APPLICATIONS OF GROUP THEORY TO QUANTUM PHYSICS

ALGEBRAIC ASPECTS

by

Louis Michel†

TABLE OF CONTENTS

0. INTRODUCTION 38

1. COVARIANCE IN QUANTUM THEORY AND ITS MATHEMATICAL TOOLS 43

1. 1. What Is Quantum Mechanics 43
1. 2. Group Invariance 45
1. 3. G-Vector Spaces 47
1. 4. Unitary Groups U(n) and Permutation Groups S(n) 51
1. 5. More Algebras and More Tensor Operators. Pseudo Roots of SU(n) 54
1. 6. More on SU(2) and its Tensor Operators 57

2. ATOMIC AND MOLECULAR PHYSICS 62

2. 1. Group Theory and Atomic Physics 62
2. 2. The Correspondence Principle 62
2. 3. Particle of Mass m in a Spherically Symmetric Potential 64
2. 4. The Hydrogen Atom 65
2. 5. The Helium Atom 71
2. 6. Pauli Principle. The Electron Spin 73
2. 7. Atomic Shell Structure - Periodic Table 74
2. 8. Atomic States in a Given Shell - Spin Orbit Coupling 78
2. 9. Spin and Euclidean or Galilean Invariance 79
2.10. Molecules 81
2.11. Measurement of Spin and Statistics of Nuclei by the Study of Diatomic-Molecule Spectra 82

3. NUCLEAR PHYSICS: STRONG AND WEAK INTERACTIONS 84

3. 1. The Set of Known Nuclei 84
3. 2. Isospin 86

† Institut des Hautes Etudes Scientifiques, 91 - BURES-SUR-YVETTE - France.

3. 3. U(4) Invariance 89
3. 4. Shell Model 92
3. 5. The Hadrons 95
3. 6. The Other Particles and the Other Interactions 98

4. RELATIVISTIC INVARIANCE. THE DISCRETE SYMMETRIES C, P, T 101

4. 1. The Poincaré Group and its Automorphisms; Zeeman Theory 101
4. 2. Relativistic Invariance and Internal Symmetries 102
4. 3. Irrep of $\overline{P}_0$ 105
4. 4. Particle States and Irreps of $\overline{P}_0$ 108
4. 5. Particle Polarization 110
4. 6. Invariance Under $P \times Z_2(C)$; PCT Theorem 113
4. 7. How to Observe Violation 115
a. Action of P, T, C on Observables 115
b. Parity Violation 116
c. Time Reversal Invariance 118
4. 8. CP Violation 119

5. THE INTERNAL SYMMETRIES OF HADRONS 121

5. 1. SU(3) Symmetry 121
a. The Octets 121
b. The Mass Operator 122
c. The First Baryon Decuplet 124
d. Other SU(3) Multiplets 124
e. Cross-Sections and Decays of Resonances 125
5. 2. Geometry on the SU(3) - Octet 125
5. 3. Electromagnetic and Weak Interactions in SU(3) 128
a. Electromagnetic Interaction 128
b. Weak Interaction 129
5. 4. Critical Orbits of a G-Invariant Function on a Manifold M 131
5. 5. SU(3) × SU(3) Symmetry 134
5. 6. SU(6), Quarks, Current Algebra, Boot-Strap, Etc. 138

6. ACKNOWLEDGMENTS 143

0. INTRODUCTION

Since you mathematicians and we physicists came here to meet together, there is no need to emphasize that we both believe that the progress of physics requires for its theoretical formulation more and more advanced mathematics. I thought fit however to give you the opportunity to read what Dirac wrote on this subject, 38 years ago, as an introduction to the very paper where he predicted the existence of the "antielectron", which we now call positron.

Notwithstanding Dirac's prediction, when positrons were observed one year later by Blackett and Occhialini, and by the Joliot-Curies, they were not immediately recognized. And Anderson who was the first to identify a positron (in cosmic rays) did not know Dirac's paper. This illustrates the communication difficulties which existed and still exist between theoretical and experimental physicists. You should also expect them between mathematicians and physicists. (Not to speak of the difficulties due to my use of English.) You and I are here determined to overcome them, but I beg your patience in advance.

Extract from *Proc. Roy. Soc.*, Ser. A, 130, 60 (1930):

Quantised Singularities in the Electromagnetic Field

By P. A. M. DIRAC, F. R. S., St. John's College, Cambridge

§ 1. Introduction

The steady progress of physics requires for its theoretical formulation a mathematics that gets continually more advanced. This is only natural and to be expected. What, however, was not expected by the scientific workers of the last century was the particular form that the line of advancement of the mathematics would take, namely, it was expected that the mathematics would get more and more complicated, but would rest on a permanent basis of axioms and definitions, while actually the modern physical developments have required a mathematics that continually shifts its foundations and gets more abstract. Non-euclidean geometry and non-commutative algebra, which were at one time considered to be purely fictions of the mind and pastimes for logical thinkers, have now been found to be very necessary for the description of general facts of the physical world. It seems likely that this process of increasing abstraction will continue in the future and that advance in physics is to be associated with a continual modification and generalization of the axioms at the base of the mathematics rather than with a logical development of any one mathematical scheme on a fixed foundation.

There are at present fundamental problems in theoretical physics awaiting solution, e.g., the relativistic formulation of quantum mechanics and the nature of atomic nuclei (to be followed by more difficult ones such as the problem of life), the solution of which problems will presumably require a more drastic revision of our fundamental concepts than any that have gone before. Quite likely these changes will be so great that it will be beyond the power of human intelligence to get the necessary new ideas by direct attempts to formulate the experimental data in mathematical terms. The theoretical worker in the future will therefore have to proceed in a more indirect way. The most powerful method of advance that can be suggested at present is to employ all the resources of pure mathematics in attempts to perfect and generalize the mathematical formalism that forms the existing basis of theoretical physics and after each success in this direction, to try to interpret the new mathematical

features in terms of physical entities (by a process like Eddington's Principle of Identification).

A recent paper by the author† may possibly be regarded as a small step according to this general scheme of advance. The mathematical formalism at that time involved a serious difficulty through its prediction of negative kinetic energy values for an electron. It was proposed to get over this difficulty, making use of Pauli's Exclusion Principle which does not allow more than one electron in any state, by saying that in the physical world almost all the negative-energy states are already occupied, so that our ordinary electrons of positive energy cannot fall into them. The question then arises as to the physical interpretation of the negative-energy states, which on this view really exist. We should expect the uniformly filled distribution of negative-energy states to be completely unobservable to us, but an unoccupied one of these states, being something exceptional, would make its presence felt as a kind of hole. It was shown that one of these holes would appear to us as a particle with a positive energy and a positive charge and it was suggested that this particle should be identified with a proton. Subsequent investigations, however, have shown that this particle necessarily has the same mass as an electron†† and also that, if it collides with an electron, the two will have a chance of annihilating one another much too great to be consistent with the known stability of matter.†††

It thus appears that we must abandon the identification of the holes with protons and must find some other interpretation for them. Following Oppenheimer,†††† we can assume that in the world as we know it, all, and not merely nearly all, of the negative-energy states for electrons are occupied. A hole, if there were one, would be a new kind of particle, unknown to experimental physics, having the same mass and opposite charge to an electron. We may call such a particle an antielectron. We should not expect to find any of them in nature, on account of their rapid rate of recombination with electrons, but if they could be produced experimentally in high vacuum they would be quite stable and amenable to observation. An encounter between two hard γ-rays (of energy at least half a million volts) could lead to the creation simultaneously of an electron and antielectron, the probability of occurence of this process being of the same order of magnitude as that of the collision of the two γ-rays on the assumption that they are spheres of the same size as classical electrons. This probability is negligible, however, with the intensities of γ-rays at present available.

The protons on the above view are quite unconnected with electrons. Presumably the protons will have their own negative-energy states, all of which normally are occupied, an unoccupied one appearing as an antiproton.

Let me just remind you that antiprotons were first observed twenty-four years later (1955).

There will be many advanced seminars on the applications of group theory to quantum physics. So I believe that these lectures must be introductory, and that I have to present concepts that will be used by all physicists here. That will be Part 1.

---

† *Proc. Roy. Soc., Ser. A*, 126, 360 (1930).

†† H. Weyl, *Gruppentheorie und Quantenmechanik*, 2nd ed., p. 234 (1931).

††† I. Tamm, *Z Physik*, 62, 545 (1930); J. R. Oppenheimer, *Phys. Rev.*, 35, 939 (1930); P. Dirac, *Proc. Camb. Philos. Soc.*, 26, 361 (1930).

†††† J. R. Oppenheimer, *Phys. Rev.*, 35, 562 (1930).

The ultimate goal of these lectures will be to bring you to the present problems on the subject, mainly in the field of the fundamental particle physics. Then there might be some overlap with Professor O'Raifeartaigh's lectures, but there should be no inconvenience to see some aspects of physics from probably two different points of view. We have to face the fact that fruitful discussions of frontier problems of physics between mathematicians and physicists are difficult, because these problems often cannot be presented in a formalized language, but only through some physical analogy. So, obviously, to understand what is the problem, one must know some physics!

In these lectures I will therefore present a quick survey of applications of group theory to atomic, molecular and nuclear physics. Often, I will even follow an historical approach. Indeed, physicist minds are partly conditioned by the recent history of physics. But I also hope to use the power of your language, mathematics, to convey to you a maximum of physics in a minimum of time. Of course I shall have succeeded only if I have also been able to convince you that physics is fascinating!

It is fit to end this introduction by the history of the birth of our subject. Less than three years after the first paper on quantum mechanics (W. Heisenberg, *Z. Phys.*, 33, 879 (1925), there appeared the first two papers devoted to the application of group theory to quantum mechanics:

- E. P. Wigner, "Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen", *Z. Phys.*, 43, 624 (1927).
- F. Hund, "Symmetriecharaktere von Termen bei Systemen mit gleichen Partikeln in der Quantenmechanik", *Z. Phys.*, 43, 788 (1927).

Wigner will surely be the most quoted author on our subject. Let us just say that, with J. von Neumann, he applied group theory to atomic spectra ("Zur Erklärung einiger Eigenshaften der Spektren aus der Quantenmechanik des Drehelektrons I., II., III., *Z. Phys.*, 47, 203; 49, 73; 51, 844 (1928)), and published a self-contained book on this question: E. P. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*, Vieweg, Braunschweig (1931).

It is remarkable that two famous mathematicians, Hermann Weyl and Van der Waerden, also published very early books on our subject:

- H. Weyl, *Gruppentheorie und Quantenmechanik*, Hirzel, Leipzig (1928).
- Van der Waerden, *Die Gruppentheoretische Methode in der Quantenmechanik*, Springer, Berlin (1932).

Then the excellent, but more elementary book, by E. Bauer, *Introduction à la Théorie des Groupes et ses Application à la Physique Quantique*, Presses Universitaires de France, Paris (1933), continued a list of books which, today, may have reached several dozen.

For the interested mathematicians I would still recommend the two very first books, but in their second, revised and enlarged edition: H. Weyl, *The Theory of Groups and Quantum Mechanics*, Methuen, London (1931); Paper Back reprint, Dover, New York (1949); the translation by J. J. Griffin of Wigner's book, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra*, Academic Press, New York (1959).

For The Mathematician Readers

Physics will be injected in these notes as needed. However, it seems convenient to gather here some information on physical constants which might be useful at any time.

We will study quantum phenomena. In atomic, nuclear, fundamental particle physics, the key number to pass from macroscopic scale is the Avogadro number:

$$a = 6.0228 \times 10^{23} \sim 6.03 \times 10^{23} \quad . \tag{0.1}$$

It is the number of atoms in a mass of one gram of hydrogen.

A hydrogen atom is made of one proton (mass m_p) and one electron (mass m_ε).

$$\frac{m_p}{m_\varepsilon} = 1836.5 \quad .$$

These two particles are electrically charged, p^+, ε^-, the absolute value of this charge is

$$e = \frac{1 \text{ Faraday}}{a} = \frac{96{,}600}{6.03 \times 10^{23}} \text{ Coulombs} \quad .$$

The most convenient unit systems, for us, will use

$$\hbar = (\text{Planck constant}) \times (2\pi)^{-1} = 1$$
$$c = (\text{velocity of light}) \qquad\quad = 1 \quad .$$

In this system $e = (137.04)^{-1/2}$, indeed

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137.04}$$

is a dimensionless number. Atoms of the other elements are made of Z electrons and a nucleus which contains Z protons and N neutrons; Z is the atomic number, $A = Z + N$ the atomic mass number, e.g., for hydrogen $Z = 1$; hydrogen has 3 isotopes $N = 0$, $A = 1$; $A = 2$, deuterium; $A = 3$, tritium (unstable, lifetime 12 years). For uranium $Z = 92$; the most abundant isotope is $A = 238$. The neutron has no electric charge and its mass is nearly equal to that of the proton,

$$\frac{m_n}{m_\varepsilon} = 1839.0 \quad .$$

So the atom mass is practically concentrated in its nucleus.

Many more particles will be introduced, e.g., the photon γ, and the four different neutrinos all with zero rest-mass and zero charge, etc... see 3.5 and 3.6.

We remind the reader that in relativistic physics mass is not conserved; mass is a form of energy. The energy of a particle of mass m, velocity v is $E = mc^2(1-(v/c)^2)^{-1/2}$. The rest energy $(v = 0)$ is mc^2.

We need to choose another unit to complete our unit systems. The best choice for atomic physics is the electron mass m_ε. Then the other units are:

momentum	$m_\varepsilon c$
energy	$m_\varepsilon c^2$
length	$\hbar/m_\varepsilon c = 3.86 \times 10^{-11}$ cm
time	$\hbar/m_\varepsilon c^2 = 1.28 \times 10^{-21}$ sec.

However, due to the nature of its measurement, the most common energy unit used for particles is the electron volt (eV). It is the energy that a particle with the universal electric charge e gains by traversing an electric field of potential difference one volt.

The conversion with the preceeding unit system is

$$m_\varepsilon c^2 = 0.511 \times 10^6 \text{ eV} \sim 1/2 \text{ MeV}$$

$$m_p c^2 = 938.256 \text{ MeV} \sim 1 \text{ GeV} = 10^9 \text{ eV} \quad .$$

Note that $1 = \hbar \sim (10^{-13} \text{ cm}) \times (200 \text{ MeV})$

Before 1932, only two kinds of interactions were known, gravitation and electromagnetism. In the static approximation the two interactions can be described by proportional potentials K/r where r is the distance. So the absolute ratio of the (attractive) gravitational energy to the (repulsive) electrostatic energy between two protons is independent of their distance.

It is

$$\frac{G' m_p^2}{e^2} = \frac{\alpha'}{\alpha} = \frac{137}{175} \times 10^{-36} \quad .$$

Thus, gravitation will be completely neglected in these lectures.†

† The gravitational energy of a system increases roughly as the square of the number N of nucleons while, in neutral matter, the electrostatic energy is roughly proportional to N. So gravitation becomes important only for masses as large as that of asteroids, planets (we know it on the earth!) or stars. It is not a coincidence that most stars have a number of nucleons $\sim \alpha'^{-3/2} = 10^{57}$ (see for instance E. E. Salpeter, "Dimensionless Ratio and Stellar Structure", in *Frontier in Physics, Bethe Festschrift*, p. 463, R. Marshak Editor).

The binding energy of atoms, molecules, solids, etc., is of electromagnetic origin. This energy can be released in chemical form, with an order of magnitude:

$$a \times 1 \text{ eV} = 23 \text{ cal/mole}$$

which is $\sim 10^{-9}$ to 10^{-11} the rest-mass energy. If the energy we receive from the sun came from chemical reactions, the sun would produce it for less than 10^5 years!

There are two other known kinds of interactions: the nuclear interaction (see part 3 and 5), stronger than the electromagnetic interaction at distance smaller than 10^{-13} cm; the Fermi or "weak interaction" (see 3.6 and 5), which is very short range. Both interactions are important in stars and nuclear reactions and can yield an energy up to 10^{-3} the rest mass energy.

For the Physicist Reader

All mathematical terms used here are not defined. Of course many of them are known to physicists (e.g., for the notion of root vectors of Lie algebra, see Salam's lectures in *High Energy Physics and Elementary Particles, Trieste Seminar 1962* (International Atomic Energy Agency, Vienna (1963)). Some terms (used mainly in I) come from a modern mathematical terminology. They were not absolutely necessary and they are used explicitly as synonyms of other terms generally used by physicists. Physicists should know the proper mathematical terms of the mathematical concepts they need: indeed their students, and even their young children know them and physicists want to communicate with their students and their children!

An excellent and elementary exposition of this modern mathematical language is given in the text *Algebra* by S. Mac Lane and G. Birkhoff, Macmillan, New York, (1967), particularly Chapter I; note also the list of symbols, p. XVII to XIX.

1. COVARIANCE IN QUANTUM THEORY AND ITS MATHEMATICAL TOOLS

1.1. What Is Quantum Mechanics

Less than two years after the first paper (quoted above) of Heisenberg on quantum mechanics, J. von Neumann answered this question in three successive papers in *Göttingen Nachrichten*, (1927), (pp. 1, 245, 273) expanded in a book: *Mathematische Grundlagen der Quanten Mechanik*, (1930) (English translation, Princeton University Press (1955)).†

† He later published with G. Birkhoff, "The Logic of Quantum Mechanics", *Ann. of Math.*, 37, 935 (1936). This subject is still controversial and lively.

Two early books on quantum mechanics by physicists are reedited and still very advisable reading: P. A. M. Dirac, *The Principles of Quantum Mechanics*, Clarendon Press, Oxford 1st ed. (1930) 4th ed. (1958). W. Pauli, "Prinzipien der Quanten Theorie", *Handbuch der Physik*, 1, Springer (1958) 1st ed. (1933).†

If you have not read these books it is not too late to do it, but today let us just give a mini-description of quantum mechanics.

a) To each physical system corresponds a separable complex Hilbert space $\mathcal{H}$. A physical state is represented by a vector $x\rangle \in \mathcal{H}$. (Normed to 1 for convenience: $\langle x,x\rangle = 1$.)

b) Each physical observable $\underline{a}$ (e.g., energy, electric charge, etc.) is represented by a self adjoint operator A on $\mathcal{H}$. The spectrum of A is the set of possible values of $\underline{a}$.

c) Quantum mechanics does not predict, in general, the value of $\underline{a}$ for the state $x\rangle$, it gives only its expectation value:

$$\langle x,Ax\rangle = \mathrm{Tr}\, A\, P_x \tag{1.1}$$

where P_x is the Hermitian projector ($P_x = P_x^*$) onto the one dimensional space spanned by $x\rangle$. Note that unit eigen vectors of P_x (with eigen-value 1) differ only by a scalar phase factor and describe the same state since they yield the same physical predictions. The projectors P_x are themselves observables. Indeed

$$\mathrm{Tr} P_x P_y = |\langle x,y\rangle|^2 \tag{1.2}$$

is the probability to observe in the state $x\rangle$ (respectively $y\rangle$) the physical system which is known to be in the state $y\rangle$ (respectively $x\rangle$). Part of the art of the quantum physicist is to code what he sees in nature into vectors of Hilbert space! This always requires "physical approximations".

When we can describe a state by a rank one projector (or a vector up to a phase) we say that we have a pure state and that we have a complete information on it.

More often our information on the state is only partial. In the simplest case we know only a set of probabilities c_i (with $\Sigma c_i = 1$) for the system to be in the set of orthogonal pure states P_i (i.e., $P_i P_j = \delta_{ij} P_j$) so the expectation value is

$$\Sigma_i c_i\, \mathrm{Tr}\, A\, P_i = \mathrm{Tr}\, A\, R \tag{1.3}$$

with

$$R = \Sigma_i c_i P_i \qquad \mathrm{Tr}\, R = \Sigma_i c_i = 1 \ . \tag{1.3'}$$

Since $0 \leq c_i$, the self adjoint operator R is positive and it is called the

---

† There are also books on the mathematical foundation of quantum mechanics by mathematicians: G. Mackey, L. Schwartz.

density matrix† of the mixture (= not pure) state of the system. Pure states are extremal points of the convex domain of states.

This leads us to a natural generalization. One defines a Banach*-algebra $\mathcal{B}$ with unit 1, generated by the observables.†† (More specifically it is usually a C*-algebra). Then a state is a linear functional ϕ on $\mathcal{B}$ which is positive, that is $\forall A \in \mathcal{B}$, $\phi(A^*A) \geq 0$. For systems with a finite number of degrees of freedom this is not an essential generalization. It becomes so for infinite degrees of freedom as in quantum field theory and statistical mechanics. Classical statistical mechanics can also be put in the same mathematical mould with an abelian algebra.†††

1.2 Group Invariance

We assume that there is a relativity group G for every physical theory considered here. That is G acts on a physical system S, and there is an isomorphism between the physics of S (its Hilbert space of states $\mathcal{H}$, its algebra of observables $\mathcal{B} \subset L(\mathcal{H})$††††, etc...) and the physics of $\underline{g}(S)$, the transform of S, by $\underline{g} \in G$ (e.g., $\underline{g}$ can be a rotation). This will be called the "active" point of view of G-invariance. The "passive" point of view for a transformation group is simply the isomorphism between the physical description of the same system S by two observers choosing different coordinate frames, G-transforms of each other.

For any $g \in G$, we denote by P_{gx_i} the transformed of the state P_{x_i}. To say that G is an invariance group is equivalent to saying that all probabilities of Equation (1.2) are invariant

$$\forall x\rangle \in \mathcal{H},\quad g \in G,\quad \mathrm{Tr} P_{gx_1} P_{gx_2} = \mathrm{Tr} P_{x_1} P_{x_2} \tag{1.4}$$

---

† J. von Neumann introduced the density-matrix in 1927 in the papers quoted above.

†† Quite early physicists also considered non-associative algebras formed by the observables and introduced Jordan algebras. The first fundamental paper on those algebras is by P. Jordan, J. von Neumann and E. Wigner; "On an Algebraic Generalization of the Quantum Mechanical Formalism", *Ann. of Math.*, 35, 29 (1934).

††† I. E. Segal advocated twenty years ago the use of C*-algebra for quantum physics. The fundamental paper showing the benefits from this choice (physical approximation and Feld's ε-equivalence; introduction of super-selection rules) is that of R. Haag and D. Kastler, "An Algebraic Approach to Quantum Field Theory", *J. Math. Phys., Supplement* 848 (1964). Most of the C*-algebra physics is published in the journal: *Communications in Mathematical Physics*, and is written in a rigorous mathematical style. For statistical mechanics, see D. Ruelle, *Statistical Mechanics*, Benjamin, New York (1969). Soon there will appear in the collection of C. N. R. S. Colloquia (France) "Rigorous Results on Interacting Systems with Infinite Degrees of Freedom".

†††† We denote by $L(\mathcal{H})$ the space of linear operators on $\mathcal{H}$.

or

$$|\langle gx_1, gx_2\rangle|^2 = |\langle x_1, x_2\rangle|^2 \tag{1.4'}$$

This means that G acts on $\mathcal{H}$ by isometries.

Wigner proved in his book (Appendix to Chapter 20)† that $x\rangle \rightsquigarrow gx\rangle$ is either a unitary operator $U(g)$ or an antiunitary operator $V(g)$ on $\mathcal{H}$. We recall that an antiunitary operator V has the characteristic properties

$$x\rangle,\ y\rangle \in \mathcal{H},\ V(\alpha x\rangle + \beta y\rangle) = \bar{\alpha} Vx\rangle + \bar{\beta} Vy\rangle \tag{1.5}$$

$$\langle Vx, Vy\rangle = \overline{\langle x, y\rangle} = \langle y, x\rangle \tag{1.5'}$$

$$\text{it has an inverse} \tag{1.5''}$$

Given an isometry on $\mathcal{H}$, there is a simple criterion†† for deciding whether it is realized by a unitary operator U or an antiunitary operator V. In either case U or V is defined up to a scalar phase factor. The product of two antiunitary operators is a unitary operator.

Let $\mathcal{V}(\mathcal{H})$ be the group of unitary and antiunitary operators on $\mathcal{H}$ and $\mathcal{U}(\mathcal{H})$ the subgroup of unitary operators. $\mathcal{U}(\mathcal{H})$ is an invariant subgroup of $\mathcal{V}(\mathcal{H})$ since it is a subgroup of index two. We assume that G acts effectively on $\mathcal{H}$, i.e., no other element than $1 \in G$ acts trivially on $\mathcal{H}$. The $U(g)$'s or $V(g)$'s for $g \in G$ generate a subgroup $E(G)$ of $\mathcal{V}(\mathcal{H})$ which is an extension of G by the group U_1 (phase multiplication of the vectors of $\mathcal{H}$, leaving invariant the states) with the action

$$G \xrightarrow{f} \text{Aut } U_1$$

where $\text{Ker } f$ is the invariant subgroup of index two $G_+ \subset G$ which acts by unitary transformations and the non-trivial element of $\text{Im } f$ is the complex conjugation $\alpha \rightsquigarrow \bar{\alpha} = \alpha^{-1} \in U(1)$.

We can also say that G_+ acts by a linear unitary projective representation and Wigner has coined the word projective "corepresentation" for the action of G (when G is strictly larger than G_+).

