Watson and F. H. C. Crick, *Nature* **171**, 964 (1953).
3. F. H. C. Crick, L. Barnett, S. Brenner and R. J. Watts-Tobin, *Nature* **192**, 1227 (1961).
4. M. W. Nirenberg and J. H. Matthaei, *Proc. Nat. Acad. Sci. USA* **47**, 1588 (1961).
5. P. Lengyel, J. F. Speyer and S. Ochoa, *Proc. Nat. Acad. Sci. USA* **47**, 1936 (1961).
6. M. W. Nirenberg, *Sci. Am.* **208**, 3, 80 (1963).
7. M. W. Nirenberg and P. Leder, *Science* **145**, 1399 (1964).
8. H. P. Ghosh, D. Söll and H. G. Khorana, *J. Mol. Biol.* **25**, 275 (1967).
9. S. Weinberg, *The First Three Minutes* (Basic Books, New York, 1977).
10. M. Eigen, B. F. Lindemann, M. Tietze, R. Winkler-Oswatitsch, A. Dress and A. von Haeseler, *Science* **244**, 673 (1989).
11. R. Grantham, *Science* **185**, 862 (1974).
12. O. T. Avery, C. M. MacLeod and M. McCarty, *J. Exp. Med.* **79**, 137 (1944).
13. F. H. C. Crick, *J. Mol. Biol.* **19**, 548 (1966).
14. F. Jacob and J. Monod, *J. Mol. Biol.* **3**, 318 (1961).
15. A. L. Lehninger, D. L. Nelson and M. M. Cox, *Principles of Biochemistry*, 2nd edn. (Worth Publ., New York, 1993).
16. L. Stryer, *Biochemistry*, 4th edn. (Freeman, New York, 1995).
17. B. Lewin, *Genes V* (Oxford University Press, Oxford, 1994).
18. W. S. Klug and M. R. Cummings, *Concepts of Genetics*, 3rd edn. (MacMillan, New York, 1991).
19. S. Ozawa, T. H. Jukes, K. Watanabe and A. Muto, *Microbiol. Rev.* **56**, 229 (1992).
20. F. H. C. Crick, *J. Mol. Biol.* **38**, 367 (1968).
21. M. O. Bertman and J. R. Jungck, *Some Unresolved Mathematical Problems in Genetic Coding*, Not. AMS **25**, A-174 (1978).
22. B. L. van der Waerden, *Die gruppentheoretische Methode in der Quantenmechanik* (Springer, Berlin, 1932); revised English transl.: *Group Theory and Quantum Mechanics* (Springer, Berlin, 1974).

23. H. Weyl, *Gruppentheorie und Quantenmechanik* (Hirzel, Leipzig, 1931); English transl.: *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1950).
24. H. Weyl, *The Classical Groups, their Invariants and Representations* (Princeton University Press, Princeton, 1939).
25. E. P. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren* (Vieweg, Braunschweig, 1931).
26. E. P. Wigner, *Ann. Math.* **40**, 149 (1939).
27. W. Pauli, *Z. Phys.* **36**, 336 (1926).
28. M. Gell-Mann and Y. Ne'eman, *The Eightfold Way* (Benjamin, New York, 1964).
29. A. Arima and F. Iachello, *The Interacting Boson Model* (Cambridge University Press, Cambridge, 1987).
30. F. Iachello and R. D. Levine, *Algebraic Theory of Molecules* (Oxford University Press, Oxford, 1995).
31. F. Iachello, *The Interacting Boson Model*, PhD Thesis (MIT, Cambridge, MA, 1969).
32. H. Feshbach and F. Iachello, *Phys. Lett.* **45B**, 7 (1973); *Ann. Phys. (NY)* **84**, 211 (1974).
33. A. Arima and F. Iachello, *Phys. Rev. Lett.* **35**, 1069 (1975); *Ann. Phys. (NY)* **99**, 253 (1976).
34. A. Arima, T. Ohtsuka, F. Iachello and I. Talmi, *Phys. Lett.* **66B**, 205 (1977).
35. F. Iachello, *Chem. Phys. Lett.* **78**, 581 (1981).
36. F. Iachello and R. D. Levine, *J. Chem. Phys.* **77**, 3046 (1982); *ibid.* **79**, 2515 (1983).
37. E. S. Bernardes, Y. M. M. Hornos and J. E. M. Hornos, *Chem. Phys.* **213**, 17 (1996).
38. J. E. M. Hornos and Y. M. M. Hornos, *Phys. Rev. Lett.* **71**, 4401 (1993).
39. J. E. M. Hornos and Y. M. M. Hornos, *J. Biol. Phys.* **20**, 289 (1994).
40. J. Maddox, *Nature* **367**, 111 (1994).
41. I. Stewart, *New Scientist* **141**, 1915, 16 (1994).
42. A. O. Barut and R. Rączka, *Theory of Group Representations and Applications*, 2nd edn. (World Scientific, Singapore, 1986).