Wigner also showed from physical arguments that antiunitary operators are to be used with transformations which reverse the direction of time, this in order that energy be positive: indeed, the time translation t is represented by the operator e^{iHt}; if $t \to -t$, i has to go to $-i$ in order that both H and e^{iHt} be invariant.

---

† A more explicit proof of Wigner's theorem has been given by V. Bargmann, *J. Math. Phys.*, 5, 862 (1964). See also proofs of slight generalizations by U. Uhlhorn, *Arkiv for Fysik*, 23, 307 (1963). In the framework of Birkhoff and von Neumann axiomatics, the equivalent theorem has been proven by G. Emch and C. Piron, *J. Math. Phys.*, 4, 469 (1963).

†† See Bargmann:† for any triplet of vectors $x\rangle$, $y\rangle$, $z\rangle$, $\langle x,y\rangle\langle y,z\rangle\langle z,x\rangle$ is invariant under a unitary transformation U and is transformed into its complex conjugate under the antiunitary transformation V.

Continuous projective linear unitary representation of finite groups or Lie groups are well known. For instance, for the three-dimensional rotation group $SO(3,R)$ these projective representations are in a one to one correspondence with the "linear irreducible unitary representations" (= irrep through all these lectures) of $SU(2)$ the universal covering of $SO(3,R)$. This justifies the introduction of spinors in quantum physics.

In Part 4 we will study invariance under the relativity groups of non-relativistic (= Newtonian) mechanics and of special relativity theory. But there are other invariance groups in physics. For instance the permutation group $S(n)$ acting on n identical particles (as the electrons of an atom). In nuclear physics and fundamental particle physics we shall meet many "approximate invariances". The corresponding invariance group is most often a $U(n)$ or $SU(n)$ group (group of unitary $n \times n$ matrices, with determinant 1 for $SU(n)$) with $n = 1, 2, 3, 4, 6$. We shall have more to say for the word "approximate" symmetry.

We will also have to study invariance under a group G when G is a symmetry group for a physical system, e.g., the symmetry group (one of the crystallographic group) of a crystal. This example raises a fascinating question about group invariance in physics. Surely the interaction between atoms are translation invariant (and may be invariant under a larger transformation group). How is it possible that atoms aggregate to form a crystal whose lattice is invariant only under a subgroup of a translation group? When such a phenomenon occurs, i.e., when a stable state has a lesser symmetry than that of the physical laws we will say that we have a broken symmetry.†

We will continue this Part 1 by introducing some mathematical tools that we will use quite frequently.

1.3 G-Vector Spaces

Let G be a given group. If you like you can say that we consider a category whose objects are vector spaces E (over a given field K) with a linear action of G on E (i.e., $G \ni x \rightsquigarrow g(x) \in L(E)$, where $L(E)$ is the algebra of endomorphisms of E, with $xy \rightsquigarrow g(x)g(y) = g(xy)$.

The morphisms of the category are the vector space homomorphisms $E \xrightarrow{f} E'$ compatible with the group action, i.e., they are the commutative diagrams for every $x \in G$, of vector space homomorphisms. We will call these morphisms G-homomorphisms of G-vector spaces.

---

† This short section on group invariance is too sketchy. Much more should be said of the symmetry of physical laws (e.g., E. P. Wigner, "Symmetry and Conversation Laws", *Proc. Nat. Acad. Sci., U.S.A.*, 51, 956 (1964)) without which symmetries of states, that we have considered, would not last. Of course much more will be said in these lectures.

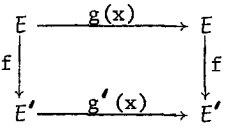


Diagram 1.

Of course we could have also said that we study bimodules (G- and K-modules) or even more simply that we are interested in the linear representations of G; and G-homomorphisms are also called "intertwining" operators. Note that the G-morphisms from E_1 to E_2 form a vector space that we denote $\mathrm{Hom}(E_1, E_2)^G$. Indeed, it is the subspace of the invariant vectors of $\mathrm{Hom}(E_1, E_2)$; they are the intertwining operators for the two representations of G on E_1 and E_2.

Given G-vector spaces, E_1, E_2 ... all vector spaces which can be formed functorially from them are also G-vector spaces, e.g., $E_1 \otimes E_2$, $\mathrm{Hom}(E_1, E_2)$, $L(E) = \mathrm{Hom}(E, E)$, the vector space of the tensor algebra on E: $T(E) = \oplus_{n=0}^{\infty} E^{(n)}$ when $E^{(n)} = E \otimes E \otimes \ldots \otimes E$, n factors, (and $E^{(0)} = K$), etc.

Given a physical system, let $\mathcal{H}$ be the Hilbert space of its state vectors. Assume that $\mathcal{H}$ is a G-vector space. So is $L(\mathcal{H})$. We are then led to the study of the objects (of the category of G-vector spaces) "above" $L(\mathcal{H})$. They are called in the physical literature "Tensor operators on $\mathcal{H}$". (A notable exception is the book devoted to this subject, *Irreducible Tensorial Sets*, by V. Fano and G. Racah, Academic Press, New York (1959)). *By definition*, for physicists, an "E_i-*tensor operator*" *is a G-morphism (or intertwining operator) from* E_i *to* $L(\mathcal{H})$. If the representation of G on E_i is irreducible, then the corresponding G-morphism is called in physics an "irreducible tensor operator". If G acts trivially on E_i then we have "scalar tensor operator". (Just try to remember that tensor operators on $\mathcal{H}$ are not operators on $\mathcal{H}$!)†

It is time to specify the field K. Generally, of course, it is the field of complex numbers since $\mathcal{H}$ is a complex Hilbert space. However, reality is also essential in physics. So often E is a real vector space and the "E-tensor operator" is a G-homomorphism T of real vector spaces from E to the real vector space of self-adjoint operators on $\mathcal{H}$. Of course it is always possible later to enlarge the field from $\mathbb{R}$ to $\mathbb{C}$.

When G is a Lie group we consider, of course, only continuous differentiable representations so a G-vector space is also a $\underline{g}$-module for the Lie algebra $\underline{g}$ of G. We denote G the vector space of $\underline{g}$. Among the G-tensor operators on $L(\mathcal{H})$ there is a particular one F' which is also the Lie algebra representation of $\underline{g}$ on $\mathcal{H}$. When the representation of G on $\mathcal{H}$ is unitary, then $F = iF'$ has self-adjoint operators for images which satisfy

† "Scalar" is often used by physicists for "invariant"!

$$[F(a),F(b)] = (F(a)F(b) - F(b)F(a)) = iF(a \wedge b) \quad . \tag{1.6}$$

When G is respectively the group of rotations, space translations, time-translations, etc., F corresponds respectively to the observables; angular-momentum, momentum, energy, ... In the technical sense of 1.1, what we called observables there, are the elements of the image of F, i.e., for instance, the component of the angular momentum or of the momentum, in a given direction. But I hope it is by now clear that G-morphisms on $L(\mathcal{H})$ are what correspond to the physical concepts with a tensorial character with respect to a group G (other examples: velocity, magnetic moment, electric quadrupole moment, energy-momentum tensor, tensor of inertia, etc.).

Let R and U (unitary) be the representations of G respectively on E and on $\mathcal{H}$. By definition of the E-tensor operator T

$$\forall x \in E, \forall g \in G, \quad U(g)T(x)U^{-1}(g) = T(R(g)x) \quad . \tag{1.7}$$

If D and $F' = iF$ are the corresponding representations of the Lie algebra

$$D(a) = \frac{d}{d\alpha} R(e^{\alpha a})\Big|_{\alpha=0} ; \; iF = \frac{d}{d\alpha} U(e^{\alpha a})\Big|_{\alpha=0} \tag{1.8}$$

then an equivalent definition of the E-tensor operator T is

$$\forall x \in E, \forall a \in \underline{g}, \quad [F(a),T(x)] = iT(D(a)x) \quad . \tag{1.9}$$

In a nutshell, I would say that much of the application of group theory to quantum physics consist in the study of the "tensor-operators" on the G-vector (Hilbert) space $\mathcal{H}$ of a physical system. They form a ring† (and an algebra). Let T_1 and T_2 be respectively E_1 and E_2-tensor operators on $\mathcal{H}$, then

$$E_1 \oplus E_2 \ni x \oplus y \rightsquigarrow T_1(x) + T_2(y) \quad \text{defines a } E_1 \oplus E_2\text{-tensor operator}$$

$$E_1 \otimes E_2 \ni x \otimes y \rightsquigarrow T_1(x)T_2(y) \quad \text{defines a } E_1 \otimes E_2\text{-tensor operator}$$

that we denote respectively $T_1 \oplus T_2$ and $T_1 \otimes T_2$. The latter is generally reducible and can be decomposed into a direct sum of irreducible "tensor operators".

I believe that many problems arise which have not been systematically studied by physicists although they work very much with this ring (for fixed G, $\mathcal{H}$ and action of G on $\mathcal{H}$).

For instance, if G is simple, and T is a G-tensor operator and $\forall x, y \in G$, $[T(x),T(y)] = 0$. I believe this implies $\dim \mathrm{Hom}(G,\mathcal{H})^G$ is infinite.††

Of course the subalgebra generated by an element is well known; given an E-tensor operator T there is a functorial G-morphism $\hat{T}$ from the tensor algebra

---

† For infinite dimensional $\mathcal{H}$, the operators $T(x)$ are not bounded so their product is not always well defined. I will forget here this difficulty which has to be faced in quantum mechanics and is considered in O'Raifeartaigh's lectures.

†† C. Moore proved it during the Rencontres.

$T(E)$ on E to $L(\mathcal{H})$, which is moreover an algebra homomorphism. If i is the canonical injection of E into $T(E)$(Im $i \simeq E^{(1)}$), then the Diagram 2 is commutative.

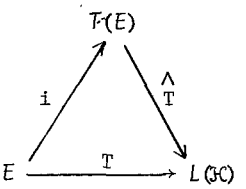


Diagram 2.

In the particular case where T is the representation F (up to a factor i, (see Equation (1.6)) of $\underline{g}$ on $\mathcal{H}$, then it appears also in the representation of $U(G)$, the universal enveloping algebra of $\underline{g}$

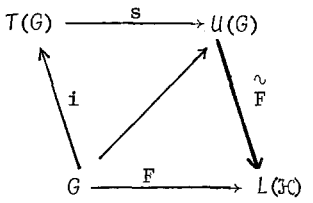


Diagram 3.

$\hat{F}$ (in diagram 2) $= \tilde{F} \cdot s$.

A remarkable "scalar tensor operator" is the Casimir† operator.

Let G be a semi-simple Lie group. Let $a \rightsquigarrow D(a)$ the adjoint representation of the Lie algebra $\underline{g}$ on its vector space G

$$D(a)b = a \wedge b, \quad [D(a),D(b)] = D(a \wedge b) \quad . \tag{1.10}$$

The symmetrical bilinear Cartan-Killing form

$$\beta(a,b) = \mathrm{Tr}D(a)D(b) \tag{1.11}$$

is non-degenerate. Therefore, it defines a G-isomorphism i' between G and its dual G'. This also defines the isomorphism $i' \otimes I$, (I = identity)

$$G \otimes G \xleftarrow{i' \otimes I} G' \otimes G \xleftarrow{j} \mathrm{Hom}(G,G) \quad .$$

The well known canonical homomorphism j is also a G-homomorphism. The identity operator 1 on G is an invariant G-vector $\in \mathrm{Hom}\,(G,G)^G$. So

$$c = (i' \otimes I) \cdot j(1) \tag{1.12}$$

is an invariant vector of $E \otimes E \subset T(E)$ with a fixed normalisation and $\hat{F}(c)$ is the Casimir operator on $\mathcal{H}$.

† Casimir is a physicist with a high position in Philips.

It occurs that neither physicists, nor some mathematicians (cf., Bourbaki, *Groupes et Algèbres de Lie*, Chapter I §3 No. 7) use this canonical normalization for c. In the physics literature nowadays, the images by $\tilde{F}$ of a set of algebraically independant elements of the center of $U(\underline{g})$ are called "the Casimir operators".

In order to induce physicists to use the more canonical point of view exposed here, let us end this section by a very simple theorem proven elaborately in particular cases in the physics literature.

Theorem

If G has no non-trivial one-dimensional representation, and if T is for G a non-invariant irreducible E-tensor operator on $\mathcal{H}$, a finite dimensional space, then $\forall a \in E$, tr T(a) = 0. Indeed, the field ($\mathbb{C}$ for instance) is a trivial one-dimensional G-vector space, and "trace" $\in \mathrm{Hom}(L(\mathcal{H}),\mathbb{C})^G$ since $T \in \mathrm{Hom}(E,L(\mathcal{H}))^G$, then "trace T" = "trace" $\cdot$ $T \in \mathrm{Hom}(E,\mathbb{C})^G = 0$ by our hypothesis.

1.4. Unitary Groups U(n) and Permutation Groups S(n)†

We have to survey briefly some results on irreps of U(n) and S(n) that we shall use very much in these lectures. The irreps of S(n) can be labeled by integer partitions of n

$$[\lambda_1^{\alpha_1} \ldots \lambda_i^{\alpha_i} \ldots \lambda_k^{\alpha_k}] \quad \text{with} \quad \lambda_1 > \lambda_2 \ldots \lambda_k > 0$$

and

$$\Sigma_{i=1}^{k} \alpha_i \lambda_i = n \quad .$$

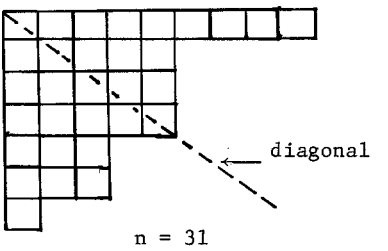
There is a more picturesque notation of $[\lambda_1^{\alpha_1} \ldots \lambda_k^{\alpha_k}]$ which is an ideogram made with n small squares, α_1 lines of λ_1 squares, α_2 lines of λ_2 squares, etc. and called a Young diagram.

Example of $[\lambda_i^{\alpha_i}]$

$\lambda_1 = 9$, $\alpha_1 = 1$, $\lambda_2 = 5$, $\alpha_2 = 3$

$\lambda_3 = 3$, $\alpha_3 = 2$, $\lambda_4 = 1$, $\alpha_4 = 1$

$n = 9 + (3 \times 5) + (2 \times 3) + 1 = 31$



The Young diagrams contain a qualitative information, the more horizontal (vertical) the diagram the more symmetrical (antisymmetrical) are the vectors of the representations.

---

† For an exposition of the representations of U(n) and S(n) see Weyl's book, Chapter V. A survey for the needs of physicists has been made by C. Itzykson and M. Nauenberg, *Rev. Mod. Phys.*, 38, 95 (1966).

There are only two one-dimensional irreps of $S(n)$

$[n] =$ completely symmetrical
n squares

$[1^n] =$  n squares completely antisymmetrical.

The representations of $S(n)$ are self contragredient. To each irrep of $S(n)$ we can associate a complementary representation

$$[\lambda_i^{\alpha_i}]^c = [\hat{\lambda}_i^{\hat{\alpha}_i}] \quad \text{with} \quad \hat{\lambda}_i = \Sigma_{j=1}^{k+1-i} \alpha_j, \ \hat{\alpha}_i = \lambda_{k-i+1} - \lambda_{k-i+2} \ . \tag{1.13}$$

Its Young diagram is simply obtained by a symmetry through the diagonal.

We recall that the tensor product of two irrep contains $[n]$ (resp., $[1^n]$) only if the two irreps are equivalent (resp. complementary) and then it contains $[n]$ (resp., $[1^n]$) once only.

We will also use a shorter symbol $[\ ]_\lambda$ for a linear unitary representation of $S(n)$.

We call factorial a group representation which is direct sum of equivalent irreps.

Let $\mathcal{H}^{(1)}$ be a Hilbert space and

$$\mathcal{H}^{(n)} = \overset{n}{\otimes} \mathcal{H}^{(1)} = \mathcal{H}^{(1)} \otimes \ldots \otimes \mathcal{H}^{(1)} \otimes \mathcal{H}^{(1)} \quad (n \text{ factors}). \tag{1.14}$$

By permutation of the factors, $S(n)$ acts linearly on $\mathcal{H}^{(n)}$ through a representation that we denote $[\ ]_{\mathcal{H}^{(n)}}$ and which can be decomposed canonically into factorial representation. We denote $\mathcal{H}^{(n)}_{[\ ]_\lambda}$ the subspace of $\mathcal{H}^{(n)}$ on which acts the factorial representation $\oplus\ [\ ]_\lambda$.

For instance $\mathcal{H}^{(n)}_{[1^n]}$, also denoted $\overset{n}{\wedge} \mathcal{H}^{(1)}$, and $\mathcal{H}^{(n)}_{[n]}$, also denoted $\overset{n}{\vee} \mathcal{H}^{(1)}$, are the spaces of completely antisymmetrical and symmetrical rank-n tensors on $\mathcal{H}^{(1)}$.

Let us assume that $\dim \mathcal{H}^{(1)} = k$ is finite. Then $U(k)$ acts on $\mathcal{H}^{(1)}$ and on $\mathcal{H}^{(n)}$ by $\overset{n}{\otimes} U(k)$. The decomposition of this linear representation of $U(k)$ on $\mathcal{H}^{(n)}$ into factorial representations yields the <u>same</u> subspaces $\mathcal{H}^{(n)}_{[\ ]_\lambda}$. One can therefore denote by the same symbols $[\cdot]_\lambda$ the corresponding irreps of $U(k)$.

To summarize:

$$\text{for } S(n), \ [\ ]_{\mathcal{H}^{(n)}} \sim \oplus_\lambda u_\lambda [\ ]_\lambda \tag{1.15}$$

$$\text{for } U(k), \ [\ ]_{\mathcal{H}^{(n)}} \sim \oplus_\lambda s_\lambda [\ ]_\lambda \tag{1.15'}$$

where u_λ = dimension of the irrep $[\ ]_\lambda$ of $U(k)$, s_λ = dimension of the irrep $[\ ]_\lambda$ of $S(n)$.

The above is a theorem which is a leit-motiv of Weyl's book quoted in the introduction and is implicit in the two other books. When $n > k$, irreps of $U(k)$ are labeled only by those partitions of n such that $\Sigma\alpha_i \leq k$, i.e., the Young diagrams of irreps of $U(k)$ have at most k lines, but an arbitrary number n of squares; $n = 1$, $\square$ corresponds to the k-dimensional (= fundamental) representation of $U(k)$ and $n = 0$, "•" to the trivial (= one-dimensional) representation. For example, the representations of $U(2)$ are labeled $[\lambda_1,\lambda_2]$ with the integers $\lambda_1 \geq \lambda_2 \geq 0$.

The diagram of the contragredient representation of $[\lambda_1^{\alpha_1}]$ is $[\lambda_i^{\alpha'_i}]$ with $\Sigma\alpha_i < k$, $\alpha'_1 = k - \Sigma_1^p\alpha_i$, $\lambda'_1 = \lambda_1\alpha'_j = \alpha_{p+2-j}$, $\lambda'_i = \lambda_i - \lambda_{p+1-i}$ if $i,j > 1$. Of course it is shorter to say that the Young diagram of $[\lambda_i^{\alpha'_i}]$ is the complement (upside down) of that of $[\lambda_i^{\alpha_i}]$ in the rectangle of k lines of λ_1 squares.

SU(n) Representations. The restriction of an irrep of $U(k)$ to the subgroup $SU(k)$ of $U(k)$ is an irrep of $SU(k)$. Irreps of $U(k)$ whose Young diagrams differ only on the left by a rectangular block of columns of length k yield by restriction equivalent irreps of $SU(k)$. Taking into account this remark, one unambiguously label irreps of $SU(k)$ by Young diagrams. Moreover, this yields all inequivalent irreps of $SU(k)$.

Example. The equivalence classes of irrep of $SU(2)$ obtained by restriction of the irreps $[\lambda_1,\lambda_2]$ of $U(2)$ are given by the value of the integer $\lambda_1 - \lambda_2$. So their Young diagram can be written as a horizontal line of $\lambda_1 - \lambda_2$ squares. For $SU(2)$ irreps it is customary to use the symbol

$$D_j \quad \text{with} \quad j = \frac{1}{2}(\lambda_1 - \lambda_2) \tag{1.16}$$

where j is called the spin of the representation. $2j + 1$ is its dimension. The Casimir operator of D_j, as normalized by physicists, is $j(j + 1)\mathbf{1}$, which is twice that defined by (1.9). (Indeed, physicists take as Cartan-Killing form $1/2\ \mathrm{Tr}D(a)D(b)$).

We also recall the well known decomposition

$$D_{j_1} \otimes D_{j_2} = \oplus_{j=|j_1-j_2|}^{j_1+j_2} D_j \tag{1.17}$$

Note that all representations of $SU(2)$ are self contragredient.

Representations of the Adjoint Groups $SU(n)/Z_n$. The center of $SU(k)$ is Z_k, the cyclic group of k elements, so the adjoint group of $SU(k)$ is $SU(k)/Z_k$.

The representations of this group are those of $SU(k)$ whose Young diagram has a number of squares multiple of k. For example, $SU(2)/Z_2 = SO(3)$. Its representations are D_j with integral j. Their Young diagrams contain only one line of an even number $(= 2_j)$ of squares.

The adjoint representation of $SU(n)$ or its adjoint group is that on the space of its Lie algebra, it has dimension $n^2 - 1$ and label $[2, 1^{n-2}]$; it is equivalent to its contragredient.

Remark For Any Group. For any group G, let $\mathcal{H}^{(1)}$ be the space of a linear unitary representation (it may be reducible and $\dim \mathcal{H}^{(1)}$ may be infinite). As we saw $S(n)$ and G acts on $\mathcal{H}^{(n)} = \overset{n}{\otimes} \mathcal{H}^{(1)}$. Subspaces $\mathcal{H}^{(n)}_{[\ ]_\lambda}$ of primary representations of $S(n)$ are not in general subspaces of primary representation of G. Methods for knowing the nature of the G-representation of the different $\mathcal{H}^{(n)}_{[\ ]_\lambda}$ would be interesting for the physicists, especially in some case, for $\mathcal{H}^{(n)}_{[n]}$ (bosons) and $\mathcal{H}^{(n)}_{[1^n]}$ (fermions). Here is an example of a result, proven by A. Bohr, *Mat. Fys. Medd. Dan. Vid. Selsk*, 26 (No. 14), 16.

G is $SO(3)$, $\mathcal{H}^{(1)}$ is the five-dimensional Hilbert-space of the representation D_2. For any n the representation of $SO(3)$ on $\mathcal{H}^{(n)}_{[n]} = \overset{n}{\vee} \mathcal{H}^{(1)}$ does not contain D_1 in its reduction into direct sum of irreps. (Physically, a nucleus with spin 0 ground state has no spin 1 state corresponding to collective excitations.)

Of course we also can add that if an irrep of G appears on $\mathcal{H}^{(n)}$ only once, then it acts either on $\mathcal{H}^{(n)}_{[n]}$ or on $\mathcal{H}^{(n)}_{[1^n]}$.

1.5. More Algebras and More Tensor Operators. Pseudo Roots of $SU(n)$

An algebra on the vector space E is an element of $\mathrm{Hom}(E \otimes E, E)$. The algebra is symmetrical, (respectively, antisymmetrical) if it is an element of $\mathrm{Hom}(E \vee E, E)$, (resp., $\mathrm{Hom}(E \wedge E, E)$). Similarly we can define a co-algebra, symmetrical, antisymmetrical co-algebra as an element of $\mathrm{Hom}(E, E \otimes E)$, $\mathrm{Hom}(E, E \vee E)$ or $\mathrm{Hom}(E, E \wedge E)$.

If E is the space of a linear representation of G, elements of $\mathrm{Hom}(E \otimes E, E)^G$, resp., $\mathrm{Hom}(E, E \otimes E)^G$ are algebras, resp., co-algebras, whose group of automorphisms contains G.

When G is a semi-simple compact Lie group, a necessary condition for $\dim \mathrm{Hom}(E \otimes E, E)^G > 0$ is that the representation on E has a null weight. For example, for the space G of the adjoint representation, $\dim \mathrm{Hom}(G \wedge G, G)^G = 1$ for all simple compact Lie groups and the corresponding antisymmetrical algebra is the Lie algebra itself.

In V we shall see two examples of symmetrical algebras uniquely defined on a real irrep space E of $G = SU(3) \times SU(3)$, with $\dim \mathrm{Hom}(E \vee E, E)^G = 1$. For the adjoint representation of a simple compact Lie algebra $\dim \mathrm{Hom}(G \vee G, G)^G = 0$ or 1. It has the latter value for the $SU(n)$, $n > 2$. Let us give some properties of this symmetrical algebra of $SU(n)$, since it has been used in the physics literature of elementary particles, after its introduction by Gell-Mann. What follows is extracted from a preprint written in collaboration with L. A. Radicati.