43. T. Bröcker and T. tom Dieck, *Representations of Compact Lie Groups* (Springer, Berlin, 1985).
44. M. Hamermesh, *Group Theory and its Application to Physical Problems* (Addison-Wesley, Reading, 1962).
45. S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces* (Academic Press, New York, 1978).
46. J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory* (Springer, Berlin, 1972).
47. A. L. Onishchik and E. B. Vinberg (eds), *Lie Groups and Lie Algebras III*, Encyclopaedia of Mathematical Sciences, Vol. 41 (Springer, Berlin, 1994).
48. B. G. Wybourne, *Classical Groups for Physicists* (John Wiley, New York, 1974).
49. I. M. Gel'fand and M. L. Zetlin, *Finite-dimensional Representations of the Group of Unimodular Matrices* and *Finite-dimensional Representations of the Group of Orthogonal Matrices*, in: Collected Papers, Vol. 2, ed. S. D. Gindikin, V. M. Guillemin, A. A. Kirillov, B. Kostant and S. Sternberg, pp. 653–656 and 657–661 (Springer, Berlin, 1988).
50. M. Cerkaski, *J. Math. Phys.* **28**, 989 (1987).
51. E. B. Dynkin, *AMS Transl.* **6**, 111 (1957).
52. E. B. Dynkin, *AMS Transl.* **6**, 245 (1957).
53. W. G. McKay and J. Patera, *Tables of Dimensions, Indices and Branching Rules for Representations of Simple Lie Algebras* (Marcel Dekker, New York, 1981).

54. L. Braggion, *Procura por Simetrias de Lie na Evolução do Código Genético (Search for Symmetries in the Evolution of the Genetic Code)*, MSc dissertation, University of São Paulo, São Carlos, 1998.
55. F. Antoneli, *Subalgebras Maximais das Álgebras de Lie Semisimples, Quebra de Simetria e o Código Genético (Maximal Subalgebras of Semisimple Lie Algebras, Symmetry Breaking and the Genetic Code)*, MSc Dissertation, University of São Paulo, São Paulo, 1998.
56. F. Antoneli, L. Braggion, M. Forger and J. E. M. Hornos, *Extending the Search for Symmetries in the Genetic Code*, in preparation.
57. M. Forger, Y. M. M. Hornos and J. E. M. Hornos, *Phys. Rev.* **E56**, 7078 (1997).
58. E. Chacon and M. Moshinsky, *Bases of Irreps of the Chain of Lie Algebras* $sp(6) \supset sp(4) \oplus su(2) \supset su(2) \oplus su(2) \oplus su(2)$, in: Topics in Theoretical Physics, eds. V. C. Aguilera-Navarro, D. Galetti, B. M. Pimentel and L. Tomio, pp. 337–345 (Editora IFT, São Paulo, 1996).
59. M. Barbosa, *Construção da Representação Simplética Irredutível para o Modelo Algébrico de Evolução do Código Genético (Construction of the Irreducible Symplectic Representation for the Algebraic Model of the Evolution of the Genetic Code)*, MSc Dissertation, University of São Paulo, São Carlos, 1997.
60. V. G. Kac, *Adv. Math.* **26**, 8 (1977).
61. V. G. Kac, *Representations of Classical Lie Superalgebras*, in: Proceedings of the VIth International Conference on Differential Geometric Methods in Theoretical Physics, Bonn, Germany 1977, Lecture Notes in Mathematics, Vol. 676 (Springer, Berlin, 1978), pp. 597–626.
62. M. Forger and S. Sachse, *Lie Superalgebras and the Multiplet Structure of the Genetic Code I: Codon Representations*, Preprint RT-MAP 9802, IME-USP, University of São Paulo, São Paulo, October 1998.
63. M. Forger and S. Sachse, *Lie Superalgebras and the Multiplet Structure of the Genetic Code II: Branching Schemes*, Preprint RT-MAP 9901, IME-USP, University of São Paulo, São Paulo, April 1999.
64. J. D. Bashford, *Supersymmetry and the Genetic Code*, PhD Thesis, University of Tasmania, 1995.
65. J. D. Bashford, I. Tsohantjis and P. D. Jarvis, INSERT TITLE! Preprint UTAS-PHYS-96-02; *Phys. Lett.* **A233**, 481 (1997).
66. M. Golubitsky, I. Stewart and D. G. Schaeffer, *Singularities and Groups in Bifurcation Theory*, Vol. II (Springer-Verlag, Berlin, 1988).
67. M. Forger, *J. Math. Phys.* **39**, 1107 (1998).
68. Y. M. M. Hornos and J. E. M. Hornos, *Symmetry in the Evolution of Non-Standard Genetic Codes*, Preprint, University of São Paulo, São Carlos, 1999.