Let G_n be the $n^2 - 1$ real vector space of the $n \times n$ traceless hermitian matrices x. The action of $u \in SU(n)$ on G_n (vector space of the Lie algebra) is $x \overset{u}{\rightsquigarrow} uxu^{-1} = uxu^*$. The euclidean scalar product

$$(x,y) = \frac{1}{2} \text{ trace } xy \tag{1.18}$$

is invariant (= $1/n$ the Cartan-Killing bilinear form). The $SU(n)$ Lie algebra law is

$$x \wedge y = -\frac{i}{2}(xy - yx) \equiv -\frac{i}{2}[x,y] \tag{1.19}$$

and the symmetrical algebra law is†

$$x \vee y = \frac{1}{2}\{x,y\} - \frac{2}{n}(x,y)\mathbb{1} \quad \text{where} \quad \{x,y\} = xy + yx \ . \tag{1.19'}$$

Note that for $n = 2$ it is trivial: $x \vee y = 0$.

In the physics literature (mainly for $n = 3$) one introduces an orthonormal basis $(e_i, e_j) = \delta_{ij}$ $(i,j = 1,\ldots,n^2 - 1)$ and uses traditionally the notation f_{ijk}, d_{ijk} for the structure constants

$$e_i \wedge e_j = \Sigma_k f_{ijk} e_k, \quad e_i \vee e_j = \Sigma_k d_{ijk} e_k \ .$$

Let us use $F(a)$, $D(a)$ for the linear mappings of G

$$F(a)x = a \wedge x, \quad D(a)x = a \vee x \tag{1.20}$$

(the matrices are $F(e_j)_{ik} = f_{ijk}$, $D(e_j)_{ik} = d_{ijk}$).

With the scalar product (1.18), $F(a)$ is antisymmetric and $D(a)$ is symmetric. D and F are tensor-operators $\in \mathrm{Hom}(G, L(G))^{SU(n)}$ so from the theorem at the end of 1.3, $\mathrm{trace}\ D(a) = 0$. As is well known, in the Lie algebra $SU(n)$, the centralizer of an element x, i.e., the set $\{y, y \wedge x = 0\}$ is a Lie subalgebra of dimension $\geq n - 1$. When its dimension is $n - 1$ it is abelian and it is called the Cartan subalgebra C_x of x. (All Cartan subalgebras are transformed into each other by the group.) C_x is spanned by the $n - 1$ linearly independent vectors x, $x \vee x$, $(x \vee x) \vee x = x \vee (x \vee x)$, $((x \vee x) \vee x) \vee x$, etc., up to $n - 1$ factors and C_x is also a subalgebra for the law " $\vee$ ". The roots of $SU(n)$ are solutions of the equation $r^n - (r,r)r^{n-2} = 0$. We shall normalize them by $(r,r) = 1$. In a

† This is not a Jordan algebra. However, one could have started from the n^2 dimensional representation realized by the $n \times n$ hermitian matrices. The corresponding symmetrical algebra is a Jordan algebra.

Cartan algebra C, there are $n(n-1)$ normalized roots r_k, (if r is a root, $-r$ is also one), for every $a \in C$, the spectrum of $F(a)$ has $n-1$ zeros for the eigen space C and on the orthogonal space $C^{\perp}$ the spectrum is the set

$$\text{Spectrum } F(a)\big|_{C^{\perp}} = \{i(a,r_k)\} \quad . \tag{1.21}$$

Define (for $n > 2$)

$$\sqrt{\frac{n-2}{n}}\, q_k = r_k \vee r_k = (-r_k) \vee (-r_k) \tag{1.22}$$

then

$$(q_k,q_k) = 1 \tag{1.22'}$$

and they are idempotents of the $\vee$-algebra

$$q_k \vee q_k = \frac{n-4}{\sqrt{n(n-2)}}\, q_k \quad . \tag{1.23}$$

We will call them "pseudo roots" (they are weights of $SU(n)$) for they satisfy for every $a \in C$

$$\text{Spectrum } D(a)\big|_{C^{\perp}} = \{\frac{n-2}{n}(q_k,a) = (a,r_k \vee r_k)\} \tag{1.24}$$

(all the eigen values have at least multiplicity 2).

Let us denote by $\lambda \in \mathrm{Hom}(G \wedge G,G)^{SU(n)}$, $\nu \in \mathrm{Hom}(G \vee G,G)^{SU(n)}$ the vector space homomorphisms

$$\lambda(x \otimes y) = x \wedge y,\ \nu(x \otimes y) = x \vee y$$

and consider their right inverse

$$\lambda \circ \lambda' = \text{Identity on } G,\ \nu \circ \nu' = \text{Identity in } G \quad . \tag{1.25}$$

Note that λ and λ' can be defined for any semi-simple Lie algebra. As we said λ' and ν' define co-algebras on G. If T is a G-tensor operator, using the mapping $\hat{T}$ of Diagram 2, one can define the G-tensor operators

$$T \wedge T = \hat{T} \circ \lambda' \quad \text{and} \quad T \vee T = \hat{T} \circ \nu' \tag{1.26}$$

and by recursion

$$(\ldots(T \tau_1 T) \tau_2 T)\ldots) \tau_k T \tag{1.26'}$$

where "τ_i" is either " $\wedge$ " or " $\vee$ ". For physicists who need to see coordinates, in $SU(3)$ octet space

$$\lambda'(e_i) = \Sigma_{j,k} - \frac{1}{3} f_{ijk} e_j \otimes e_k,\ \nu'(e_i) = \Sigma_{jk} \frac{3}{5} d_{ijk} e_j \otimes e_k \quad .\dagger$$

---

† The f_{ijk} and d_{ijk} are the structure constants introduced by Gell-Mann.

If we set $T(e_i) = T_i$, then

$$(T \wedge T)_i = \Sigma_{j,k} - \frac{1}{3} f_{ijk} T_j T_k, \quad (T \vee T)_i = \Sigma_{jk} \frac{3}{5} d_{ijk} T_j T_k \quad .$$

Note of course that we can define $T \wedge T$, $T \vee T$ for any real irreducible E-tensor operator when G is a compact group when $\dim \mathrm{Hom}(E T E, E)^G = 1$ (τ is $\wedge$ or $\vee$) since there is the irrep of G on E is orthogonal and leaves invariant a euclidean scalar product. Indeed, λ or ν are then surjective and are isomorphisms between $(\mathrm{Ker}\ \lambda)^{\perp}$ and E (resp., $(\mathrm{Ker}\ \nu)^{\perp}$ and E) so we can define their right inverse.

Let us consider the more particular case when the G-morphism T is F itself (see Equation (1.6)), i.e., the representation (up to the factor i) of the Lie algebra on $\mathcal{H}$. Then $F \wedge F = iF$. When $SU(3)$ is used for elementary particles, $F \vee F$ is often called the D-coupling operator (see 5.1b). For $SU(2)$, in order to follow the tradition started in elementary school, we denote by $\times$ the Lie algebra law (= vector product)

$$[F(\vec{a}), F(\vec{b})] = iF(\vec{a} \times \vec{b}) \tag{1.27}$$

and by ε_{ijk} the structure constants

$$e_i \times e_j = \Sigma_k \varepsilon_{ijk} e_k \quad . \tag{1.28}$$

So if $\vec{A}$ is a vector operator (with $A(e_i) = A_i$

$$(\vec{A} \times \vec{A})_i = \sum_{jk} \varepsilon_{ijk} A_j A_k = \frac{1}{2} \sum_{jk} \varepsilon_{ijk} [A_j, A_k] \tag{1.29}$$

Remark. Given two G-tensor operators A and B, we can also define

$$A \vee B = A \otimes B \circ \nu', \quad A \wedge B = A \otimes B \circ \lambda'$$

and in particular $\vec{A} \times \vec{B}$. This reduces to Equation (1.26) when $A = B$.

1.6. More on SU(2) and its Tensor Operators

For $SU(2)$ the symmetrical algebra $\vee$ on the adjoint representation G is trivial, $\dim(G \otimes G, G)^{SU(2)} = \dim(G \wedge G, G)^{SU(2)} = 1$

Much more generally, given any three irrep on E_{j_1}, E_{j_2}, E_{j_3},

$$\dim \mathrm{Hom}(E_{j_1} \otimes E_{j_2}, E_{j_3})^G = \Delta(j_1, j_2, j_3) = 0 \quad \text{or} \quad 1 \tag{1.30}$$

where $\Delta(j_1, j_2, j_3) = 1$ if $|j_1 - j_2| \leq j_3 \leq j_1 + j_2$ (triangular relation), 0 otherwise. This property, under an equivalent formulation, is called the Wigner-Eckart theorem by physicists, and groups with the property (1.30) have been called simply reducible by Wigner.

Let us give here two references that we shall quote often in this section.

A. *Quantum Theory of Angular Momentum* - a collection of reprints and original papers edited by L. C. Biedenharn and H. Van Dam, Academic Press, New York (1965).

B. *Spectroscopic and Group Theoretical Methods in Physics, Racah Memorial Volume*, North Holland, Amsterdam (1968).

In B p. 131-136, Wigner proves the following theorem for finite groups.

Theorem

Let G be a finite group and H a subgroup. The following conditions are equivalent

a) The restriction to H of any irrep of G is multiplicity free when decomposed into irrep of H;

b) The ring of conjugation classes by H of elements of G is abelian.

Let us explain a) and b) in more detail.

a) Given an irrep of G on $\mathcal{H}$, its restriction as a representation of H is generally reducible. To say that it is multiplicity free means that in its decomposition into irreps of H no such irrep appears more than once. Equivalently, one can say that the commutant of the representation of H (i.e., the set of all bounded elements of $L(\mathcal{H})$ which commute with every operator of the representation of H; this set is an algebra) is an abelian algebra. That last condition can be used as definition of multiplicity free for any linear representation of any group.

b) Given $a \in G$, the conjugation class of a by H is the set $A = \{hah^{-1}, \forall h \in H\}$. Given two such classes we define as $A \cdot B$ the set $\{ab, a \in A, b \in B\}$. Condition b) states that for any pair of classes, $A \cdot B = B \cdot A$. It seems easy to extend Wigner's proof to compact groups. Examples of pairs of group and subgroup which satisfy this theorem: $S(n)$ and $S(n-1)$, $U(n)$ and $U(n-1)$.†

From the group law one verifies that the direct product $SU(2) \times SU(2)$ and its diagonal subgroup satisfies b); by Wigner's theorem this implies (1.30). It would be interesting to extend, if possible, Wigner's proof to type I locally compact groups.††

---

† After the lecture, Professor G. Mackey gave a proof for compact groups, using his theory of induced representations.

†† Wigner proved in 1941 (paper reproduced in reference A., see 1.6) for finite groups another property equivalent to a) and b). Let $\zeta(g)$ the number of square roots of g in the finite group G, and $v(g)$ the number of elements of G which commute with g. In a finite group $\sum_{g \in G} v(g)^2 - \zeta(g)^3 \geq 0$. The equality occurs if and only if G is simply reducible.

Another property of SU(2) that we have mentioned is that any irrep of SU(2) is equivalent to its contragredient. For any irrep D_j of SU(2) this defines an isomorphism $C: E_j \overset{C}{\to} E'_j$ between the E_j space of D_j and its dual E'_j with the canonical property

$$E_j \overset{C}{\to} E'_j;\ C^T = (-1)^{2j}C \tag{1.31}$$

for the transposed C^T of C. Physicists normalize C by

$$C^T C = 1 \quad . \tag{1.32}$$

We are by now ready to give you a glimpse of the calculus developed independently by Wigner and Racah for the necessity of physics in order to exploit completely rotational invariance in atomic physics (and later on in nuclear physics and many other branches of quantum physics). Many of the numbers observed in atomic spectra (spacing between neighbors in a family of lines, relative intensity of these lines, etc.) turn out to be algebraic functions of the coefficients defined by Wigner and Racah. Since they are so useful, the literature on Wigner and Racah coefficients is abundant. They have been found to possess unexpected symmetries, there are unproven conjectures on them. However, the language of this physics folklore seems to be unknown to the mathematically minded ethnographer.

I hope there will be much discussion on this subject in this Rencontres. The rest of this section should help to start it.

To write Wigner's "three j" coefficients, physicists choose a base in each Hilbert space E_j, carrier of the irrep D_j, for every j. It is formed by the eigen vectors of a chosen U(1) (Cartan subgroup of SU(2)) ordered in terms of decreasing eigen value μ (going from j to -j by integer steps). It is obvious that most of their properties are base independent.

Consider an element of the one-dimensional vector space

$$(E'_{j_1} \otimes E'_{j_2} \otimes E_{j_3})^G = \mathrm{Hom}(E_{j_1} \otimes E_{j_2}, E_{j_3})^G \tag{1.33}$$

and denote it

$$(\mathring{j}_1 \mathring{j}_2 \mathring{j}_3) \quad . \tag{1.34}$$

The isomorphism C and its inverse, defined in (1.31), (1.32), transform the tensor (1.20) into its following images

$$(\mathring{j}_1 \mathring{j}_2 \mathring{j}_3) \in (E'_{j_1} \otimes E'_{j_2} \otimes E'_{j_3})^G = \mathrm{Hom}(E_{j_1} \otimes E_{j_2} \otimes E_{j_3}, \mathbb{C})^G \tag{1.35}$$

$$(\mathring{j}_1 \mathring{j}_2 \mathring{j}_3) \in (E_{j_1} \otimes E_{j_2} \otimes E_{j_3})^G = \mathrm{Hom}(\mathbb{C}, E_{j_1} \otimes E_{j_2} \otimes E_{j_3})^G \tag{1.36}$$

$$(\mathring{j}_1 \mathring{j}_2 \mathring{j}_3) \in (E'_{j_1} \otimes E_{j_2} \otimes E_{j_3})^G = \mathrm{Hom}(E_{j_1}, E_{j_2} \otimes E_{j_3})^G \tag{1.37}$$

and so on.

Equations (1.35), (1.36) show that $(\mathring{j}_1\mathring{j}_2\mathring{j}_3)$ (resp., $(j_{\circ 1}j_{\circ 2}j_{\circ 3})$) belong to a one-dimensional representation of the permutation group of the three factor spaces labeled by j_1, j_2, j_3. Computation shows that the representation is

$$\begin{array}{l} \square\square\square \text{ (symmetrical) if } j_1 + j_2 + j_3 \text{ is even} \\ \begin{smallmatrix}\square\\\square\end{smallmatrix} \text{ (antisymmetrical) if } j_1 + j_2 + j_3 \text{ is odd .} \end{array} \tag{1.38}$$

The composition of the two homomorphisms

$$E_{j_1} \otimes E_{j_2} \otimes E_{j_3} \xrightarrow{(\mathring{j}_1\mathring{j}_2\mathring{j}_3)} \mathbb{C} \xrightarrow{(j_{\circ 1}j_{\circ 2}j_{\circ 3})} E_{j_1} \otimes E_{j_2} \otimes E_{j_3} \tag{1.39}$$

is an element of $\mathrm{Hom}(E_{j_1} \otimes E_{j_2} \otimes E_{j_3}, E_{j_1} \otimes E_{j_2} \otimes E_{j_3})^G$ that we denote $(\mathring{j}_1\mathring{j}_2\mathring{j}_3)(j_{\circ 1}j_{\circ 2}j_{\circ 3})$.

Wigner proved (see reference A and Equation (24.18b) of his book quoted in the Introduction)

$$\int_{SU(2)} D_{j_1}(g) \otimes D_{j_2}(g) \otimes D_{j_3}(g) d\mu(g) = (\mathring{j}_1\mathring{j}_2\mathring{j}_3)(j_{\circ 1}j_{\circ 2}j_{\circ 3}) \tag{1.40}$$

where $d\mu(g)$ is the invariant $SU(2)$ measure of mass $\int_{SU(2)} d\mu(g) = 1$.

This also defines for you, up to a sign, which element of the one-dimensional vector space $(E_1 \otimes E_2 \otimes E_3)^G$ has been chosen by physicists for $(\mathring{j}_1\mathring{j}_2\mathring{j}_3)$.

Of course tensors can have their indices contracted (notation $\times$); for instance $(\mathring{a}\mathring{b}\overset{\times}{\circ})(\overset{\times}{c}\mathring{p}\mathring{q})$ is the composed homomorphism

$$E_a \otimes E_b \otimes E_p \otimes E_q \xrightarrow{(\mathring{a}\mathring{b}\mathring{c}) \otimes I_p \otimes I_q} E_c \otimes E_p \otimes E_q \xrightarrow{(\mathring{c}\mathring{p}\mathring{q})} \mathbb{C} \quad .$$

Wigner's notation is very handy!

Note that equation (1.40) yields

$$\int_{SU(2)} \chi_{j_1}(g)\chi_{j_2}(g)\chi_{j_3}(g) d\mu(g) = (\overset{\times}{j}_1\overset{\times}{j}_2\overset{\times}{j}_3)(\overset{j_1}{\times}\overset{j_2}{\times}\overset{j_3}{\times}) = \Delta(j_1, j_2, j_3) \tag{1.41}$$

where χ_j is the character of D_j.

Approximate expression, asymptotic expression, for large j's exist for the components of those tensors in the basis described above (see the thorough bibliography of reference A (see 1.6)). Regge (paper reproduced in A (see 1.6)) has found a 72 elements group of symmetry $\sim \mathrm{Aut}(S(3) \times S(3))$ for the set of components

$\begin{pmatrix} \mu_1\mu_2\mu_3 \\ j_1 j_2 j_3 \end{pmatrix}$ of $(\overset{\circ}{j}_1\overset{\circ}{j}_2\overset{\circ}{j}_3)$.†

In 1941 Racah and Wigner (both papers reproduced in A (see 1.6)) introduced a "six j" symbol (numerical function of six irrep of SU(2)), also known as recoupling coefficient. It is canonical.

Consider the sequence of SU(2)-homomorphisms

$$E_e \xrightarrow{(\overset{\circ}{e}fa)} E_f \otimes E_a \xrightarrow{I_f \otimes (\overset{\circ}{a}bc)} E_f \otimes E_b \otimes E_c \xrightarrow{(fb\overset{\circ}{d}) \otimes I_c}$$

$$E_d \otimes E_c \xrightarrow{(ad\overset{\circ}{e})} E_e \quad . \tag{1.42}$$

Since E_e is the carrier of an irrep, this SU(2)-homomorphism must be a multiple of the identity operator on E_e.

Its trace defines (up to a sign that I do not guarantee here) the six j's symbol

$$\begin{Bmatrix} abc \\ def \end{Bmatrix} = (-1)^{b+c-d+e+f} (\overset{\times}{e}f\overset{\times}{a})(\overset{\times}{a}b\overset{\times}{c})(\overset{\times\times}{f}b\overset{\times}{d})(\overset{\times\times}{a}c\overset{\times}{e}) \quad . \tag{1.43}$$

Wigner has shown that for given a, b, d, e, $\{^{abc}_{def}\}$ is an orthogonal matrix with indices c, f. He also proved the relation (in his book, Chapter 24)

$$\begin{Bmatrix} abc \\ def \end{Bmatrix}^2 = \iiint \chi_a(r)\chi_b(s)\chi_c(t)\chi_d(st^{-1})\chi_e(tr^{-1})\chi_f(rs^{-1})d\mu(r)d\mu(s)d\mu(t) \quad .$$

Asymptotically its value is a rapidly oscillating function of some variables, but an average over some range of one argument yields, when a, b, c, d, e, f form the edges of a tetrahedron, the asymptotic value

$$\begin{Bmatrix} abc \\ def \end{Bmatrix}^2 \to (24\pi V)^{-1}$$

where V is the volume of the tetrahedron.

G. Ponzano and T. Regge (in reference B (see 1.6), first paper) have conjectured precise asymptotic formulae for $\{^{abc}_{def}\}$ whether or not the value of the arguments can be the length of the edges of a tetrahedron.

Furthermore, Regge has found (paper reproduced in A (see 1.6)) the largest linear group acting on the Z-module generated by the symbols a/2, b/2, c/2, d/2, e/2, f/2 and having $\{^{abc}_{def}\}$ as invariant. It is the group $S(3) \times S(4)$ which includes the permutation group of the columns.

---

† Most of these symmetries appear naturally (see Bargmann's paper in A (see 1.6)); for the others see G. Flamand, *Ann. Inst. H. Poincaré*, 7, 353 (1967).

Neatest and very symmetrical expressions for the $(j_1 j_2 j_3)$ and $\begin{Bmatrix} j_1 j_2 j_3 \\ j_4 j_5 j_6 \end{Bmatrix}$ symbols can be found in the paper of V. Bargmann (last paper reproduced in book A (see 1.6)) who uses Hilbert spaces of analytic functions as spaces of SU-2 irreps.

2. ATOMIC AND MOLECULAR PHYSICS

2.1. Group Theory and Atomic Physics

The application of group theory to atomic physics is essentially of this type; only the even part $f_+(\vec{r}) = 1/2(f(\vec{r}) + f(-\vec{r}))$ (respectively, the symmetric part $f_+(\vec{r}_1,\vec{r}_2) = 1/2(f(\vec{r}_1,\vec{r}_2) + f(\vec{r}_2,\vec{r}_1))$ of $f(\vec{r})$ (resp., $f(\vec{r}_1,\vec{r}_2)$) contributes to the integral over the whole space $\int f(\vec{r})d^3\vec{r}$ (resp., $\int f(\vec{r}_1,\vec{r}_2)d^3\vec{r}_1 d^3\vec{r}_2$). This is the explanation of two empirically known facts (before 1926), the Laporte selection rule for atomic spectra and the partition of the helium spectrum into two independent subsets (attributed to ortho and parahelium). Of course, these examples are the simplest because they are based on invariance under a two element group (Z_2). We will have to use invariance under $SO(3)$, $S(n)$ and $U(2)$ for atoms and invariance under subgroups of $SO(3)$ for molecules.

2.2. The Correspondence Principle

We had a general description of quantum mechanics, but now we have to know how to study a given physical system. There does not seem to exist an axiomatic formulation of the question, so here again, physics is still an art! However, when the system under consideration, with a finite number of degrees of freedom can be described by classical Hamiltonian mechanics, the "correspondence principle" tells physicists how to treat it quantum-wise.

Let $h(p_k,q_e)$ be the classical Hamiltonian and

$$\frac{dp_k}{dt} = \dot{p}_k = -\frac{\partial h}{\partial q_k},\quad \dot{q}_\ell = \frac{\partial h}{\partial p_\ell}$$

the Hamiltonian equations. The corresponding observables P_ℓ, Q_ℓ in quantum mechanics form the abstract algebra with unit

$$P_k Q_\ell - Q_\ell P_k = [P_k,Q_\ell] = -i\hbar\delta_{k\ell}\mathbb{1}$$

$$[P_k,P_\ell] = 0 = [Q_k,Q_\ell] \tag{2.1}$$

where $2\pi\hbar$ is Planck's constant.

In the cases we shall study, h is a sum $h = h' + h''$ where h' is a function of the p's and h'' is a function of the q's. Then $H = H' + H''$ where H' and H'' are the same functions respectively of the P's and the Q's. There is yet no synthetic formulation of quantum mechanics as there is for classical mechanics by symplectic manifolds (see, however, work in progress by Kostant, Souriau). We also know that the relations between the classical and quantum treatment of the same problem are not simple (see e.g., Van Hove's work in 1951 comparing the two automorphism groups).

The Hamiltonian operator is the generator of the group of time translations

$$[H,Q_k] = i\hbar\dot{Q}_k, \quad [H,P_\ell] = i\hbar\dot{P}_\ell \quad . \tag{2.2}$$

A representation of the algebra defined by (2.1) and (2.2) was obtained, independently of Heisenberg's work by Schrödinger, using the concept of de Broglie's waves. Indeed, the algebra (2.1) is realized by self-adjoint operators of $L(\mathcal{H})$ where $\mathcal{H}$ is the Hilbert space of square integrable functions $\Psi(q_i)$. Then

$$Q_k\Psi = q_k\Psi, \quad P_\ell\Psi = \frac{\hbar}{i}\frac{\partial}{\partial q_\ell}\Psi \quad . \tag{2.3}$$

The Ψ are also functions of the time (t) and the Schrödinger equation is

$$H = i\hbar\frac{\partial}{\partial t} \quad . \tag{2.4}$$

This representation raises some analysis problems. On the other hand, von Neumann's theorem (J. von Neumann: "Die Eindeutigkeit der Schrödingerschen Operatoren", *Math. Annalen*, 104, 570 (1937)) tells us that all irreducible representations of the algebra defined by Equation (2.1) are equivalent when e^{iP_k}, e^{iQ_ℓ} are realized by unitary operators.†

Quantum mechanics was also discovered by Dirac who gave the neatest formulation of the "correspondence principle".†† In classical Hamiltonian mechanics one has also a Lie algebra, that of the Poisson brackets (P.B). Let f, g be two functions of the p's and the q's,

$$\text{P.B.}(f,g) = \Sigma_\ell \frac{\partial f}{\partial q_\ell}\frac{\partial g}{\partial q_\ell} - \frac{\partial f}{\partial q_\ell}\frac{\partial g}{\partial p_\ell} \quad . \tag{2.5}$$

---

† For systems with an infinite number of degree of freedom, as they appear in statistical mechanics and field theory, this is far from true. Infinities of irreducible representations of (2.2) have been given first by Friedrichs, Van Hove, Gärding and Wightman, Segal and several other physicists and mathematicians. An excellent thin book on the subject is by A. Guichardet, *Algèbres d'Observables Associées aux Relations de Commutation*, Armand Colin, Paris, (1969). (See also G. Mackey, *Duke Math. J.*, 16, 313 (1949)).

†† Historically, the expression "correspondence principle" had a more restricted meaning.

The Lie algebra of the corresponding quantum observable is

$$[F,G] = i\hbar \quad \text{quantum observable of} \quad P.B.(f,g) \quad . \tag{2.5'}$$

As you surely know $|\Psi|^2 \Pi dq_k$, where Ψ is a solution of Schrödinger Equation (2.4), is the density of probability to find the system at the coordinate $\{q_k\}$. This of course appeals very much to physicists. As mathematicians you will like just as well to work with the abstract algebra. As a short, but fundamental illustration of the use of that algebra, let us prove the Heisenberg uncertainty relations.

Let A, B be the self adjoint operators corresponding to the observables a, b. If $x>$ is a given state of the physical system we study, we have seen that $<xAx>$ is the expectation value of "a" for $x>$ and the mean square dispersion of probability is given by

$$(\Delta a)_x = |<x,(A - <x,Ax>)^2 x>|^{1/2} = |<x,\hat{A}^2 x>|^{1/2} = ||\hat{A}x|| \tag{2.6}$$

where

$$\hat{A} = A - \mathbb{1}<xAx> \quad . \tag{2.6'}$$

By Schwarz' inequality

$$(\Delta a)_x (\Delta b)_x = ||\hat{A}x|| \cdot ||\hat{B}x|| \geq |<\hat{A}x,\hat{B}x>| \geq \frac{1}{2}|<x,[A,B]x>| \tag{2.7}$$

If A and B satisfy the same canonical relations as the P's and Q's we do obtain

$$(\Delta a)_x (\Delta b)_x \geq \frac{1}{2}\hbar \quad . \tag{2.8}$$

2.3 Particle of Mass m in a Spherically Symmetric Potential

Let V(r) be a spherical symmetric potential, where r denotes $|\vec{r}|$. The Hamiltonian of the particle is

$$H = \frac{1}{2m}\vec{P}^2 + V(r) \quad , \tag{2.9}$$

which is invariant under the orthogonal group O(3). Using the vocabulary of 1.3, H, $\vec{P}^2$, V(r) are "scalar operators"; $\vec{P}$, $\vec{R}$ and $\vec{R} \times \vec{P} = \vec{L}$ are (polar and axial) vector operators. (So we put an arrow on them!) If, $\vec{a}$, $\vec{b}$, etc., are vectors of the three-dimensional vector space E_3 of the adjoint representation of O(3), we should write the canonical commutation relations (2.1)

$$[\vec{P}(\vec{a}),\vec{Q}(\vec{b})] = i\hbar\mathbb{1}\,\frac{1}{2}\beta(\vec{a},\vec{b}) = i\hbar\vec{a},\vec{b}\mathbb{1} \quad , \tag{2.10}$$

where the Cartan-Killing form β has been defined in (1.11).†

† See also the Appendix on commutation relations at the end of 2.

From (2.10) and the definition by the correspondence principle of the angular momentum operator (see end of 1.5), $\vec{L} = \vec{R} \times \vec{P}$, we obtain

$$[\vec{L}(\vec{a}),\vec{L}(\vec{b})] = i\hbar\vec{L}(a \wedge b) \quad , \tag{2.11}$$

which confirms that the vector-operator representing the angular momentum is the representation (up to i) of the O(3) Lie algebra on the Hilbert space of our problem.

Some physicists write $\vec{P} \cdot \vec{a}$, $\vec{L} \cdot \vec{n}$ for $\vec{P}(\vec{a})$, $\vec{L}(\vec{n})$. But do not be surprised if in all physics text books an orthonormal basis of vectors notations Q_i, P_j, L_k are used for $\vec{Q}(\vec{e}_i)$, $\vec{P}(\vec{e}_j)$, $\vec{L}(\vec{e}_k)$, etc.

The operators corresponding to the observables which are constants of motion generate the algebra $\{H\}'$, the commutant of H. Hence, the equation that one deduces from (2.10) and the definition of $\vec{L}$

$$\vec{a} \in E_3, \ [L(\vec{a}),H] = 0 \quad \text{or symbolically} \quad [\vec{L},H] = 0 \tag{2.12}$$

means both that the Hamiltonian is invariant under rotations and that the angular momentum is a constant of motion.

The Casimir operator (with the physicists' normalization) of O(3) is $\vec{L}^2 = \Sigma_{i=1}^3 L_i^2$. As is well known, its values for irreducible representations of SU(2) are $j(j+1)\hbar^2$ where $2j$ is an integer ≥ 0; and $2j+1$ is the dimension of the representation. Only integer values of j appear in the SO(3) irrep. When the state vector is an eigenvector of $\vec{L}^2$ with eigenvalue $j(j+1)\hbar^2$, we say shortly that the corresponding angular momentum is $j\hbar$.

2.4. The Hydrogen Atom

Consider two particles of mass m_1, m_2 electric charge Ze, -e (Z is a positive integer). The total Hamiltonian for this system of two particles is

$$h_{tot.} = \frac{\vec{p}_1^2}{2m_1} + \frac{\vec{p}_2^2}{2m_2} - \frac{Ze^2}{r} \quad , \tag{2.13}$$

where $r = |\vec{r}|$ with $\vec{r} = \vec{r}_2 - \vec{r}_1$.

Introduce the center of mass

$$\vec{r}_0 = (m_1\vec{r}_1 + m_2\vec{r}_2)(m_1 + m_2)^{-1} \tag{2.14}$$

and $\vec{r}$ as new variables instead of $\vec{r}_1$ and $\vec{r}_2$; let $\vec{p}_0$ and $\vec{p}$ the conjugate variables. Then

$$h_{tot.} = \frac{\vec{p}_0^2}{2(m_1 + m_2)} + (\frac{\vec{p}^2}{2m} - \frac{Ze^2}{r}) = h_{cm} + h \quad , \tag{2.15}$$

where

$$m = m_1m_2(m_1 + m_2)^{-1} \quad . \tag{2.15'}$$

The motion of the center of mass is described by h_{cm} while h corresponds to the internal energy of the system. So quantum-wise, we have to study the spectrum of

$$H = \frac{\vec{P}^2}{2m} - \frac{Ze^2}{R} \quad , \tag{2.16}$$

for obtaining the energy of the hydrogen atom levels. The first quantum study of the hydrogen atom was made by Pauli, *Z. Phys.*, 36, 336 (1926) before Schrödinger's equation was published. Pauli did study the abstract algebra generated by $\vec{R}$, $\vec{P}$, H and Equations (2.1), (2.2), and (2.15). The angular momentum $\vec{L} = \vec{R} \times \vec{P}$ is a constant of motion. Another constant of motion is the Runge-Lenz vector

$$\vec{A} = \frac{1}{2}(\vec{L} \times \vec{P} - \vec{P} \times \vec{L}) + \frac{\lambda}{R}\vec{R} \quad \text{with} \quad \lambda = mZe^2 \quad . \tag{2.17}$$

Note that

$$\frac{1}{2}(\vec{L} \times \vec{P} - \vec{P} \times \vec{L}) = (\vec{R} \cdot \vec{P})\vec{P} - \vec{R}(\vec{P}^2) - i\hbar\vec{P} = \vec{P}(\vec{P} \cdot \vec{R}) - (\vec{P}^2)\vec{R} + i\hbar\vec{P} \tag{2.18}$$

so we can check that

$$[\vec{A},H] = 0, \ [\vec{L},H] = 0 \quad . \tag{2.19}$$

We recall that ε_{ijk} = sign of the permutation $\binom{123}{ijk}$ or 0 if two indices are equal. From now on we will use the Einstein summation convention, i.e., summation of repeated indices is implied, and we find

$$[L_i,L_j] = i\hbar\varepsilon_{ijk}L_k, \ [L_i,A_j] = i\hbar\varepsilon_{ijk}A_k \tag{2.20}$$

$$[A_i,A_k] = -i\hbar 2mH\varepsilon_{ijk}L_k \tag{2.21}$$

$$\vec{L} \cdot \vec{A} = \vec{A} \cdot \vec{L} = 0 \tag{2.22}$$

$$\vec{A}^2 - 2mH(\vec{L}^2 + \hbar^2) = (Ze^2m)^2\mathbb{1} \quad . \tag{2.23}$$

Let us just consider the bound states of the hydrogen atom. They correspond to the spectrum of $H < 0$. Let P_- be the projector on the bound states. For any X write $X^- = XP_-$. From (2.19) when X is $\vec{L}(\vec{a})$ or $\vec{A}(\vec{b})$,

$$P_-XP_- = XP_- = P_-X = X^- \quad .$$

Furthermore, $-2mH^-$ is an inversible positive operator. Let $(-2mH^-)^{-1/2}$ be the positive square root of its inverse and define $K_i^- = A_i^-(2mH^-)^{-1/2}$. Then Equations (2.20′) to (2.23) read

$$[\tfrac{1}{\hbar}L_i^-,\tfrac{1}{\hbar}L_j^-] = \tfrac{i}{\hbar}\varepsilon_{ijk}L_k^-, \ [\tfrac{1}{\hbar}L_i^-,\tfrac{1}{\hbar}K_j^-] = \tfrac{i}{\hbar}\varepsilon_{ijk}K_k^- \tag{2.20′}$$

$$[\tfrac{1}{\hbar}K_i^-,\tfrac{1}{\hbar}K_j^-] = \tfrac{i}{\hbar}\varepsilon_{ijk}L_k^- \tag{2.21′}$$

$$\vec{L}^- \cdot \vec{K}^- = \vec{K}^- \cdot \vec{L}^- = 0 \tag{2.22′}$$

$$\tfrac{1}{\hbar 2}K^{-2} + \tfrac{1}{\hbar 2}\vec{L}^{-2} = (\tfrac{Ze^2m}{\hbar})^2(-2mH)^{-1} \quad . \tag{2.23′}$$

We last define

$$\vec{J}^{(\pm)} = \frac{1}{2\hbar}\vec{L}^{(-)} \pm \frac{1}{2\hbar}\vec{K}^{(-)} \tag{2.24}$$

so the previous equations read

$$[J_i^{(\pm)}, J_j^{(\pm)}] = i\varepsilon_{ijk}J_k^{(\pm)}, \quad [J_i^{(+)}, J_j^{(-)}] = 0 \tag{2.25}$$

$$\vec{J}^{(+)2} = \vec{J}^{(-)2} = \frac{1}{4}\left(\left(\frac{Ze^2m}{\hbar}\right)^2(-2mH)^{-1} - 1\right) \quad . \tag{2.26}$$

The spectrum of this operator is $j(j+1) = (n^2-1)/4$ with $2j+1 = n$, positive integer. So the energy spectrum of the bound states of the hydrogen atom is

$$(n \text{ positve integer}, \ \varepsilon_n = -\frac{(Ze^2)^2}{\hbar}\frac{m}{2n^2} = \frac{-Z^2}{n^2}\left(\frac{e^2}{\hbar c}\right)^2\frac{mc^2}{2} = \frac{-1}{2n^2}(Z\alpha)^2mc^2) \tag{2.27}$$

where

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137.03\ldots} \tag{2.28}$$

in rationalized units of charge, α is the fine structure constant, a dimensionless fundamental constant of physics.

Some Physical Comments. The ratio binding energy/electron rest-mass energy is the number

$$\frac{\varepsilon_n}{mc^2} = -\frac{(Z\alpha)^2}{2n^2} \quad .$$

The value of every physical observable we can compute will appear as the product of a pure number and the quantity of same physical dimensions built with the constants e, $\hbar$, m, c. Example: length $\hbar/mc = 3.86 \times 10^{-11}$ cm; energy $mc^2 = .51 \times 10^6$ eV; time $\hbar/mc^2 = 1.28 \times 10^{-21}$ sec. The pure number is a function of α only. It is the value of the observable in the unit system $\hbar = m = c = 1$ that we will use, and α is the value of e^2 in this system. For instance

$$\left\langle\frac{1}{R}\right\rangle = Z\alpha\frac{mc}{\hbar} \sim \left(\frac{1}{2}\ 10^{-8}\ \text{cm}\right)^{-1} = \left(\frac{1}{2}\ \text{Ångström}\right)^{-1} \quad .$$

We have studied not only the bound state of the hydrogen atom $p^+e^-(m_p = 1836\ m_e)$ (the nucleus can also be a deuteron $\sim 2m_p$), but also that of positronium $e^+e^-(m_1 = m_2)$, munomium $\mu^+e^-(m_\mu = 207\ m_e)$, μ-atom, Π-atom, ionized Helium ion He^+, etc.

More On The Group Aspect. The states of energy ε_n are eigen states of $\vec{J}^{(+)2}$ and $\vec{J}^{(-)2}$ and they form the space $\mathcal{K}_n$ of the irrep (j,j) of $SO(4)$; $\mathcal{K}_n$ has dimension

$$(2j+1)^2 = n^2 \quad . \tag{2.29}$$

The Lie algebra of the physical rotation $(\vec{L})$ is the diagonal of $SU(2) \oplus SU(2) = SO(4)$, so the representation of the rotation group in $\mathcal{H}_n$, space of the irrep $(j,j) = (\frac{n-1}{2}, \frac{n-1}{2})$ of $SO(4)$ reduces to

$$(j,j)\big|_{SO(3)} = \oplus_{\ell=0}^{2j} D_\ell \tag{2.30}$$

i.e.,

$$\ell = 0, 1, \ldots, n-1 \quad . \tag{2.30'}$$

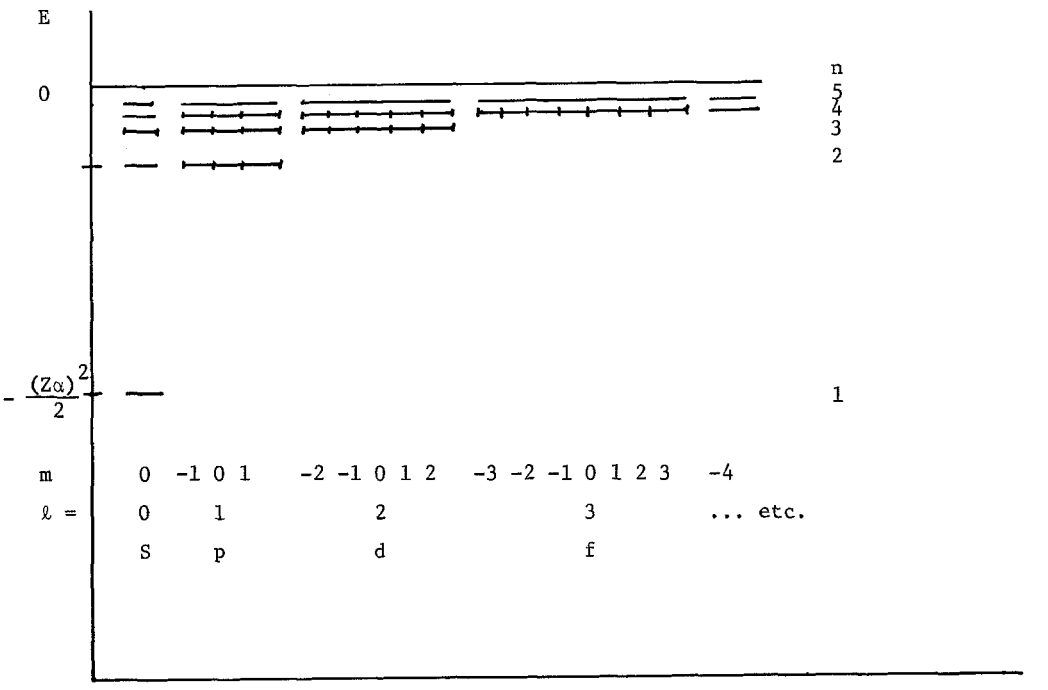


FIGURE 2.1. LOWEST STATES, IN A LINEAR ENERGY SCALE OF THE HYDROGEN ATOM

There is an infinite number of states with energy between $-\varepsilon$ and 0 because of the dependence in r^{-1} of the potential for $r \to \infty$. The eigenvectors of the abelian algebra generated by H, $\vec{L}^2$, $\vec{L}(\vec{e}_3)$ form an orthogonal basis for the Hilbert space of bound states. A complete set of labels for them is the quantum numbers n, ℓ, m; n = positive integer, ℓ and m integers $0 \le \ell \le n-1$, $-\ell \le m \le \ell$, corresponding to the eigenvalues $Z(\alpha)^2/2n^2$, $\ell(\ell+1)$, m of $(mc^2)^{-1}H$, $\hbar^{-2}\vec{L}^2$, $\hbar^{-1}\vec{L}(\vec{e}_3)$

Note that the trivial representation appears once only for each n, and from the Frobenius reciprocity theorem we know that

$$\oplus_{2j=0}^{\infty}(j,j) = U^{D_0} \tag{2.31}$$

the induced representation of $SO(4)$ by the trivial representation of $SO(3)$. In

other words, $P_-\mathcal{H} = L_2$ (functions on S_3) since the sphere S_3 is the homogeneous space $SO(4)/SO(3)$. (This was exploited by V. Fok, Z. *Physik*, 98, 145 (1935), see also L. Hulthen, Z. *Physik*, 86, 21 (1933).)

From Mackey's theory of induced representations, (2.31) is also the content of the representation of $R^4 \square SO(4) = E_4$ (euclidean group in four dimensions) induced by the trivial representation of $R^4 \square SO(3)$, the stabilizer of any chosen vector $\neq 0$ of R^4. This is an irreducible representation of E_4. We can also consider $P_-\mathcal{H}$ as the space of an irrep of $SO(4,1)$ (obtained by deformation of the irrep of E_4 considered above). However, the physical meaning of the generators (representing the elements of Lie algebra) of E_4 or $SO(4,1)$ is not very transparent!

The spectrum of H on $P_+\mathcal{H}$, (positive energy) is $(+0,\infty)$. One cannot speak of eigenvectors of H for the positive energy = unbounded states of a proton and an electron and one has to study their scattering. However, since $HP_+ = P_+H$ is a positive operator one can define $\vec{K}^+ = \vec{A}P_+(2mHP_+)^{-1/2}$ and $1/\hbar\ \vec{L}P_+$ and $1/\hbar\ \vec{K}^+$, which generate a $SO(3,1)$ Lie algebra as was noted and exploited by V. Bargmann, Z. *Physik*, 99, 576 (1936). Thus, $P_+\mathcal{H}$ is a direct integral of (infinite dimensional unitary) irreps of $SO(3,1)$, the Lorentz group. It is also the space of an irrep of the inhomogeneous $SO(3,1)$ that we call the Poincaré group (it is an irrep of spin zero and fixed mass $m > 0$).

What we observe mainly in atoms are their emission or absorption of photons of frequency

$$\nu = \frac{1}{2\pi\hbar}(\varepsilon_{n_1} - \varepsilon_{n_2}) \quad . \tag{2.32}$$

So the wave length is

$$\frac{c}{\nu} = \pi(Z\alpha)^{-2}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)^{-1}$$

which is $\sim \frac{4}{Z\alpha}\, 10^3$ times the "size" of the atom.

All large enough frequencies of (2.32) were already seen in the spectrum of atomic hydrogen. In fact, there is a "fine structure" which corresponds to a relative splitting in the energy of the states with $\ell \neq 0$ of the order of $\alpha^2 \sim 1/2\ 10^{-4}$.

The theory even predicts the intensity of the spontaneous emission of photons. Since its wave length is large compared to the atomic size, the light wave is a dipole emission† and the probability for spontaneous emission of a photon from

---

† Electromagnetic waves, predicted by Maxwell's equations, were produced by Hertz with an oscillating electric dipole. An example of such a dipole would be a charge $-e$ rotating around a charge $+e$ with a given frequency ν. That light was an electromagnetic wave was a Maxwell hypothesis and Selenyi, by clever experiments, verified in 1913 that light emitted by atoms was a dipole radiation. (Higher multipoles occur for more complicated charge distributions for which the

state $x>$ to state $y>(E_x > E_y)$ is

$$\lambda_{xy} = \frac{4}{3}(E_x - E_y)^3 |<x\vec{e}\vec{R}y>|^2 \quad . \tag{2.33}$$

(Note that $<x\vec{e}\vec{R}x>$ is the expectation value of the electric dipole of a state and, as we shall see, it vanishes.) The intensity of the emitted light by N_x atoms in state $x>$ is

$$i_{xy} = N_x \frac{4}{3}(E_x - E_y)^4 e^2 \Sigma_k \mathrm{Tr} P_x R_k P_y R_k \quad . \tag{2.34}$$

The Wigner-Eckhart theorem (see beginning of 1.6) predicts that for all vector operators, the matrix elements between two given eigenstates $x>$, $y>$ of $\vec{L}^2$, are colinear.

Example. For $x>$, $y>$ eigenstates of H

$$\frac{-i\hbar}{m} <x\vec{P}y> = <x[H,\vec{R}]y> = (E_x - E_y)<x\vec{R}y> \quad . \tag{2.35}$$

Consider from now on states which are eigenstates of $\vec{L}^2$ (eigenvalue $\ell(\ell+1)$). Note that

$$<x\vec{L}y> = 0 \quad \text{if} \quad \ell_x \neq \ell_y \quad , \tag{2.36}$$

while for a general vector operator

$$<x\vec{R}y> = 0 \quad \text{if} \quad \ell_x + \ell_y = 0 \quad \text{or} \quad |\ell_x - \ell_y| > 1 \quad . \tag{2.37}$$

However, we should have taken into account the parity operation

$$\Pi(\vec{r}) = -\vec{r} \quad .$$

The corresponding Π operator satisfies

$$\Pi^2 = I, \ \Pi\vec{R}\Pi = -\vec{R}, \ \Pi\vec{P}\Pi = -\vec{P} \tag{2.38}$$

but, since $\vec{L}$ is an axial vector

$$\Pi\vec{L}\Pi = \vec{L} \quad . \tag{2.39}$$

Eigenvectors of $\vec{L}^2$ have a well defined parity (the eigenvalue of Π). Looking at their realization by spherical harmonics, one finds

$$\Pi\vec{L}^2 = (-1)^{\ell}\vec{L}^2 \quad . \tag{2.40}$$

Thus, when x, y have a well defined angular momentum

$$<x\vec{R}y> = -<x,\Pi\vec{R}\Pi y> = -(-1)^{\ell_x+\ell_y}<x\vec{R}y>$$

so

$$<x\vec{R}y> = 0 \quad \text{if} \quad \ell_x + \ell_y = \text{even} \tag{2.41}$$

which is Laporte selection rule we spoke of in 2.1. The set of both equations (2.37) and (2.41) is equivalent to: no electric dipole transition: $<x\vec{R}y> = 0 \Leftrightarrow |\ell_x - \ell_y| \neq 1$.

---

dipole moment $\Sigma_i \vec{r}_i q_i = 0$; see work of Mie, Poincaré, Rayleigh, etc., on multipole expansion - it's applied group theory!) The trouble was: classically atoms should always radiate and use up their energy fast. Quadrupole radiation in atoms can be observed in exceptional cases (rare-earth, atoms in interstellar vacuum). In a radiation field, electromagnetic emission of photons can be induced and become intense: laser!

How Does This Theory Compare With Experiments? It is both very good and very poor. Within an accuracy of 10^{-4} the agreement for the values of the binding energy is perfect. The value predicted by the present theory of quantum electrodynamics gives a correction in $(Z\alpha)^4/n^3$ (i.e., a relative correction of $(Z\alpha)^2/n \sim 10^{-4}$) so that levels with different ℓ and same n have a small difference in energy.†

What is very bad is the counting of the number of levels. This can be seen by putting the hydrogen atom in a constant electromagnetic field ($\vec{F}$, electric and $\vec{B}$, magnetic). Then one must add to H

$$H_{em} = -\frac{3e}{2}\vec{K}\cdot\vec{F} + \frac{e}{2mc}\vec{L}\cdot\vec{B} \quad . \tag{2.42}$$

The effect of $\vec{F}$ (Stark effect) is well reproduced, but not that of $\vec{B}$. Indeed, levels of the same ℓ should split into $2\ell + 1$ levels separated by $eB/2mc$. They do split, but in an even number of levels!! This is due to the electron spin which we have not yet taken into account (see 2.6). One should also take into account the proton spin with effects

$$\sim \left\langle \frac{e^2}{m_e m_p} \frac{1}{R^3} \right\rangle \sim \alpha(Z\alpha)^3 m_e m_p^{-1}) \quad .$$

2.5. The Helium Atom

It has a nucleus of charge $Ze = 2e$ (mass $\sim 4m_p$) and 2 electrons. After separation of the center of mass motion, the Hamiltonian for the internal energy is

$$H = H_1 + H_2 + \frac{e^2}{R_{12}} \tag{2.43}$$

where $H_i = P_i^2/2m - Ze^2/R_i$, the hydrogen Hamiltonian and the operator R_{12} corresponds to $r_{12} = |\vec{r}_2 - \vec{r}_1|$ the relative distance of the two electrons. If we neglect the term in e^2/R_{12} (this is a better than 10% approximation) our problem is solved. We will consider only bound states. Let $\mathcal{H}^{(1)}$ be the Hilbert space of the bound states of hydrogen atom. Our simplified helium atom has Hilbert space $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(1)}$ with Hamiltonian $H_0 \otimes I + I \otimes H_0$ where H_0 is that of hydrogen. So the binding energy is $-(Z\alpha)^2(1/n_1^2 + 1/n_2^2)/2$ i.e., the sum of the binding energies for the two electrons.

We assume here that the term e^2/R_{12} is a perturbation in the technical sense (see Kato's book for mathematical rigor). This term breaks the SO(4) invariance, so the electron levels with different ℓ and same n no longer have the

† The difference between the two levels $n = 2$, $\ell = 1$ and 2 predicted by the theory of quantum electrodynamics is essentially $Z^4\alpha^5 2^{-3}$, i.e., $\sim 10^3$ megacycles and the agreement with experiment is of the order of 10^{-1} megacycles $\sim 10^{-15}\, mc^2/\hbar$. Quantum electrodynamics is not yet well defined for the mathematicians! Refined predictions for positronium, muonium, etc., are also very precisely verified.

same energy. (As we shall see later in 2.6, for a given n, E increases with ℓ.) What is left is angular momentum and parity conservation

$$[\vec{L},H] = 0, \quad [\Pi,H] = 0 \tag{2.44}$$

and the indistinguishability of the two electrons

$$[S_{12},H] = 0 \tag{2.45}$$

where S_{12} is the operator permuting the two electrons

$$S_{12}^2 = I, \quad S_{12}(A \otimes B)S_{12} = B \otimes A \in L(\mathcal{H}^{(1)} \otimes \mathcal{H}^{(1)}) \quad . \tag{2.45'}$$

The decomposition of the tensor product $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(1)}$ into the direct sum of symmetrical and antisymmetrical tensor spaces

$$\mathcal{H}^{(1)} \otimes \mathcal{H}^{(1)} = \mathcal{H}^{(1)} \vee \mathcal{H}^{(1)} \oplus \mathcal{H}^{(1)} \wedge \mathcal{H}^{(1)} \tag{2.46}$$

that we also wrote

$$\mathcal{H}^{(2)} = \mathcal{H}_{[2]} \oplus \mathcal{H}_{[1^2]} \tag{2.46'}$$

give the decomposition into eigenspaces of S_{12}. Let x, y be states of the hydrogen atom. Which of the two states $x \vee y = 1/\sqrt{2}(x \otimes y + y \otimes x)$ or $x \wedge y = 1/\sqrt{2}(x \otimes y - y \otimes x)$ yield the smallest expectation value for the positive operator e^2/R_{12}? It is obviously $x \wedge y$ because the two-electron wave function vanishes when e^2/R_{12} is very large (while that of $x \vee y$ has generally a maximum when $R_{12} = 0$). This symmetry character yields a new selection rule for the dipole radiation; the matrix element of the transition operator is proportional to

$$\langle \Psi | \vec{R}_1 + \vec{R}_2 | \Psi' \rangle \quad . \tag{2.47}$$

Since $\vec{R}_1 + \vec{R}_2$ is symmetrical, Ψ and Ψ' must have the same symmetry character $\varepsilon = \varepsilon'$ ($\varepsilon^2 = 1$) since

$$\langle \Psi, (\vec{R}_1 + \vec{R}_2)\Psi' \rangle = \langle \Psi, S_{12}(\vec{R}_1 + \vec{R}_2)S_{12}\Psi' \rangle = \varepsilon\varepsilon' \langle \Psi (\vec{R}_1 + \vec{R}_2)\Psi' \rangle \quad . \tag{2.47'}$$

As we announced in 2.1, this shows that the helium levels are to be divided in two sets according to their symmetry characters, and electric dipole transitions occur only within each set. Let me remind you that helium got its name because it was observed in the sun before being observed on earth. Its spectrum appears to be composed of two spectra, one for orthohelium ($\varepsilon = +1$), one for parahelium ($\varepsilon = -1$). This was a complete mystery before quantum mechanics. The explanation was given by Heisenberg in 1926, "Über die Spektra von Atomsystem mit zwei Elektronen", *Z. Physik*, 39, 499 (1926)). It also explained that the orthohelium has more levels; those of the type $x \otimes x$, as for instance the lowest level ($n = 1$, $\ell = 0$ for each electron). It is observed that corresponding rays (e.g., transitions $(1,0) \vee (n,\ell) \to (1,0) \vee (n',\ell')$ and $(1,0) \wedge (n,\ell) \to (1,0) \wedge (n',\ell')$ with $n' \neq 1, \ell' \neq 0$) of parahelium are about three times more intense than those of orthohelium. To explain it, the electron spin will have to be taken into account (see also 2.9).

2.6. Pauli Principle. The Electron Spin

We want to pass now to the case of n-electron atoms. The internal energy Hamiltonian is

$$H^{(n)} = \Sigma_{i=1}^{n} H_i + \sum_{1\leq i<j\leq n} \frac{e^2}{R_{ij}} \quad \text{with} \quad H_i = \frac{P_i^2}{2m} - \frac{Ze^2}{R_i} \quad . \tag{2.48}$$

Of course $H^{(n)}$ is invariant under the permutation group $S(n)$ of the n electrons. It is also the case for the electric dipole operator $e(\Sigma_i R_i)$ and for all observables. Identical particles cannot be distinguished from each other and every prediction of the theory must be invariant under $S(n)$.

When we consider states of Z (or n) distinguishable particles, we considered (with success for the helium atom) the Hilbert space tensor product of the $\mathcal{H}$ for each particle. Consider again $\mathcal{H}^{(n)} = \overset{n}{\otimes} \mathcal{H}^{(1)}$ for n identical particles. $S(n)$ acts on $\mathcal{H}^{(n)}$ by the representation $s \to S(s)$. Invariance under $S(n)$ of all observables requires that they are in the commutant $\{S(s)\}'$ of the set $\{S(s), s \in S(n)\}$ of operators. As we saw the rank one projectors wich represent physical state are observables of the theory so $\forall s$, $S(s)P_x S(s) = P_x$ for any vector x which represents a state. This requires that the vector $x\rangle$ belongs either to $\mathcal{H}_{[n]} = \overset{n}{\vee} \mathcal{H}^{(1)}$ (completely symmetrical) or to $\mathcal{H}_{[1^n]} = \overset{n}{\wedge} \mathcal{H}^{(1)}$ (completely antisymmetrical). The other spaces $\mathcal{H}_{[\ ]_\lambda}$ of the other factorial representations of $S(n)$ are excluded as space of physical states.

We have used both $\mathcal{H}_{[2]}$ and $\mathcal{H}_{[1^2]}$ for the helium atom. However, the use of $\mathcal{H}_{[n]} (n > 2)$ for atoms does not represent nature. Indeed, the ground state of any atom would have all electrons with the same binding energy (of the order of $(Z\alpha)^2/2$). Experimentally, only two electrons have this binding energy (X-ray spectrum for Z-large enough). The necessary energy (called ionization energy) for removing a first, a second, a kth ..., the Zth electrons of any neutral atom increase irregularly from a fraction of α^2 to $Z(\alpha)^2/2$. Moreover, as we shall see, vectors of some other $\mathcal{H}^{(n)}_{[\ ]_\lambda}$ do appear! The solution to this puzzle is that $\mathcal{H}^{(1)}$ is not the Hilbert space of the bound states of one electron in a constant potential. The electron has another degree of freedom, the spin and the Hilbert space of its states has to be changed into a new

$$\mathcal{H}^{(1)} = L^{(1)} \otimes K^{(1)}$$

where $L^{(1)}$ is the $L_2(R^3,t)$ previously called $\mathcal{H}^{(1)}$ and $K^{(1)}$ is a two-dimensional Hilbert space. Pauli was the first in 1924 to introduce the spin as an intrinsic angular momentum and magnetic moment for the nuclei, but it was Goudsmit and Uhlenbeck who introduced in 1925 the spin as an intrinsic angular momentum $\hbar/2$ for the electron. This explained the number of energy levels which appear in the Zeeman effect, but it did not explain the magnitude of their splitting. Indeed, the magnetic moment produced by an electric charge e moving with an angular momentum

$\vec{j}$ (Ampere's law!) is

$$\vec{\mu}/\frac{e}{2mc} = g\vec{j}/\hbar \tag{2.49}$$

where $e\hbar/2mc$ is the Bohr magneton. For the orbital momentum, $g(|\vec{j}/\hbar|$ integral) $g = 1$, but for the spin $|\vec{j}/\hbar| = 1/2$, g appeared to be 2. This was a mystery solved by Thomas in 1925. It is a relativistic effect.

It is an experimental fact that we have to use Fermi Statistics for electrons, i.e., the Hilbert space of electronic states of an n-electron atom is

$$\mathcal{H}^{(n)}_{[1^n]} = \overset{n}{\Lambda}\mathcal{H}^{(1)} \subset \oplus_\lambda (L^{(n)}_{[\ ]_\lambda} \otimes K^{(n)}_{[\ ]^c_\lambda}) \tag{2.50}$$

where $\mathcal{H}^{(1)}$ is the (new) one electron Hilbert space defined in (2.48).

Since $\dim K^{(1)} = 2$, the Young diagram of $[\ ]^c_\lambda$ has only two lines of length $\lambda_1 \geq \lambda_2 \geq 0$. Of course $\lambda_1 + \lambda_2 = n$; we will show that $\lambda_1 - \lambda_2$ is the chemical valence.

The diagram of $[\ ]_\lambda$ in $L^{(n)}_{[\ ]_\lambda}$ is the one symmetric through the diagonal. It has two columns $\lambda_1 \geq \lambda_2 \geq 0$. In other words, it has λ_2 lines of length 2 and $\lambda_1 - \lambda_2$ lines of length 1. That means that it cannot be completely symmetrical in more than two electrons, i.e., there can be only two electrons at most in each orbital state; then two electrons must have "different spin states", or more exactly, their spin-state has to be antisymmetrical. This is the Pauli principle, discovered by Pauli (*Z. Physik*, 31, 765 (1925)).

FIGURE 2.2. YOUNG DIAGRAM OF THE ORBITAL PART OF n-ELECTRON STATE: $L^{(n)}_{[\ ]_\lambda}$. HERE $n = 21$.

2.7. Atomic Shell Structure - Periodic Table

We can now clearly describe the lowest state of an n-electron atom. The orbital part is a vector of $\mathcal{H}^{(n)}$ obtained by filling all the lowest energy states, putting only two electrons in each electronic orbital state. Of course, in atoms

with $n > 1$ electrons,† the two specific properties of the $1/R$ potential disappear. The number of bounded states is finite, and the $SO(4)$ degeneracy no longer exists, i.e., states with the same n and different ℓ have different energy. The observed order of increasing energy for the states is given in Table 2.1.

TABLE 2.1. THE ELECTRON STATES ARE LISTED IN ORDER OF INCREASING ENERGY

n	1	2	2	3	3	4	3	4	5	4	5	6	5	4	6	7	...
ℓ	0	0	1	0	1	0	2	1	0	2	1	0	2	3	1	0	...
Spectroscopist Notation	1s	2s	2p	3s	3p	4s	3d	4p	5s	4d	5p	6s	5d	4f	6p	7s	...
$2(2\ell + 1)$	2	2	6	2	6	2	10	6	2	10	6	2	10	14	6	2	...
Total	2	4	10	12	18	20	30	36	38	48	54	56	66	80	86	88	...

Note that for a given n, the energy increases with ℓ. This is of course predicted by computation and can be qualitatively understood. Consider a nucleus of charge Ze and k electrons. At infinity an electron will feel the coulomb potential $(Z - k)e/r$, but when it gets near, the probability of finding some of the k electrons at the distance $r \sim n(Z\alpha)^{-1}\hbar(mc)^{-1}$ is not negligible, the $(k + 1)^{th}$ electron feels a potential above the Coulomb potential. It feels less the difference if it is in a $\ell = 0$ state (more concentrated inside the sphere of radius r) than in a $\ell = 1$ state, $\ell = 2$ state, etc., where the concentration of probability is more and more on the surface of the sphere. Note also from Table 2.1 that "p shells" ($\ell = 1$) are filled by the 10th, 18th, 36th, 54th, 86th electron and this is just the atomic number of the "rare gas" elements, very inactive chemically, Neon, Argon, Krypton, Xenon, and Radon. We can even give the Mendeleev's periodic table, in terms of electron states, if we know the order of increasing binding energy of the states (n,ℓ). Using the spectroscopic notation

$$\begin{matrix} \ell = 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ \quad s & p & d & f & g & h & i \end{matrix} \quad . \tag{2.51}$$

This order is, from Table 2.1, (n,ℓ): 1s, 2s, 2p, 3s, 3p, 4s $\sim$ 3d, 4p, 5s $\sim$ 4d, 5p, 6s, 5d $\sim$ 4f, etc. The sign $\sim$ indicates roughly the same energy so that the two shells are filled simultaneously. Indeed, for the element of $\ell = 2$ (d-shell) the + (++) sign indicates 1 (or 2) more electrons in the d-shell, taken from the s

† Because of the electrostatic repulsion among the electrons.

shell. We can now construct a periodic Table.

TABLE 2.2. PERIODIC TABLE (UP TO ELEMENT 56)

	ℓ=0	2	ℓ=1,2(2ℓ+1) = 6						ℓ=2,2(2ℓ+1) = 10										ℓ=3,2(2ℓ+1)=14
electron state	s		p						d										f
	1	2	1	2	3	4	5	6	1	2	3	4	5	6	7	8	9	10	
n=1	H 1	He 2																	
n=2	Li 3	Be 4	B 5	C 6	N 7	O 8	F 9	Ne 10											
n=3	Na 11	Mg 12	Al 13	Si 14	P 15	S 16	Cl 17	A 18	Sc 21	Ti 22	V 23	+ Cr 24	Mn 25	Fe 26	Co 27	Ni 28	+ Cu 29	Zn 30	
n=4	K 19	Ca 20	Ga 31	Ge 32	As 33	Se 34	Br 35	Kr 36	Y 39	Zr 40	+ Nb 41	+ Mo 42	+ Tc 43	+ Ru 44	+ Rh 45	++ Pd 46	+ Ag 47	Cd 48	14 rare earths 58 to 71
n=5	Rb 37	Sr 38	In 49	Sn 50	Sb 51	Te 52	I 53	Xe 54	La 57	Hf 72									
n=6	Cs 55	Ba 56																	

In d-shell, + (or ++) means 1 (or 2) more electrons in n-d (coming from the (n + 1),s) state).

The atomic state of an atom is labeled by the filled states, e.g., Oxygen: $(1s)^2(2s)^2(2p)^4$, i.e., 8 electrons. In general, the electrons fill up all states of lower energy and fill incompletely the last "shell"; for example, in the case of Oxygen, we can add 2 more electrons in the 2p state. The question which arises is which state is the atom ground state for an incomplete shell? We can label this state by a Young diagram (let us do it for the first elements).

TABLE 2.3. YOUNG DIAGRAM OF FUNDAMENTAL STATES

Z =	1	2	3	4	5	6	7
Name	H	He	Li	Be	B	C	N
Ydiag of orbital state $L^{(Z)}_{[\ ]_\lambda}$	1s	1s	1s, 2s	1s, 2s	1s, 2s, 2p	1s, 2s, 2p	1s, 2s, 2p
Ydiag of spin state $K^{(Z)}_{[\ ]^c_\lambda}$							

Now we can generalize what we say for helium. Given k electrons in the same energy state (i.e., $k \leq 2(2\ell + 1)$), the lowest energy state of this k electron configuration is the most antisymmetrical in the coordinates ($k = 2$, $k = 3$, ...), so it is the most symmetrical for the spin coordinates ($k = 2$, $k = 3$, etc.). This is illustrated in Table 2.3 for the 2p-electrons for $f = 1, 2, 3$. When $k = 4$, it is not possible to have a completely antisymmetric tensor on a $2\ell + 1 = 3$ dimensional Hilbert space, so we can give the successive atomic states of p shells.

TABLE 2.4. np-SHELL STATES

nb of electrons k =	1	2	3	4	5	6
filled shells						
space state						
filled shells						
spin state						
valence	1 (or 3)	2(or 4)	3	2	1	0
n = 2	B	C	N	O	F	Ne
n = 3	Al	Si	P	S	Cl	A
n = 4	Ga	Ge	As	Se	Br	Kr
n = 5	In	Sn	Sb	Te	I	Xe

The ionization energy (energy necessary to extract one electron from the p-shell) is increasing with k, as we expect, along a given p-shell, except for the four electron state, because it is the first one not completely antisymmetric.

Although the energy of ns-states is lower than that of the np-state, a low excited state of atoms with $k = 1$ or 2 is $k = 1$; $(ns)(np)^2$; $k = 2$; $(ns)(np)^3$, that is, an ns-electron goes to an np-state. This increases the valence by two and gives more bounded molecules.

There would be a lot more to say, even from group theory, about the Mendeleev Table. For instance did you notice that the ferromagnetic elements (Ni, Co, Fe and also Mn in alloys are together in an incomplete 3-d shell, etc.? But we shall stop here.

2.8. Atomic States in a Given Shell - Spin Orbit Coupling

The Pauli principle, expression of the Fermi statistics, simplifies the study of atoms. Instead of studying an n-electron system, within a good approximation we can (for not too highly excited states) consider k electrons outside a closed shell which has angular momentum zero, electric charge (n - k)e (with a distribution depending on the electron wave function). This closed shell can be considered as a spherical potential and the Hilbert space of states for k electrons in an $n - \ell$ shell is

$$\mathcal{H}^{(k)} = \overset{k}{\wedge} (\mathcal{H}_{2\ell+1} \otimes K_2) \tag{2.52}$$

to a good approximation, an electron state is a kth order decomposable antisymmetric tensor, $x \wedge y \wedge z \ldots$ (k factors).

Example. $\ell = 1$, $k = 2$, $\dim \mathcal{H}^{(1)} = \dim \mathcal{H}_3 \otimes K_2 = 6$; and $\dim \mathcal{H}^{(2)} = 15 = \binom{6}{2}$ for $k = 6$, $\dim \mathcal{H}^{(6)} = 1$ (complete shell again). Each decomposable tensor of $\mathcal{H}^{(k)}$ can be given a name or a label. That is what the spectroscopist does using a complete set of observable $\vec{L}^2$, $\vec{S}^2$, $\vec{J}^2 = (\vec{L} + \vec{S})^2$, J_z, that is the orbital angular momentum ℓ, the spin angular momentum s, the total angular momentum j, and its projection on axis j_3.

TABLE 2.5. THE $1S(np)^2$-STATES IN ORDER OF INCREASING ENERGY

Spectrocopist notation	$^3P_0$		$^3P_1$		$^3P_2$		$^1D_2$		$^1S_0$
L and space symmetry	1 		1 		1 		2 		0 
S and spin symmetry	1 		1 		1 		0 		0 
J = total ang. mom.	0		1		2		2		0
2J+1=nb of states	1	+	3	+	5	+	5	+	1 =15
$\langle \vec{L}.\vec{S} \rangle$	-2		-1		1		0		0

Remark on Table 2.5. Note that the space-antisymmetric  states (P-states) are below the  states as we already emphasized. For the symmetric states, the S-state which feels more the repulsion than the D-states, is above them. Why do the P-states appear in order of increasing J? This is the small spin-orbit effect that we can explain in the following way.

The orbital state of angular momentum $L\hbar$ produces a magnetic moment $(e\hbar/2mc)\vec{L}$, while the spin state of angular momentum $\vec{S}\hbar$ produces a magnetic moment $g(e\hbar/2mc)\vec{S}$ with $g = 2$ (see (2.49)). The interaction between the two magnetic moments has for matrix element†

$$\frac{e^2}{2}\left(\frac{\hbar}{mc}\right)^2 \left\langle \frac{\vec{L} \cdot \vec{S}}{R^3} \right\rangle \quad . \tag{2.53}$$

For a state $|j,\ell,s\rangle$ the expectation value of $\vec{L} \cdot \vec{S}$ is easy to compute from

$$\vec{J}^2 = (\vec{L} + \vec{S})^2 = \vec{L}^2 + 2\vec{L} \cdot \vec{S} + \vec{S}^2 \quad . \tag{2.54}$$

And the expectation value for state $|j,\ell,s,j_z\rangle$ (when j_z is the eigenvalue of J_z) is

$$\langle \vec{L} \cdot \vec{S} \rangle = \frac{1}{2}(j(j+1) - \ell(\ell+1) - s(s+1)) \tag{2.55}$$

where

$$|\ell - s| \leq j \leq \ell + s\ell,\ j + s \text{ integers} \geq 0 \quad . \tag{2.56}$$

This explains the value of $\langle \vec{L} \cdot \vec{S} \rangle$ in Table 2.5.

We have seen that for hydrogen the $\langle n\ 1/R^3\ n \rangle \sim (Z\alpha/n)^3$, so the expectation value of the spin orbit term is

$$\sim \langle L \cdot S \rangle \frac{(Z\alpha)^2}{2n^2} \frac{Z\alpha^2}{h} \sim \varepsilon_n \frac{\langle L \cdot S \rangle}{n} Z\alpha^2 \sim 10^{-4}\ \varepsilon_n \quad .$$

This is an order of magnitude. In the sodium atom (alcaline = hydrogen like) spectrum $[(1s)^2(2s)^2(2p)^6](3s)$ (fundamental state) and the $[\ ](3p)$ state has the largest splitting, i.e., nearly 10^{-3}, so the very bright $3p - 3s$ (yellow) Na-line is a doublet.

2.9. Spin and Euclidean or Galilean Invariance

In Chapter 2, Sections 2.7 and 2.8, we have mainly used spin as a new degree of freedom for the electron. This new degree of freedom has two discrete values (often called "up" and "down" in the physics manual) so the corresponding Hilbert space K_2, of complex valued functions defined in a 2-element set, has dimension 2. The atom Hamiltonian (2.48) is independent of this spin degree of freedom, i.e., it is of the form $H \otimes I$ acting on the space $\oplus_\lambda (L^{(n)}_{[\ ]_\lambda} \otimes K^{(n)}_{[\ ]^c_\lambda})$ of Equation (2.50). The permutations of the n electron spins are represented by operators of the form $I \otimes S(s)$ which commute with H. So eigenstates of H can have well defined $[\ ]^c_\lambda$. The simplest illustration is the helium atom $(n = 2)$.

† This is a short for $\vec{L} \otimes \vec{S}$ applied to $\mathcal{K}^{(k)}_{2\ell+1[\ ]_\lambda} \otimes K^{(k)}_{[\ ]^c_\lambda}$.

The set of states with $[\ ]^c_\lambda = [2] =$ ⊡ was called parahelium, with $[\ ]^c_\lambda = [1^2] =$ ⊟ was called orthohelium. Since $\dim K^{(2)}_{\square\square} = 3$ and $\dim K^{(2)}_{\square} = 1$, helium states (which are tensor products $x \wedge y$ or $x \vee y$ of different ($x \neq y$ hydrogen states) have the statistical weight (for instance in the Boltzman distribution of thermodynamic equilibrium) of 3 for parahelium and 1 for orthohelium; this explains that spectral lines of the former are three times more intense than the corresponding lines of the latter. Using the considerations of Chapter 1, Section 1.4, on the relations between the unitary groups and the permutation groups, we could also consider the action of the group $U(2)$ on the two-dimensional space $K^{(1)}_2$. Its action on $K^{(n)}_{[\ ]^c_\lambda}$ is through the factorial representation $[\ ]^c_\lambda$, and this action is on $\mathcal{H}^{(n)} \subset L^{(n)}_{[\ ]^c_\lambda} \otimes K^{(n)}_{[\ ]^c_\lambda}$ of the type $I \otimes (\oplus D_{[\ ]^c_\lambda}) = I \otimes (\oplus_s D_s)$; it commutes with the Hamiltonian $H \otimes I$.

These two points of view are formally equivalent for the classification of quantum states, but the $SU(2) \subset U(2)$ has a deeper meaning. It is related to the Euclidean or Galilean invariance. Let G be either group, and $\bar{G}$ its universal covering, i.e., there is a surjective homomorphism $\bar{G} \overset{\pi}{\to} G$. (As we have seen in Chapter 1, Section 1.2, and will see again in Chapter 4, it is an extension of the relativity group which acts through a linear representation on the Hilbert space of states. See also O'Raifeartaigh.) There is also a surjective homomorphism $\bar{G} \overset{\varphi}{\to} SU(2)$. (In the Euclidean case for instance $\bar{G} = R^3 \square SU(2)$; $\square$ = semi-direct product, where $SU(2)$ is the covering of the rotation group.) This gives us the action of $\bar{G}$ on $\mathcal{H}^{(1)}$; the one particle-state $\Psi(\vec{x},t;\sigma) \in L_2(\vec{x},t) \otimes K_\sigma$ transfoms into

$$(U(\bar{g})\Psi)(\vec{x}t,\sigma) = \Psi(\pi(\bar{g})^{-1} \cdot (x,t);\ \varphi(\bar{g})^{-1} \cdot \sigma) \quad . \tag{2.57}$$

Often physicists prefer to write equivalently $\mathcal{H}^{(1)}$ as the Hilbert space of square integrable function Ψ_σ of $\vec{x},t$ with value in the two-dimensional Hilbert space K_σ. Then (2.56) reads

$$(U(\bar{g})\Psi_\sigma)(\vec{x},t) = \sum_{\sigma'=1,2} D_{1/2}(\varphi(\bar{g}))_{\sigma\sigma'}\Psi_{\sigma'}(\pi(\bar{g})^{-1} \cdot (\vec{x},t)) \quad . \tag{2.58}$$

To summarize, the spin is related to (essentially the rotation part of) Euclidean (and a portion of the larger Galilean) invariance; and it is an intrinsic angular momentum for the electron. We will study it in 4.5. The value $g = 2$ for the corresponding electron intrinsic magnetic moment is, however, a relativistic effect (see Figure 2.5).

Conservation of angular momentum implies only that $\vec{J} = \vec{L} + \vec{S}$ (orbital + spin angular momentum) be a constant of motion. In atoms $\vec{L}$ and $\vec{S}$ are separately conserved to a good approximation only because H is spin-independent (see Equation (2.48).

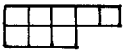
2.10. Molecules

The interaction which binds N atomic nuclei and n electrons into a neutral molecule (or a charged molecular ion) is essentially the electrostatic (= coulomb) interaction. Instead of treating directly a $N + n$ body problem, one uses the Born-Oppenheimer approximation where the (heavy) nuclei are considered fixed. Take for example the Hamiltonian of the hydrogen molecule (subscript $A = I$, II for the two protons, $i = 1, 2$ for the two electrons, $r_{A,i} = |\vec{r}_A - \vec{r}_i|$, etc.).

$$H = \frac{1}{2M}(\vec{p}_I^{\,2} + \vec{p}_{II}^{\,2}) + \frac{1}{2M}(\vec{p}_1^{\,2} + \vec{p}_2^{\,2}) - e^2\Big(\sum_{\substack{A=I,II \\ i=1,2}} \frac{1}{r_{A,i}}\Big) + \frac{e^2}{d} + \frac{e^2}{r_{12}} \tag{2.59}$$

where $d = |\vec{r}_I - \vec{r}_{II}|$ the distance between the two hydrogen nuclei is considered as a parameter in the Born-Oppenheimer approximation (and the kinetic energy of the nuclei will be neglected). When d is very large, a state of (2.59) is in $\mathcal{H}^{(2)}$, the tensor product of two hydrogen atom-Hilbert spaces. Consider first the space dependence $L_2^{(2)}(\vec{r},r)$ and the two-dimensional subspace $h = (x \otimes y) \oplus (y \otimes x)$ where x, y are hydrogen states. These two states have the same energy $\varepsilon = \varepsilon_x + \varepsilon_y$. However, in this basis, for h, the Hamiltonian operator $H|_h$ is not exactly diagonal when d is finite because each electron feels also the attraction of the other nucleus, so

$$H|_h = \begin{pmatrix} \varepsilon & \rho \\ \rho & \varepsilon \end{pmatrix} \tag{2.60}$$

(since it is Hermitian), and its eigenvalues are $\varepsilon \pm |\rho|$. Hence, the two eigenstates of $H|_h$ are $1/2(x \wedge y)$ and $1/2(x \vee y)$ and they have an energy difference of $2|\rho|$. When $d \to \infty$, $|\rho| \to 0$ and so does $e^2/d - |\rho|$. When $d \to 0$, $e^2/d - |\rho| \to \infty$. But there is a domain for d for which $e^2/d - |\rho| < 0$, and a value of d for which $e^2/d - |\rho|$ is minimum. The ground state is of the type $x \otimes x$, and from Fermi statistics the two-electron spins form an antisymmetrical state. Hydrogen (or alcaline) form a similar type of liaison (covalent bond) with electrons of unfilled shell of atoms. The number of atoms which can be bound to an atomic (spin) state $\begin{matrix}\lambda_1 \\ \lambda_2\end{matrix}$  is $\lambda_1 - \lambda_2$ in order to form a closed "spin" shell, as was discovered empirically before 1920, and $\lambda_1 - \lambda_2$ is the "valence" of the atom. Quantum mechanics has explained qualitatively and quantitatively the covalent bond (W. Heitler and F. London, *Z. Physik*, 44, 455 (1927)). It explains, for instance, why the molecules H_2, H_2S, H_2Se are of the form $H-\!\!-O\!\!\diagdown_{H}$ with an angle $\geq 90°$ (the repulsion of the two H atoms makes the angle increasing from 90° for H_2Se, H_2S, H_2O (= 108°). It explains why NH_3 is a trihedron and CH_4 a tetrahedron, why C_2H_4 is flat $\begin{matrix} H \diagdown \\ H \end{matrix} C = C \begin{matrix} \diagup H \\ \diagup H \end{matrix}$ (σ and π electrons). It explains mesomery (e.g., for benzene), etc. Group theory is so useful for explaining molecular spectra! We

have to skip this subject for now and simply refer to an elementary but elegant textbook, *Quantum Chemistry*, by Eyring, Walter, and Kimball, Wiley, New York (1944).

The symmetry group of a molecule is a subgroup of $O(3)$, the three-dimensional orthogonal group. When its shape is known experimentally, its symmetry group G is known. Let us refer to Wigner's paper (Göttingen, (March 1930), p. 133) on the characteristic elastic vibration modes of molecule (given by the equivalent classes of G), as examples of the application of group theory. Wigner studied CH_4 (whose group G is $S(4)$) as an illustration. The H. A. Jahn, E. Teller theorem (*Proc. Roy. Soc. Ser. A*, 161, 220 (1932)) proves that the electron orbital state of "non-straight" molecules cannot transform as an irrep of G of dimension > 1. (The irrep has dimension 2 for molecules whose atoms are on a straight line.)

We will study here only one very important example.

2.11. Measurement of Spin and Statistics of Nuclei by the Study of Diatomic-Molecule Spectra

The Hamiltonian H of a diatomic molecule can be divided into

$$H = H_{\text{electronic}} + H_{\text{vibration}} + H_{\text{rotation}} + H'$$

where, to a good approximation, H' can be neglected. $H_{\text{electronic}}$ gives the electronic states of the molecule; each such state yields a distance d (between the two nuclei) which minimizes the energy. The invariance group is $O(2)$ or if the two nuclei are identical, $O(2) \times Z_2$. Binding energy for such states are typically a fraction of α^2 (few electron volts). H vibration is essentially the harmonic oscillator Hamiltonian for small oscillations around the equilibrium position fixed by the distance d. The equidistant spacing of the vibration level is small compared to α^2, and the H rotation yields for each d also rotation energies proportional to $\ell(\ell + 1)$, ℓ integer ≥ 0, and small compared to the vibrational energies (rotational bands; in spectrum). If the two nuclei of the molecule are identical, which is the symmetry of the molecular state for the permutation group $S(2)$ of these two nuclei? The symmetry depends only on the spin state of the nuclei, (each of spin j) its $SU(2)$ irrep is

$$D_s,\ 0 \leq s \leq 2j;\ \square\square\ s = 2j,\ 2j-2,\ 2j-4,\ \ldots$$
$$\begin{array}{|c|}\hline\ \\ \hline\ \\ \hline\end{array}\ s = 2j-1,\ 2j-3,\ \ldots$$

and the rotational state of the system, $\square\square$ for ℓ even, $\begin{array}{|c|}\hline\ \\ \hline\ \\ \hline\end{array}$ for ℓ odd.

Since H is independent of the nuclear spin (to a very good approximation) the symmetry character of the nuclear spin state is a constant of motion (with often a lifetime of weeks) and is, as for Helium, called ortho or para. Because of "statistics", the symmetry character of the rotational state is also a constant of motion. So the rotational spectrum of the molecule divides into two independent

sets of transitions - those between even ℓ, and those between odd ℓ. The transitions occur in both states as quadrupolar $\ell + 2 \to \ell$, with a (radio-wave) photon energy $\sim (\ell + 2)(\ell + 3) - \ell(\ell + 1) = 4\ell + 6$. The number of $\square\square$ nuclear spin states is $(2j + 1)(2j + 2)/2 = (j + 1)(2j + 1)$. The number of $\begin{smallmatrix}\square\\\square\end{smallmatrix}$ nuclear spin states is $(2j + 1)(2j)/2 = j(2j + 1)$. So if for the molecule the relative intensity of spectral (rotational) lines is $(2j$ integer $\geq 0)$, $j/j+1$ for ℓ even/ℓ odd, the nuclear spin is j, the statistics of the nuclei is $\left(\begin{smallmatrix}\square\\\square\end{smallmatrix} \times \square\square\right) / \left(\square\square \times \begin{smallmatrix}\square\\\square\end{smallmatrix}\right) = \begin{smallmatrix}\square\\\square\end{smallmatrix}$ Fermi; if it is $j/j + 1$ for ℓ odd/ℓ even, the nuclear spin is j, the statistics of the nuclei is $\left(\begin{smallmatrix}\square\\\square\end{smallmatrix} \times \begin{smallmatrix}\square\\\square\end{smallmatrix}\right) / \left(\square\square \times \square\square\right) = \square\square$ = Bose. Experimentally, only Fermi statistics is found for half odd integral j (as for the electron) and Bose statistics for integral j. We will summarize this important experimental fact by

$$\text{statistics} = (-1)^{2j} \quad . \tag{2.61}$$

For instance when only even ℓ rotational states exist, we conclude that $j = 0$, and the statistics has to be Bose.

Historically, the first nuclear spin measured (F. Rasetti, *Z. Physik*, 61, 598 (1930)) was (in 1929) that of N_{14} (nitrogen molecule N-N). Rasetti found $j = 1$ and Bose statistics. But it was then believed that the universe was made of protons p^+, electrons e^-, and photons γ, (the only particles then known, and that the nucleus N_{14} of charge 7e, contained 14 protons and 7 electrons, thus, half integral spin and Fermi statistics were expected. This measurement started a crisis in physics.

Appendix. On Commutation Relations

Professor Bargmann pointed out to me that I have spoken of the invariance group of the commutation relations only in the context of rotational invariance (see Equation (2.10)). Surely it is worth mentioning the general case: consider the relations

$$[P_i, Q_j] = i\hbar\delta_{ij}\mathbb{1} \tag{2.61}$$

$(i,j = 1$ to $n)$. Let $a = (a_1 \ldots a_n)$, $b = (b_1 \ldots b_n) \in R^n$; we can use the tensor operator notations $P(a) = \Sigma_i a_i P_i$, $Q(b) = \Sigma_j b_j Q_j$. Equation (2.61) defines a $2n + 1$ dimensional Lie algebra which is a central non-abelian extension $\mathcal{g}$ of R^{2n} by R^1 (center of $\mathcal{g}$). This extension is defined by the antisymmetrical bilinear form on $R^{2n} = R^n \oplus R^n$

$$\sigma(a \oplus b, a' \oplus b') = a \cdot b' - b \cdot a' \tag{2.62}$$

where $a \cdot b = \Sigma_i a_i b_i$. The symplectic group $Sp(n)$ which leaves this form invariant is a group of automorphism of $\mathcal{g}$.

The corresponding simply connected group G has, up to an equivalence, a unique unitary irrep (von Neumann's theorem). Its Schrödinger realization as operators on the space L^2 of functions of n variables: $x = (x_1 \ldots x_n)$ is $U_a = e^{iP(a)}$ with $(U_a f)(x) = f(x + a)$; $V_b = e^{iQ(b)}$ with $(V_b f)(x) = e^{i\hbar b \cdot x} f(x)$. Here x, $a \in E_n$, $b \in E_n'$ dual of E_n. In the case of Equation (2.10) $n = 3$. Furthermore, the rotation group $SO(3)$ leaves invariant the symmetrical linear form β on E_3 and we used the corresponding identification of E_3 and its dual.

3. NUCLEAR PHYSICS: STRONG AND WEAK INTERACTIONS

3.1. The Set of Known Nuclei

The nuclei are made of protons p and neutrons n. These two particles have similar masses $m_p = 1836.10\, m_e = 938.25$ MeV, $m_n = 939.55$ MeV. The proton has electric charge + e. Both have spin 1/2. We define a nucleus by its number Z of protons and N of neutrons, and denote it by (Z,N); it contains A = Z + N nucleons. Nuclei have bound excited states, which are unstable. The ground state itself may be unstable and the nucleus may transform spontaneously into another nucleus by one of the following types of decay.

a) β^--decay $n \to p^+ + e^- + \bar{\nu}$ ($\bar{\nu}$ = antineutrino); $(Z,N) \to (Z+1, N-1) + e^- + \bar{\nu}$
β^+-decay $(Z,N) \to (Z-1, N+1) + e^+ + \nu$ which competes with e^--capture $(Z,N) + e^- \to (Z-1, N+1) + \nu$ (which requires less energy).
The mean life τ can vary from 10^{-3} sec to 10^{20} years.

b) α-decay†: $(Z,N) \to (Z-2, N-2) + (2,2)$ for $A > 140$ nuclei, τ from seconds to 10^{20} years.
And two much rarer types:

c) neutron emission: $(Z,N) \to (Z, N-1) + n$ rare, $\tau <$ few seconds,

d) spontaneous fission into two smaller nuclei $(Z,N) \to (Z_1,N_1) + (Z_2,N_2)$.

Let us call nuclei stable if they have a half life of decay $\tau > 10^{20}$ years. 274 stable nuclei are known.

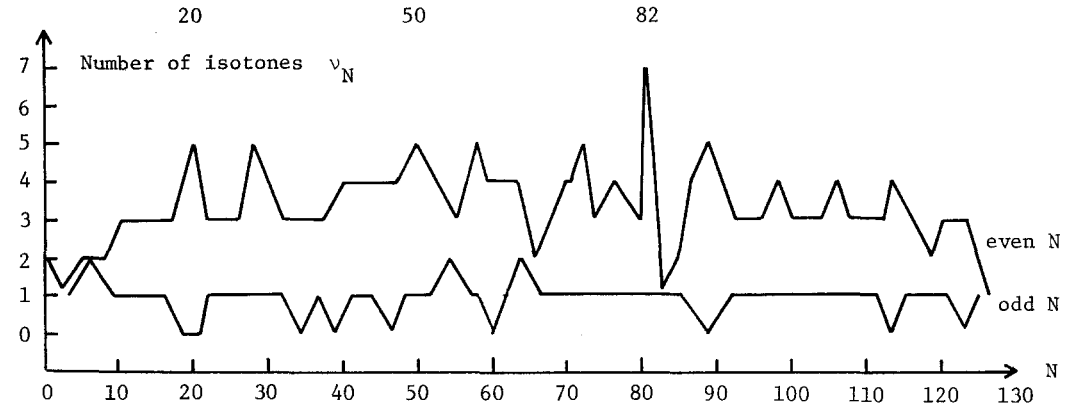
A even	Z even	N even	165		A odd	Z even	N odd	55		
	Z odd	N odd	4	(Z = N = 1,3,5,7)		Z odd	N even	50		

The much greater abundance of Z even, N even nuclei is strikingly illustrated in Figure 3.1 which gives the number of stable nuclei for given Z (isotopes) and for a given N (isotones).

It is worthwhile to note from Figure 3.1 that nuclei for Z = 20, Z = 50, (N = 82) have definitely more isotopes (or isotones) than their even-neighbors. This is also true, but less strikingly, for N = 20, N = 50 (and also N = 28). The heaviest stable nucleus is Pb_{208}, Z = 82, N = 126. Another striking feature in the distribution of stable nuclei in function of Z and N is that with two exceptions $N - Z \geq 0$ and N - Z is a slowly increasing function of A = N + Z:

N - Z = -1 for the proton (Z = 1) and He^3 (Z = 2)
N - Z = 0 for 13 nuclei;
N - Z = 1 for 16 nuclei;
N - Z increases with A on the average $(N - Z) \sim 6.10^{-3}\, A^{5/3}$.

† What was first called an α-particle has been identified with a Helium nucleus: (2,2).



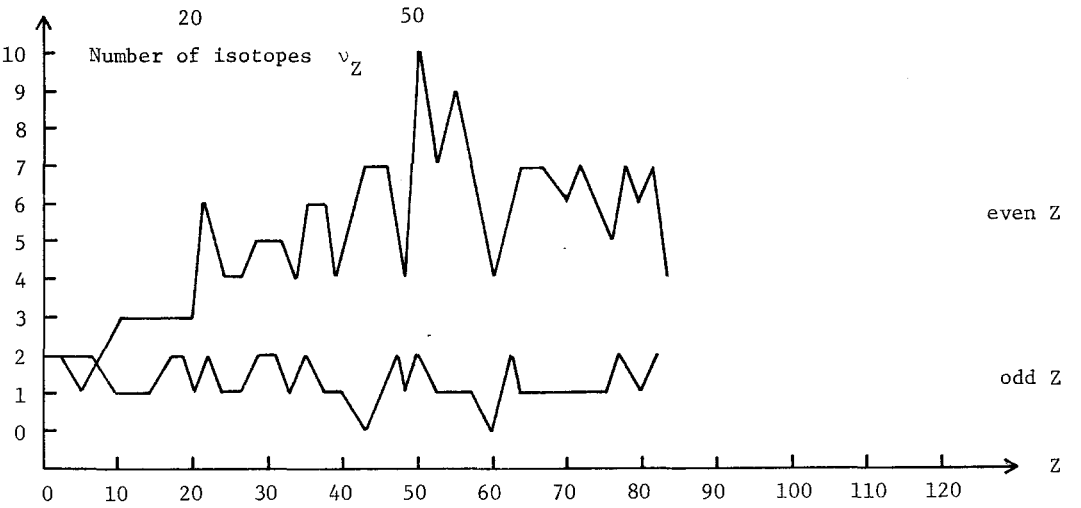


FIGURE 3.1. NUMBER ν_N AND ν_Z OF STABLE NUCLEI (Z,N) AS FUNCTIONS OF Z AND N

Note that there are no stable nuclei for Z = 43, 61, > 83, for N = 19, 21, 35, 39, 45, 61, 89, 115, 123, > 126 and none for A = N + Z = 5, 8, 147, ≥ 208. The heaviest stable nucleus is $^{126}_{82}Pb_{208}$ with Z = 82, N = 126. The most striking feature of Figure 3.1 is that ν_Z and ν_N are mainly 1, sometimes 2 or 0 for odd Z or odd N. Their value is more irregular for even Z or even N, there are relative maxima for Z = 20 = N, Z = 50 = N, N = 82 and also N = 28.

Nuclear forces are the most intense in nature, nowever, they do not bind more than 208 nucleons together†. The reason is that they have short range and also that nucleons obey Fermi statistics which, for condensed states, yield effects similar to repulsion.

More than one thousand different unstable nuclei are known. Those with a lifetime $\tau > .1$ (age of universe) and their decay products are found in nature, some are still produced in nature (C^{14}). All other are man made; more than half of those have Z-even, N-even. There exists a semi-empirical formula (Weizsäcker's) giving the binding energy of the lowest state of (stable or unstable) nuclei in function of Z, N and $A = Z + N$

$$B(Z,N) = Zm_p + Nm_n - m(Z,N) = U_\nu A - U_c Z(Z-1)A^{-1/3}$$

$$- U_s A^{2/3} - U_t \frac{(Z-N)^2}{A} + U_p \frac{(-1)^Z + (-1)^N}{2} A^{-3/4} \qquad (3.1)$$

the values of the constants U are in MeV

$$U_\nu = 14.0 \text{ MeV}, \; U_c = .61 \text{ MeV}, \; U_s = 14.0 \text{ MeV}, \; U_t = 84.2 \text{ MeV}, \; U_p = 34 \text{ MeV}$$

U_ν corresponds to maximal average binding energy by nucleon. The term U_c corresponds to the Coulomb repulsion among Z protons equally distributed in a sphere of radius proportional to $A^{1/3}$. The term U_s corresponds to a surface effect which suggests a short range for nuclear forces; U_t favors a minimum for $|Z - N|$ while U_p corresponds to pairing effects in like nucleons. As we saw, nuclei with even Z and N are more stable and more numerous than those with odd Z and/or odd N. A rule without exception is that all known Z even, N even nuclei have zero spin (= angular momentum at rest).

The distribution of nuclear spin for odd A nuclei is discussed in Section 3.4.

3.2. Isospin

As soon as the neutron was discovered (1932), Heisenberg created a formal language for the study of nuclei. Neutrons and protons are considered as the same particles, the nucleons, which have five degrees of freedom: 3 continuous in space $(\vec{x})$, a two valued one, σ, for the spin and a new one that Heisenberg simply called the fifth degree of freedom, τ, and which distinguishes neutrons and protons

† The existence of neutron stars with a radius of 10 km to 100 km and containing $\sim 10^{57}$ neutrons has been postulated. These stars seem to be observed now as "pulsars". They are indeed gigantic nuclei, but the binding energy is due both to nuclear and gravitational forces.

(*Z. Phys.*, 77, 1 (1932)); since, like the spin, it is two-valued, it is now called isospin.†

This Heisenberg convention has revealed itself more than useful. Indeed it was quickly established that nuclear forces did not distinguish between protons and neutrons: their differences (different electric charge and magnetic moment, small mass difference) are attributed mainly to electromagnetic effects and it is a reasonable approximation to neglect them.

If we denote the Hilbert space of our nucleon states by

$$\mathcal{H}^{(1)} = L_2(\vec{x},t) \otimes K_\sigma \otimes K_\tau \tag{3.2}$$

that of a number A of nucleons is

$$\mathcal{H}^{(A)} = \mathcal{H}^{(1)}_{[1^A]} = P_{[1^A]} \otimes \sum_\lambda ((L_2 \otimes K_\sigma)^{(A)}_{[\lambda]} \otimes K^{(A)}_{\tau[\lambda]^c}) \tag{3.3}$$

where $P_{[1^A]}$ is the projector on $\mathcal{H}^{(1)}_{[1^A]}$. A convenient approximation for the study of a nucleus of A nucleons is to replace the sum of 2-particle interactions,†† by an average potential (= sum of 1-particle Hamiltonians) plus a residual 2-particle potential, which is still attractive. Then the analogy with the study of atoms††† allows us to draw qualitative conclusion. Using the same type of argument as in Section 2.6 for atoms, but here with the opposite sign, we know that for the ground state $[\lambda]$ in Equation (3.3) should be as symmetrical as possible, so $[\lambda]^c$ is as antisymmetrical as it can be with the restriction that it has only two lines. This implies that the two lines are as nearly equal as possible

$$\lambda_i \geq 0;\ [\lambda]^c = [\lambda_1,\lambda_2] 0 \leq \lambda_1 - \lambda_2 = \lambda \text{ minimal};\ \lambda_1 + \lambda_2 = A \ . \tag{3.4}$$

If the nucleus has Z protons and N neutrons $(Z + N = A)$ its states are completely symmetrical in $\sup\{Z,N\}$ particles, so

$$\lambda_1 \geq \sup\{Z,N\} \tag{3.5}$$

and

$$|Z - N| \leq \lambda_1 - \lambda_2 \ . \tag{3.6}$$

---

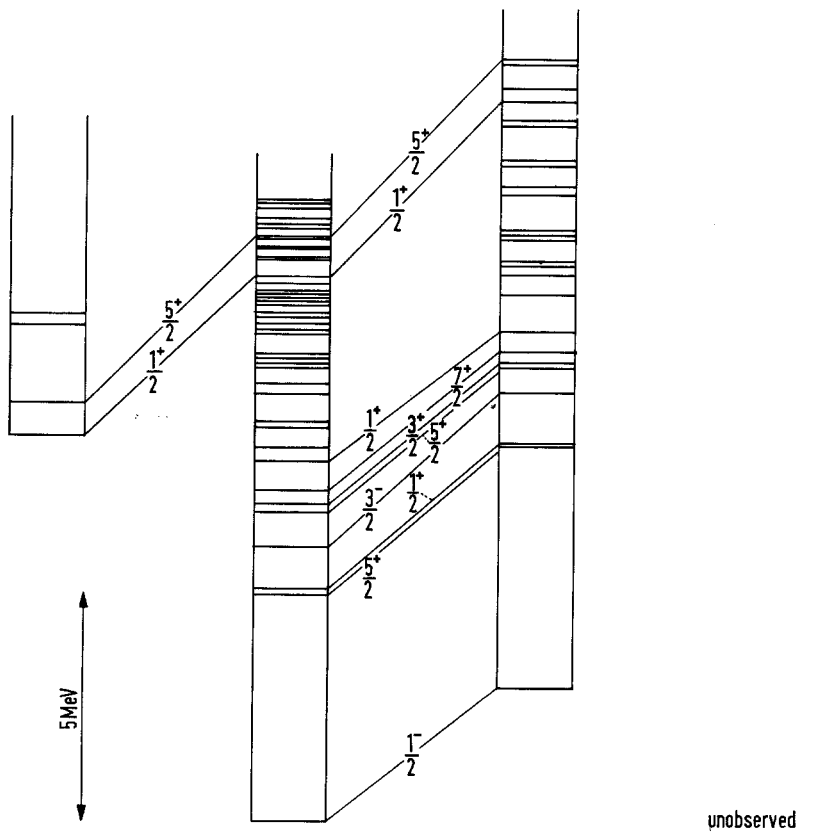
† Called isotopic spin since 1936, the name isobaric spin would have been more proper. Anyway it has been shortened into isospin by the natural evolution of language.

†† In fact physicists are more sophisticated: when a sum of 2-particle interaction does not yield a good enough approximation, one adds also the sum of all k-particle $(2 < k \leq A)$ interactions, mainly for $k = A$ (collective effects).

††† There is still a difference. Atoms of n electrons consist of $n + 1$ particles and as we have seen, the elimination of the center of mass motion is easy: one singles out the nucleus, and the electrons are all treated on the same footing. This elimination is still clumsily carried out in nuclear physics.

So (3.4) can be translated into: the most stable nuclei have as small $|Z - N|$ as possible. As we have seen, this is well verified for light nuclei, where the electromagnetic repulsion of protons in negligible; when this repulsion is taken into account $0 < N - Z$ has to be a slowly increasing function of $A = N + Z$.

In the same approximation in which n, p are considered identical, isobars (nuclei with the same number $A = Z + N$ of nucleons) should be identical. Consider Figure 3.2; it gives the energy spectrum of the known states for $A = 15$,



C^{15}	N^{15}	O^{15}	F^{15}
Z = 6	7	8	9
N = 9	8	7	6

j^p is the spin (j, 1/2 integer > 0) and the parity (±) of the state.

FIGURE 3.2. SPECTRUM OF STATES OF ISOBARS 15

States of C^{15} have isospin ≥ 3/2. Another conventional notation for nuclei is to use the chemical symbol of the corresponding atom (this gives implicitly the number Z) and write the number of nucleons $A = Z + N$ in superscript.

and the known value of spin and parity of these states. The similarity of the spectra (at least for the low lying part) for $1/2|Z - N| = 1/2$ (i.e. N^{15} and O^{15} nuclei) is striking. The essential difference is a shift upward in energy of $\sim$ 3 MeV for O^{15} which has one more proton than N^{15}. The pairs of corresponding states are called "doublets" of isospin 1/2 states.

Let us explain in detail this point of view, which exploits the relations between the permutation groups $S(n)$ and the unitary groups $U(k)$ that we have recalled in 1.4 and used in 2.9.

Nuclear interaction does not distinguish between protons and neutrons. For a nucleus this implies a property of invariance under the permutations ($\in S(A)$) of its nucleons. We could also have translated this property as follows:

All nuclear physics observables O acting on $\mathcal{H}^{(1)}$ (space of the one particle states for the nucleon) in Equation (3.3) are of the form (3.7), where

$$\mathcal{H}^{(1)} = L_2(\vec{x},t) \otimes K_\sigma \otimes K_\tau \tag{3.2}$$

$$O = N \otimes I \tag{3.7}$$

$$U(2) = I \otimes U(2) \tag{3.8}$$

i.e., they correspond to a trivial action on K_τ, the factor in this tensor product which corresponds to Heisenberg's fifth degree of freedom "proton - neutron".

The action of the group $U(2)$ on $\mathcal{H}^{(1)}$, as defined by (3.8), commutes with every observable: $U(2) \subset \{O\}'$, the commutant of the algebra of "one particle observables". The action of this $U(2)$ can be extended to every $\mathcal{H}^{(A)}$, ($A \geq 0$), Hilbert space of the A particle states. Therefore, in nuclear physics, when the non-nuclear interactions are neglected, this $U(2)$ is a subgroup of the invariance group. $\mathcal{H}^{(A)}$ has the same decomposition into spaces of factorial representation for $S(A)$ and $U(2)$ and we use the same symbols (Young diagrams) for the corresponding representations.

Since Coulomb repulsion of the protons can be neglected only in light nuclei, it was not expected that isospin conservation could be an interesting concept for heavier nuclei. However, the progress of nuclear physics in the last five years has shown that for nuclei with A up to 100, isospin is indeed a useful concept. For a non-technical review of this question, see W. R. Coker and C. F. Moore, "Isobaric Analog Resonances", *Physics Today*, 22, no. 4, 53 (1969).

3.3. U(4) Invariance

In 1936 Wigner, in his paper "On the consequence of the Symmetry of the nuclear Hamiltonian on the Spectroscopy of Nuclei", *Phys. Rev.*, 51-106 (1937)†

† Reproduced in Dyson's anthology: *Symmetry Groups in Nuclear and Particle Physics*, Benjamin, New York (1966).

studied the approximation where not only isospin dependence of nuclear forces is neglected but also the spin dependence. Then Equation (3.7) and (3.8) can be replaced by

$$\mathcal{H}^{(1)} = L_2(x) \otimes K_\sigma \otimes K_\tau \tag{3.2}$$

$$\mathcal{O} = N \otimes I \otimes I \tag{3.9}$$

$$U(4) = I \otimes U(4) \quad . \tag{3.10}$$

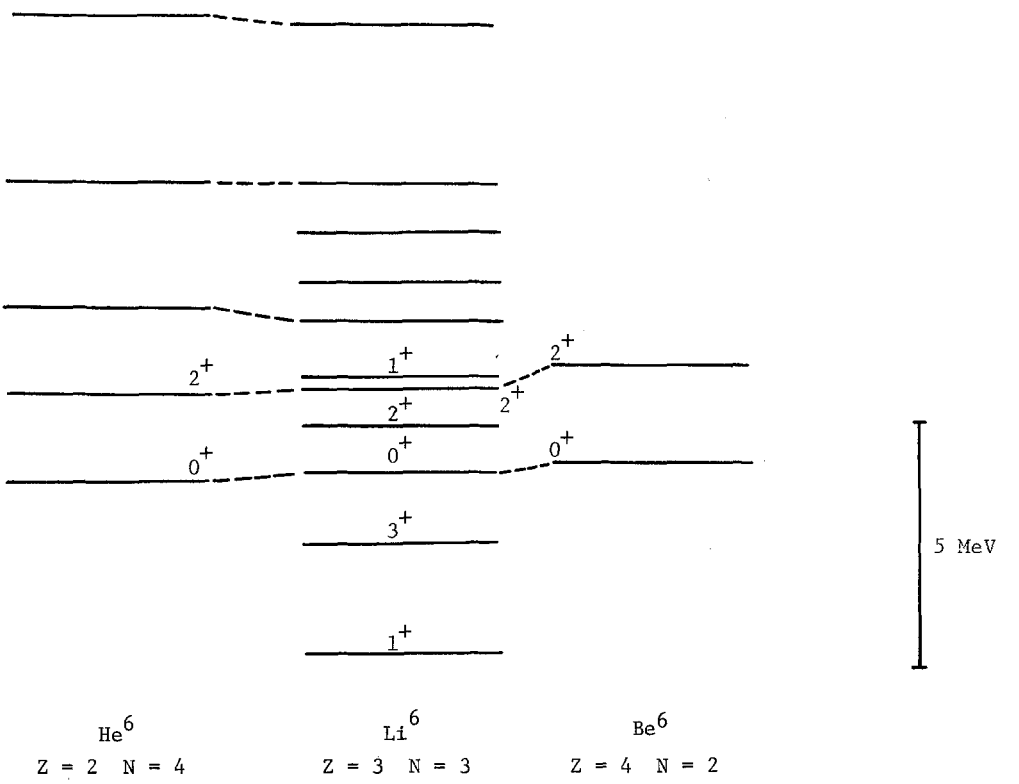
In this approximation, nuclear theory is also invariant under the group $U(4)$ acting on the four dimensional space $K = K_\sigma \otimes K_\tau$ and Equation (3.3), for the Hilbert space of A nucleon states can be replaced by

$$\mathcal{H}^{(A)} = \mathcal{H}^{(1)}_{[1^A]} = P_{[1^A]} \oplus_\lambda (L_{2[\lambda]} \otimes K_{[\lambda]^c}) \tag{3.11}$$

where the $[\lambda]$ are representations of $U(4)$.

For the most stable states, the property (used in 3.2) of the "residual" two-nucleon force to be attractive implies now that $[\lambda]$ is as symmetrical as possible, so $[\lambda]^c$ is as antisymmetrical as possible, i.e., its Young diagram has its four lines of length $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 \geq 0$ (with $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = A$) as nearly equal as possible. For $A/4$ = integer this implies $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = A/4$. This $U(4)$ irrep has dim. 1. The restriction of this representation of U(4) (acting on $K_\sigma \otimes K_\tau$) to the subgroup $SU(2) \times SU(2)$, yields a spin 0 and isospin 0 for the ground state. As we have seen, the former result is observed for all such nuclei, the latter only for light nuclei ($Z < 17$) where Coulomb repulsion of protons is not too large. For nuclei with $A = 4n + 2$, the $[\lambda]^c$ representation of lowest lying states is $\lambda_1 - 1 = \lambda_2 - 1 = \lambda_3 = \lambda_4 = n$; it has dimension $\binom{4}{2} = 6$. Its restriction to the subgroup $SU(2) \times SU(2)$ decomposes into the direct sum of two three-dimensional representations: one of spin 1, isospin 0, the other of spin 0, isospin 1. In Figure 3.3 (for which $n = 1$) this gives correctly the spin of the lowest state of Li^6 (spin 1) and He^6 and Be^6 (spin = 0). These last two levels form an isospin triplet with the third level (spin 0^+) of Li^6. The other levels whose spin are marked in Figure 3.3 belong to another equivalent representation of $U(4)$ with an angular orbital momentum (i.e., angular momentum of the space degree of freedom) $\ell = 2$. So the total angular momentum has the possible value $j = \ell = 2$ for the spin 0, isospin 1 states and $\ell - s \leq j \leq \ell + s$ i.e., $= j = 3, 2, 1$ for the spin 1, isospin 0 states i.e., those of Li^6 with no correspondents in He^6 and Be^6.

States belonging to a U(4) irrep are called supermultiplets in physics literature. The study of Galilean invariance of the theory of supermultiplets is very similar to that made in 2.9 for atomic physics.



j^p is the spin (j integer ≥ 0) and the parity $p(\pm)$ of the state.

FIGURE 3.3. SPECTRUM OF LEVELS OF THE NUCLEI WITH 6 NUCLEONS

One has to consider the covering $\overline{G}$ of the Galilean group G, $\overline{G} \overset{\pi}{\to} G$ and also the homomorphism

$$\overline{G} \overset{\phi}{\to} SU(2) \times 1 \subset U(2) \times U(2) \subset U(4)$$

The invariance group of the theory is the direct product $G \times U(4)$ and $\overline{G}$ is the subgroup $\overline{G} \overset{i}{\to} G \times U(4)$ with $i(\overline{g}) = (\pi(\overline{g}), \phi(\overline{g}))$.

The approximation of spin and isospin independence for nuclear forces leading to $U(4)$ invariance is crude and could not be expected to be very useful for nuclei with a number A of nucleons not very small. However, as for the better approximation of isospin conservation, $U(4)$ invariance has been usefully applied to nuclei with A up to 100 as shown by the statistical study of the energy of their ground state: P. Franzini and L. A. Radicati, "On the validity of the supermultiplet model", *Phys. Lett.*, 6, 322 (1963). (Reproduced in Dyson's anthology, quoted in 3.3 and 4.)

3.4. Shell Model

We have seen that nuclei with Z or $N = 8, 20, 50, 82$, and $N = 126$ seem more stable. Many properties of nuclei (e.g. binding energy per nucleon, magnetic and quadripole moment) single out these numbers. A survey of nuclear tables shows that for A = odd nuclei, the parity and spin of the ground state vary in a regular pattern which suggests very much the filling of shells (as in atomic physics). Ground state, spin j and parity $\pm$ depend only on the value of the odd number Z or N so the order of the levels, with increasing energy, is the same for protons and neutrons. The order obtained can be deduced from the three-dimensional harmonic oscillator spectrum $E_n = n\hbar\omega + E_o$ (where ω is a constant) with some modifications.

Consider the set of nine operators $T_{ij} = P_iP_j + Q_iQ_j$ where the P_i and Q_j $(i = 1,2,3)$ satisfy the canonical commutation relation

$$[P_i,Q_j] = i\hbar\delta_{ij} \quad .$$

Its use for the computation of the commutators $[T_{ij},T_{i'j'}]$ shows that the T_{ij} form a representation up to $i\hbar$ of the Lie algebra $U(3)$. The center

$$H = \mathrm{Tr}T_{ij} = T_{11} + T_{22} + T_{33}$$

is the harmonic oscillator Hamiltonian (in convenient units). From $[H,T_{ij}] = 0$ we deduce that there is a $U(3)$ invariance for the three-dimensional harmonic oscillator similar to the $SO(4)$ invariance of the hydrogen atom, 2.4, and we can find the spectrum of H by a similar method:

The ground state $n = 0$ has energy E_o. The number of states of energy E_n is $1/2(n + 1)(n + 2)$. Their orbital angular momentum ℓ satisfies $(-1)^\ell = (-1)^n$, $0 \le \ell \le n$. This corresponds to the first column of Figure 3.4. The second column gives the spectrum of the Hamiltonian

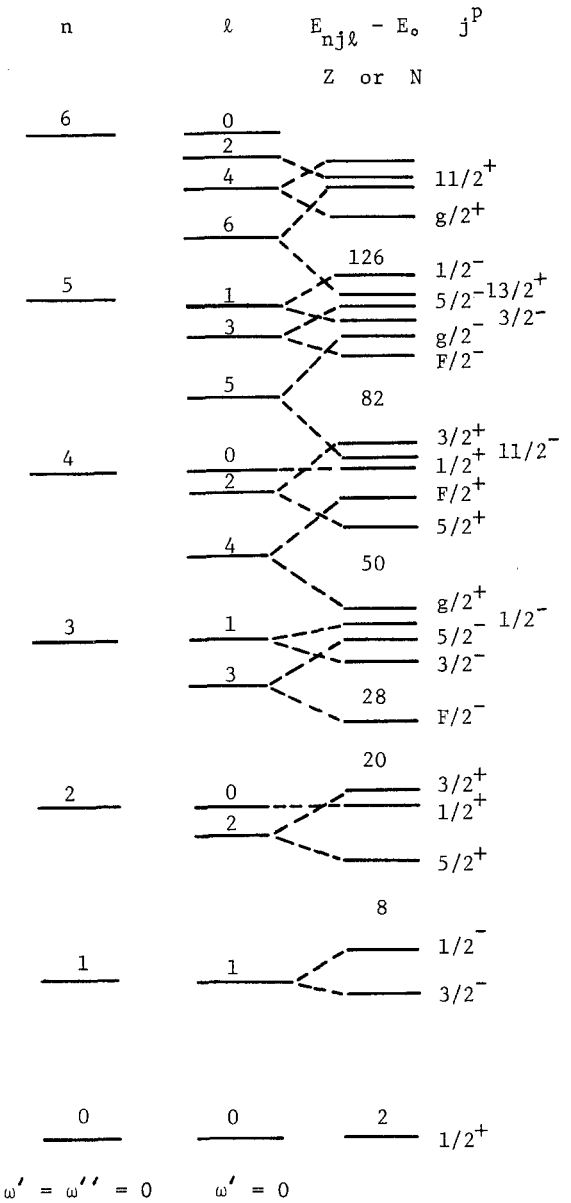
$$\hbar^{-1}H' = \omega H - \omega'\vec{L}^2 - 2\omega''\vec{L}\cdot\vec{S} \tag{3.12}$$

with ω, ω', ω'' positive constants, $\omega > \omega'$ and ω''. Using Equation (2.55) for $s = 1/2$ and when $\ell > 0$, $j = \ell + \varepsilon\, 1/2$, $\varepsilon = \pm 1$ one obtains the energy spectrum

$$\ell > 0 \qquad E_{n,j,\ell} - E_{0,\frac{1}{2},0} = n\omega - \omega'(\ell(\ell + 1) - \varepsilon\omega''(\ell + \tfrac{1}{2}),\ \varepsilon = \mathrm{sign}(j - \ell) \tag{3.13}$$

$$\ell = 0 \qquad = n\omega \tag{3.13'}$$

This Hamiltonian H' is the one-nucleon Hamiltonian in the average potential produced by the whole nucleus. As in 2.7 we can now "fill the successive shells" for protons and neutrons. Such shell-model for nuclei was proposed in 1949 (see M. Goeppert Mayer and J. H. D. Jensen, *Elementary Theory of Nuclear Shell Structure*, Wiley, New York (1955)). It is very successful in explaining the properties of the



j^p is the spin (j half integer > 0) and the parity (±) of the state.

FIGURE 3.4. ENERGY SPECTRUM OF THE ONE NUCLEON HAMILTONIAN OF EQUATION (3.12)

low lying levels of nuclei, and their decays. A more detailed book on nuclear shell structure is that of A. de Shalit and I. Talmi.

Note that a Z (or N) particle state is to a lesser degree a decomposable antisymmetrical tensor than it is for atoms (see 2.8). There is much more configuration mixing, i.e., the nuclear states are tensors which are linear combination of different decomposable tensors with the same quantum numbers.

Shell-model computations require a classification of states inside the "same shell". That was done between 1942 and 1949 by Racah (*Phys. Rev.*, 76, 1352 (1949)) who introduced the seniority quantum number. (See also work of Flowers and many references in Dyson's anthology quoted in 3.3 and 4.3.)

Part of the problem is to label unambiguously irreps D_j of SU(2) appearing in the decomposition into a direct sum of irreps of the tensor power of a given irrep $\overset{n}{\otimes} D_j$ (where J is usually integral for atoms, half integral for nuclei). The method is to find a chain of subgroups

$$H_o = U(2\overline{j} + 1) \supset H_1 \supset \ldots \supset H_k \supset SU(2) \quad \text{or} \quad SO(3) \tag{3.14}$$

(where SU(2), (2J odd) or SO(3), (2J even) is the subgroup of U(2J + 1) formed by the matrices of D_j) with the following property:

The successive restrictions of the representation of U(2J + 1)

$$\overset{n}{\otimes} \square = \oplus_\lambda \, s_\lambda [\lambda] \qquad \text{(see (1.15'))}$$

(where $\oplus_\lambda$ is over all Young diagrams of n squares and s_λ is the dimension of the corresponding irrep of $S(n)$) to the different H_i, $(0 \le 1 \le k)$ must finally yield direct sums of SU(2) irreps with multiplicity one. Table 3.1 illustrates a simple example J = 2, n = 3, the U(5) irrep  and  restricted to SU(2) yields only multiplicity one. This is not the case for . One intermediate group is necessary $H_1 = SO(5)$.

TABLE 3.1. DECOMPOSITION OF $\overset{3}{\otimes} D_2$ (3 PARTICLES IN D-SHELL)

representation of U(5)		$(\square)^3 = \square\square\square \oplus 2$ $+$
dimension		$5^3 = 35 + (2 \times 40) + 10$
Irrep of U(5),	Restriction to SO(5)	Restriction to SO(3)
	= one irrep	$\square\square\square = D_6 \oplus D_4 \oplus D_3 \oplus D_2 \oplus D_0$
	= A ⊕ B	$A = D_5 \oplus D_4 \oplus D_3 \oplus D_2 \oplus D_1$
		$B = D_2$
	= one irrep	 $= D_3 \oplus D_1$

Neglecting the s_λ multiplicity, every inequivalent irrep of SO(3)/or SU(2) has a different genealogy of irreps of H_i. Racah says that they have different seniority quantum number. For distinguishing the different seniorities Racah had to introduce the exceptional Lie group G_2 among the H_i (in the case $J = 3$, $n = 7$)! When J is half integer, one generally takes $H_1 = Sp(2J + 1)$. Of course, nuclear physicists nowadays use more refined models of nuclei (e.g., *Nuclear Structure I, II, III,* by A. Bohr and B. Mottelson, Benjamin). To go into more details is outside the scope of these lectures. We refer again to a non-technical paper by D. R. Inglis: "Nuclear Models", *Physics Today,* 22 no. 6, 29 (1969) for a recent survey.

3.5. The Hadrons

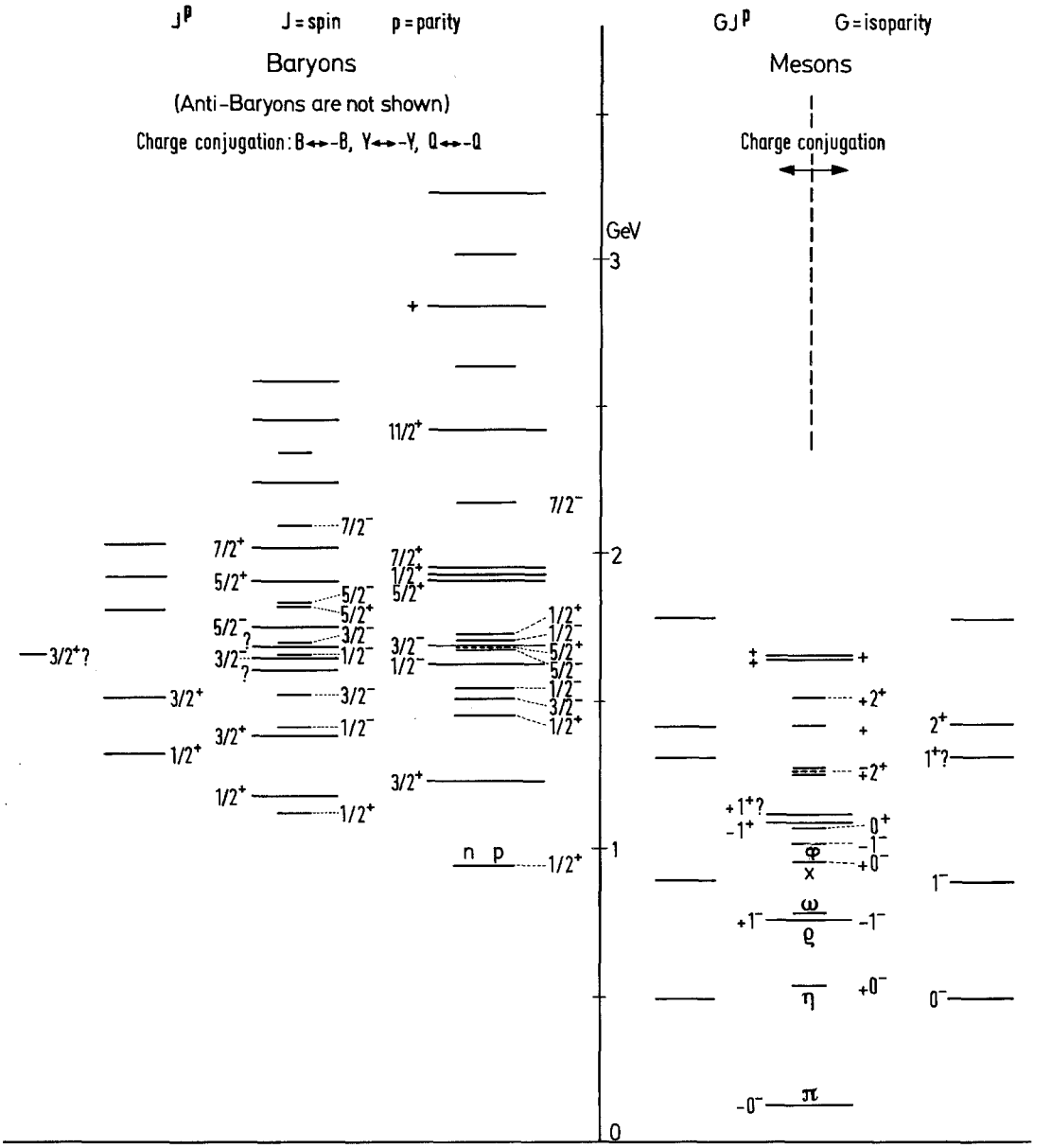
Although SU(2) invariance, through isospin considerations, is more familiar to physicists than $S(n)$ invariance for the study of the property of nuclear interaction not to distinguish between neutrons and protons, is it more fundamental? If one had to deal only with nucleons, the answer is no; both mathematical methods are physically equivalent. However, there are many more particles with strong interaction; they cannot be permuted with the nucleons but they can be attributed an isospin. Let us give as example the π-meson. In 1935, Yukawa predicted the existence of mesons which are to the nuclear interaction what photons are to the electromagnetic interaction. He predicted their electric charge $\pm$, their mass, their lifetime, their decay mode. Soon the particles were discovered but it was a case of mistaken identity with the μ-lepton! The Yukawa particle was discovered in 1947 and is called $\pi^{\pm}$. In 1937, physicists (e.g., Kemmer) showed that 3 states of charge were necessary for the meson, +, 0, -. Indeed, in order that nuclear interaction preserve isospin, they have to be invariant under the corresponding SU(2). In Yukawa's theory the meson field is coupled with the nucleonic current. This current transforms under SU(2) as a tensor operator of the space of the representation $(D_{1/2})^2 = D_1 \oplus D_0$. Then the simplest SU(2) invariant Yukawa coupling which can include electrically charged meson, is of the form

$$\int \vec{j}(x) \cdot \vec{\phi}(x) dx \tag{3.15}$$

where $\vec{j}(x)$ and $\vec{\phi}(x)$ are vector operators for the isospin SU(2) and the interaction is the scalar product of these vectors. The π°-meson so predicted in 1937 was found in 1950.

Already in 1947 two other strongly interacting particles had been found. The generic name "hadron" was given to particles with strong interaction. The rate of discovery of new hadrons has passed from 15 in the fifties to 250 in the sixties. We give their mass spectrum and their spin and parity when known, in Table 3.2;

TABLE 3.2. SPECTROSCOPY OF HADRONS



(Compiled from data of *Rev. Mod. Phys.*, January, 1969.)

different columns of this table correspond to different values of the quantum numbers preserved by the nuclear interaction also called strong interaction.

There is a charge $\underline{b}$ which is for the strong interaction what the electric charge is for the electromagnetic interaction. The "charged" particles have $b = \pm 1$; they are called baryons. The "neutral" particles $b = 0$ are called mesons. It happens that all baryons have half integral spin and all mesons have integral spin. This will be explained in 4.2 (Equation 4.8). There has been no difficulty in attributing an isospin to every baryon. Particles inside the same isospin multiplet have the same spin and parity, their masses are equal within 1% (exceptionally 3% for the π-mesons) and they have different electric charges. Isospin conservation allows us to predict some missing members of a multiplet which are then looked for and have always been found. Instead of using the value of their electric charge q the $2t + 1$ states of the same isospin multiplet can be labeled by the value t_3 of the isospin $SU(2)$ generator in the direction "3". These two labels are equivalent but different; the difference $q - t_3$ depends on the multiplet; since both q and t_3 are conserved by strong interaction, their difference

$$y = 2(q - t_3) \tag{3.16}$$

is a new strong interaction quantum number which can be attributed to each isospin multiplet†; it has integral values, and it is called the hypercharge. To include it, one had enlarged the $SU(2)$ invariance group to a $U(2)$ group††. In 5.1, we will see how $U(2)$ was enlarged further.

In Table 3.2 we have left out the antibaryons, each one is to the corresponding baryon what antiproton is to proton (see Dirac quotation in introduction). Antibaryons are obtained from baryons by the involution called C which changes the sign of the charges b, q, y and leaves invariant the mass and the spin (for more detail on C and its relation with parity, see 4.6). To each isospin multiplet with values t, y, b correspond the C-conjugated multiplet t, $-y$, $-b$. A new quantum number is then necessary for the self-conjugated multiplets ($b = 0$, $y = 0$). I introduced it in 1953††† and called it "isoparity". It is denoted G in Table 3.2.

Taking into account all quantum numbers introduced for hadrons, the invariance group should be written $(U_1 \times U_2) \square Z_2(C)$ where U_1 corresponds to the baryonic charge, $Z_2(C)$ is the two element group generated by C and $\square$ means the semi-direct product. The action of C is equivalent to the complex conjugation of the matrices of $U(1)$ and $U(2)$. Irreps of this group when $b = 0 = y$, and by

---

† Relation (3.16) was guessed by M. Gell-Mann, *Phys. Rev.*, 92, 833 (1953); see also T. Nakano and K. Nishijima, *Prog. Theo. Phys.*, 10, 587 (1953).

†† And not to $U_1 \times SU(2)$ because of the relation $(-1)^y = (-1)^{2t}$ implied by (3.16). See L. Michel reference LM I of 4.9.

††† L. Michel, *N. Cim.*, 10, 319 (1953).

(3.16), t is integral, are faithful representations of $SO(3) \square Z_2(C)$ which is isomorphic to the direct product $SO(3) \times Z_2(C)$.

Finally, let us emphasize that all particles in Table 3.2 with the exception of the proton (and the antiproton) are unstable. Neither their lifetime nor their decay modes are indicated. Most of them are even unstable by the strong interaction with a lifetime of 10^{-23} to 10^{-22} sec. They are often called resonances instead of particles because they do not completely fit this latter concept. (See 4.4.) The particles stable for strong interaction are the lowest states of the columns in Table 3.2 and the first excited states of the column $y = 0$, $b = 1$ (Σ) and $b = 0$ (η). However, to be stable, or unstable does not seem so fundamental! Indeed if the mass difference between the lowest states of two neighboring columns of baryons in Table 3.2 ($\Delta y = 1$) were $> m_k$ the highest of these lowest states would be unstable for strong interaction. The Σ is stable only because $m_\Sigma - m_\Lambda < m_\pi$; if for instance $m_\Delta - m_N < m_\pi$ were satisfied, the Δ would also be stable. The stability of η is due to the fact that both η and π have spin 0, parity $-$ (invariance under P, see 4.7a) and that η-isoparity $= +$ while π-isoparity $= -$.

3.6. The Other Particles and the Other Interactions

There are only nine known "elementary" particles which are not hadrons, i.e., have no strong interaction.

The photon, γ, with mass zero, spin 1 (see 4.4).

The 8 leptons μ^+, μ^-; e^+, e^- (electrons) and their associated zero mass neutrinos ν_{μ^+} ν_{μ^-} ν_{e^+} ν_{e^-}; they all have spin 1/2. Only the μ are unstable.

$$\mu^\pm \to e^\pm + \nu_{e^\pm} + \nu_{\mu^+}$$

because $m_\mu = 207\, m_e$.

All particles have electromagnetic interaction even when they have no electric charge (e.g. $q = 0$; baryons have magnetic moments) but it seems that Ampère's hypothesis† that the whole electromagnetic interaction is through the electromagnetic current $j^\mu(x)$ is well verified; the interaction Hamiltonian is

$$H_{em} = e \int A^\mu(x) j_\mu(x) d^3\vec{x} \tag{3.17}$$

where $A^\mu(x)$ is the photon field (electromagnetic potential vector). In units for which $\hbar = c = 1$, the universal constant e is given by $e^2 = \alpha = 1/137.039$ (see 2.4). Electromagnetic interaction is about 100 times weaker than strong interaction. This is also the order of magnitude of mass difference in an isospin multiplet.

---

† Called nowadays "minimal coupling" in the jargon.

There is another universal interaction, shared by all particles (except the photon), characterized by a universal constant G, introduced by Fermi†, whose value is

$$\frac{G}{\sqrt{2}} = 1.01 \times 10^{-5} \times m_p^2 \tag{3.18}$$

Since it is much weaker than the electromagnetic interaction, this interaction is simply called the "weak" interaction. Fermi postulated as early as 1934 that this interaction has some analogy with the electromagnetic interaction. For instance there are four electrically charged $(\pm)$ "weak" hadronic currents, respectively vectors and axial vectors for the Lorentz group, $v_{\mu\,(\pm)}(x)$, $a_{\mu\,(\pm)}(x)$ which interacts with the leptons through a leptonic current $\ell^{\mu}_{(\pm)}(x)$, and the interaction Hamiltonian being

$$H_w = \frac{G}{\sqrt{2}} \sum_{\varepsilon=\pm 1} \int \ell^{\mu}(\varepsilon)(x) h_{\mu}(\varepsilon)(x) d^{3\rightarrow}x \tag{3.19}$$

with

$$h_{\mu}(\varepsilon)(x) = v_{\mu}(\varepsilon)(x) - a_{\mu}(\varepsilon)(x), \quad (\varepsilon = \pm 1) \tag{3.20}$$

Equation (3.19) has indeed some similarity with (3.17). The fact that h_{μ} is a linear combination of a vector and an axial vector will explain the parity violation of the weak interaction (see 4.7.b).

R. P. Feynman and M. Gell-Mann (*Phys. Rev.*, 109, 193 (1958)) have found a very deep relation among the three interactions. From the unitary representation of $U(2)$ on $\mathcal{H}$, the Hilbert space of hadrons, one obtains the representation F of the $U(2)$-Lie algebra on $\mathcal{H}$. The operators corresponding to the observables y and t_3 are the self-adjoint operators

$$Y = F(y) \quad \text{and} \quad T_3 = F(t_3) \tag{3.21}$$

Since $q = t_3 + 1/2\ y$ (Equation 3.16) for all hadronic states, this relation has also to be true for the self adjoint operators representing these observables, so

$$F(q) = Q = \int j^{\circ}(x) d^{3\rightarrow}x = T_3 + \frac{1}{2} Y \tag{3.22}$$

Note that $\partial_{\mu} j^{\mu}(x) = 0 \Rightarrow Q$ is time independent $\Leftrightarrow [H,Q] = 0$. However, Q here is the total electric charge of the hadronic part of the world, it is not conserved since weak interaction can transfer it to the leptonic part of the world. It is conserved only in the approximation which neglects weak interaction.

The beauty of the discovery by Feynmann and Gell-Mann is that, when electromagnetic and weak interactions are neglected, the vector part of the weak hadronic currents $v'_{\mu}(\varepsilon)(x)$ (Equation 3.20) and the electric current $j_{\mu}(x)$ of the

---

† E. Fermi, "Versuch einer Theorie der β-strahlen", *Z. Physik*, 88, 61 (1934).

hadrons are images of the same tensor operator for the U(2) group of invariance of strong interaction for, respectively, the vectors $t_\pm$ and q of the vector space of the complexified U(2)-Lie algebras

$$t_\pm \wedge y = 0 = y \wedge t_3,\ t_\pm \wedge t_3 = \pm t_\pm \quad . \tag{3.23}$$

This implies that

$$T_\pm = F(t_\pm) = \int v'^0(\pm)(x) d^{3\rightarrow}x \tag{3.24}$$

The isospin group, which was introduced in 3.2 in such a formal and abstract fashion, becomes a physical reality since it is generated by the space integral of the weakly interacting hadronic currents! The addition of the electric charge generates the full U(2) group. When electromagnetic and weak interactions are not neglected, $\partial_\mu v^\mu(x)$ as well as $\partial_\mu j^\mu(x)$ do not vanish and the representation of U(2) on $\mathcal{H}$ becomes: 1) time-dependent for the physicists (just as $[P,Q] = i\hbar\mathbb{1}$ is true at any time with time-dependent P, Q); 2) undefined for the mathematicians (as Coleman and other physicists have shown). Have you noticed the v' instead of v in Equation (3.24)? I have shortened a long story. The Feynman-Gell-Mann hypothesis really needs the enlargement of the U(2) group to SU(3) as we will explain in 5.1 and 5.3.

To stay inside U(2), one has to decompose h_μ of Equation (3.19)

$$h_\mu(\varepsilon)(x) = h'_\mu(\varepsilon)(x) \cos\theta + h''_\mu(\varepsilon)(x) \sin\theta \tag{3.25}$$

where $h'(\varepsilon)$ has hypercharge $y = 0$ and $h''(\varepsilon)$ has $y = \varepsilon$ and θ is the Cabibbo angle†. The same decomposition appears separately for the $v_\mu(\varepsilon)(x)$ and the $a_\mu(\varepsilon)(x)$ part of h (Equation (3.20)). The angle θ has a value 15° so the $|\Delta y| = 1$ weak transitions are slower than those with $|\Delta y| = 0$ by a factor $tg^2\theta$.

They have also a different "selection rule" for isospin. As we just said, v' is a vector-operator for the SU(2) isospin group. This is also true for a' and h'. Hence

weak transitions	$\lvert\Delta Y\rvert = 0$	satisfy	$\lvert\Delta T\rvert = 0$ or 1
while weak transitions	$\lvert\Delta Y\rvert = 1$	satisfy	$\lvert\Delta T\rvert = 1/2$

i.e., h'', v'', a'' are SU(2)-spinor operators.

We have also to mention two other charges conserved by <u>all</u> known interactions (as the baryonic and electric charges). They are the two leptonic charges which seem separately conserved: that of the e-type with value $\varepsilon = \pm 1$ for e^ε, ν_{e^ε} and zero for μ, ν_μ; that of the μ-type: with value $\varepsilon = \pm 1$ for μ^ε, ν_{μ^ε} and zero for e, ν_e.

---

† N. Cabibbo, *Phys. Rev. Lett.*, <u>10</u>, 531 (1963).

4. RELATIVISTIC INVARIANCE. THE DISCRETE SYMMETRIES C. P. T.

4.1. The Poincaré Group and its Automorphisms; Zeeman Theorem

Physicists call Poincaré group the inhomogeneous Lorentz group†. We will denote its connected component by $\mathcal{P}_0$. It is the semi-direct product $T \square L$ of the connected Lorentz group L_0 by the translation group T. It has a trivial center. It is a 10 parameter real Lie group. Its universal covering $\overline{\mathcal{P}}_0$ is the semi-direct product $T \square SL(2,C)$, whose center is a two element group generated by ω = "the rotation by 2π". The group law of $\overline{\mathcal{P}}_0$ is given explicitly in Equation (4.10).

We call P the space reflection $P(\vec{r},t) = (-\vec{r},t)$ and T the time reflection $T(\vec{r},t) = (-\vec{r},t)$, $\mathcal{D}$ the group of dilations $\{\alpha > 0, \alpha(\vec{r},t) = (\alpha\vec{r},\alpha t)\}$. We call $Z_2(P)$, $Z_2(T)$, $Z_2(P) \times Z_2(T)$ the group generated by respectively P, T, P and T. We denote by $\mathcal{P}$, $\mathcal{P}^\uparrow$, $\mathcal{F}^\uparrow$, $\mathcal{F}$ the groups generated by $\mathcal{P}_0$ and respectively $Z_2(P) \times Z_2(T)$, $Z_2(P)$, $Z_2(P)$ and $\mathcal{D}$, $Z_2(P) \times Z_2(T)$ and $\mathcal{D}$. We call $\mathcal{P}$ the full Poincaré group.

It can be proven†† that all automorphisms of these groups are continuous and, if Aut G is the automorphism group of G,

$$\text{Aut } \mathcal{P}_0 = \text{Aut } \mathcal{P}^\uparrow = \text{Aut } \mathcal{P} = \text{Aut } \mathcal{F}^\uparrow = \text{Aut } \mathcal{F} = \mathcal{F} \quad . \tag{4.1}$$

Given any group G, we denote by In.Aut G the group of inner automorphisms and by Out G the quotient Out G = Aut G/In.Aut G. Note that here $\mathcal{P}_0 = \text{In.Aut } \mathcal{P}_0$ and that $\mathcal{F}$ is the semi-direct product

$$\mathcal{F} = \mathcal{P}_0 \square (Z_2 \times Z_2 \times \mathcal{D}) = \mathcal{P}_0 \square \text{Out } \mathcal{P}_0 \quad . \tag{4.2}$$

Binary Relation on Space Time E

Given $x \neq y$ two distinct points of E, we define the notations:

x T y = (y is inside the light cone of x)

s L y = (y is on the light cone of x)

---

† Lorentz transformations were introduced by Vogt in 1882 and applied by Lorentz to electromagnetism. H. Poincaré (*C. R. Acad. Sci., Paris,* 140, 1504 (1905)) required that they form a group with the rotation group and, from it deduced physical consequences. In *Rend. Circ. Mat. Palermo,* 21, 129 (1906) he included the translations and studied physical implications of invariance under the group we call here Poincaré group.

†† L. Michel, "Relations entre symétries internes et invariance relativiste", lectures published in *Application of Mathematics to Problems in Theoretical Physics, Cargèse 1965,* Lurçat editor, Gordon and Breach (1967) referred to as LM III. We will also refer to my lectures in Istanbul (1962) and Brandeis (1965) as LM I and LM II. They are both published by Gordon and Breach (Book of the lectures, for each school).

$x\ S\ y$ = (y is outside the light cone of x)

$x < y$ = (y is inside the future light cone of x)

$x <\cdot\ y$ = (y is on the future light cone of x).

Given a permutation f of the points of E, it is said to preserve the binary relations $x\ R\ y$ if $x\ R\ y \Rightarrow f(x)\ R\ f(y)$. E. C. Zeeman (*J. Math. Phys.*, 5, 490 (1964)) proved the following theorem:

Theorem 1.

The necessary and sufficient condition that f and f^{-1}, permutations of E, preserve the relation $x < y$ or the relation $x <\cdot\ y$, is $f \in \mathcal{F}^{\uparrow}$.

Zeeman also established the corollary (proof published in LM II, p. 297):

Corollary 1.

The necessary and sufficient condition that f, permutation of E, preserves the three relations $x\ T\ y$, $x\ L\ y$, $x\ S\ y$ is $f \in \mathcal{F}$.

4.2. Relativistic Invariance and Internal Symmetries†

A physical theory is relativistic if its automorphism group G contains $\mathcal{P}_0$. We are also interested in other symmetry groups, subgroups of G, and called internal symmetry groups. Note that if we consider "passive" invariance, the dilations $\mathcal{D} \subset G$.

If $\mathcal{P}_0$ is a subgroup of G, one can consider $C = C_G(\mathcal{P}_0)$, the centralizer of $\mathcal{P}_0$ in $G = \{g \in G,\ p \in \mathcal{P}_0,\ gpg^{-1} = p\}$, $N = N_G(\mathcal{P}_0)$, the normalizer of $\mathcal{P}_0$ in $G = \{g \in G,\ p \in \mathcal{P}_0,\ gpg^{-1} \in \mathcal{P}_0\}$.

That $\mathrm{Aut}\ \mathcal{P}_0 = \mathcal{F}$ is the semi-direct product (2) and that $\mathcal{P}$ has no center imply that

$$N = \mathcal{P}_0 \square (N/\mathcal{P}_0) \quad , \tag{4.3}$$

and there is a canonical homomorphism

$$N/\mathcal{P}_0 \xrightarrow{f} \mathrm{Out}\ \mathcal{P}_0 = Z_2(P) \times Z_2(T) \times \mathcal{D} \quad . \tag{4.4}$$

And for instance $Z_2(P) \subset \mathrm{Im}\ f$ means that parity is preserved in the theory. We also see that $\mathcal{D} \cap \mathrm{Im}\ f$ will give information on the mass spectrum. Indeed, a theory of mass zero particle has $\mathcal{D}$ in its automorphism group. If $\mathcal{D} \subset \mathrm{Im}\ f$ and if there is a particle of mass $\neq 0$ then there are particles with the same properties and any $m > 0$ for the mass value.

---

† We also refer the reader to the paper with the same title: L. Michel, *Phys. Rev.*, 137B, 405 (1965).

O'Raifeartaigh (*Phys. Rev. Lett.*, 14, 519 (1965)) has proven the following theorem when G is a connected Lie group:

Theorem 2.

If the restriction of an irrep (= unitary irreducible representation) of G to P_0 has an isolated point in the mass spectrum, it is the whole mass spectrum.

There have been too many papers written by physicists proving "theorems" much weaker than the following trivial lemma (LM III, p. 450).

Lemma.

Let P_0 be a subgroup of G. If there exists $p \in P_0$, $p \notin T \subset P_0$, such that $\forall g \in G$, $gpg^{-1} \in P_0$, then P_0 is an invariant subgroup of G. Indeed consider the homomorphism f, $G \xrightarrow{f}$ permutations of (G/P_0) giving the action $\forall x$, $g \in G$, $gP_0 \overset{x}{\rightsquigarrow} xgP_0$ of G on its homogenous space G/P_0. Then $p \in \text{Ker } f$ so $P_0 \cap \text{Ker } f$ is an invariant subgroup of P_0 containing p; it is P_0 and $P_0 \subset \text{Ker } f$; that implies $\forall q \in P_0$, $\forall g \in G$, $qg = gP_0$.

In my opinion, the preceding considerations are physically very poor, indeed P_0 acts on space time so if $G \supset P_0$ is an automorphism group of the theory, Zeeman's theorem implies that in order to preserve causality, G can act on space time only through a quotient subgroup either of F or P if we forget dilations. This led us to consider G as an extension of P.

We are interested in quantum mechanics. So we must use the existence of the *-algebra A of observables.

We refer the reader to the remarkable paper of Haag and Kastler, "An Algebraic Approach to Quantum Field Theory", *J. Math. Phys.*, 5, 848 (1964); there are physical arguments for A to be a C*-algebra.† Let A be its representation (obtained by a Gelfand-Segal construction) by operators on $\mathcal{H}$, the Hilbert space of states, A' its commutant, A'' the enveloping W*-algebra, $Z = A' \cap A''$ its center. The spectral resolution of Z yields superselection rules††. For instance, if the spectrum is discrete, $\mathcal{H} = \oplus_\lambda \mathcal{H}_\lambda$ and the only vectors of $\mathcal{H}$ which represent states are those belonging to one of the $\mathcal{H}_\lambda$. The $\mathcal{H}_\lambda$ are called superselection sectors.

Assume that P_0 is a subgroup of Aut A, which is implementable (i.e., its elements can be realized by operators of $L(\mathcal{H})$).

---

† This proposition was made by I. E. Segal, more than ten years earlier.

†† Concept introduced by G. C. Wick, A. S. Wightman, E. P. Wigner, *Phys. Rev.*, 88, 101 (1952). See the preprint of Doplicher, Haag and Roberts for the most recent study of this question.

Let $U(p)$ be a realization of the automorphism $p \in \mathcal{P}_0$ by an operator $U(p) \in L(\mathcal{H})$. It has to be unitary in order to be an automorphism: $\forall A \in \mathcal{A}$; $(UAU^{-1})^* = (UA^*U^{-1})$. If V is any element of the group $\mathcal{U}(\mathcal{A}')$ of the unitary operators of $\mathcal{A}'$, $U(p)V$ is just as good for representing the Poincaré transformation p. So the set:

$$\mathcal{E} = \{U(p)V,\ p \in \mathcal{P}_0,\ V \in \mathcal{U}(\mathcal{A}')\} \quad , \tag{4.5}$$

forms a group of unitary operators which is a "central extension" of $\mathcal{P}_0$ by $\mathcal{U}(\mathcal{A}')$ i.e.,

$$\mathcal{E}/\mathcal{U}(\mathcal{A}') = \mathcal{P}_0, \text{ quotient group} \quad . \tag{4.6}$$

and

$$\forall V \in \mathcal{U}(\mathcal{A}'),\ \forall U \in \mathcal{E},\ V \rightsquigarrow UVU^{-1} \text{ is an inner automorphism of } \mathcal{U}(\mathcal{A}') \ . \tag{4.6'}$$

One can prove (see Moore's lectures), that any Polish topological group $\mathcal{E}$ satisfying (4.6) and (4.6') is either the direct product $U(\mathcal{A}') \times \mathcal{P}_0$ or are of the form (see also LM II):

$$\mathcal{E}_\alpha = (\mathcal{U}(\mathcal{A}') \times \overline{\mathcal{P}}_0)/Z_2(\alpha,\omega) \quad , \tag{4.7}$$

where the two element group is generated by the element (α,ω) with $\alpha \in {}_2\mathcal{U}(Z)$, the group of square roots ($\neq e$) of the unit, in the group $\mathcal{U}(Z)$ which is the center of $\mathcal{U}(\mathcal{A}')$.† Which is the extension $\mathcal{E}_\alpha$ chosen by nature?†† The answer is the extension defined by (4.7) with

$$\alpha = e^{i\pi(B+\Sigma_i L_i)}, \quad \alpha^2 = I \tag{4.8}$$

---

† Equation (4.7) implies some topology as explained in Moore's lectures. In "Sur les extensions centrales du groupe de Lorentz inhomogène connexe", *Nucl. Phys.*, 57, 356 (1964), I have studied the same problem for abstract groups: any abelian group A is the direct sum $A = D \oplus K$ where D is the maximal divisible subgroup and K is a reduced subgroup (no infinitely divisible elements $\neq 1$). One has the relations: $H^2(\overline{\mathcal{P}}_0,A) = H^2(SL(2,\mathbb{C}),A)$, $H^2(\mathcal{P}_0,A) = H^2(\mathcal{L}_0,A) = {}_2K + H^2(\mathcal{L}_0,D)$, $H^2(SL(2,\mathbb{C}),D)^{\text{Aut }\mathbb{C}} = 0$ and of course $H^2(\mathcal{L}_0,D)^{\text{Aut}\,C} = {}_2D$. Indeed Aut $\mathbb{C}$ the group of automorphisms of the complex field act on $SL(2,C)$ and on $\mathcal{L}_0$ (exactly Aut $\mathcal{L}_0/\mathcal{L}_0$ = Aut $\mathbb{C}$). So it acts on $H^2(SL(2,\mathbb{C}),A)$ (through a trivial action on A) since the group of inner automorphisms of $\mathcal{L}_0$ acts trivially. Following the usual convention, also used in Chapter 1, $H^2(SL(2,\mathbb{C}),A)^{\text{Aut }\mathbb{C}}$ is the subgroup of fixed elements. So if $H^2(SL(2,\mathbb{C}),A) \neq 0$, the automorphisms of $\mathbb{C}$ do not pass the non-trivial extensions and the corresponding extensions are very pathological. I found this a sufficient argument for considering in physics only the extensions of Equation (4.7).

†† This was the question that Lurçat and myself asked and answered in *N. Cim.*, 21, 57 (1965) and *Comptes Rendus of the Conference of Aix-en-Provence*, p. 183, C.E.A. Saclay editor, (1962).

where B is the baryonic charge operator and L_i the (different) leptonic charges (see 3.6). Indeed this choice of extension implies the observed relation between spin and charges:

$$(-1)^{2j} = (-1)^{b+\Sigma i \ell i} \quad , \tag{4.9}$$

where j is the angular momentum of any state and b, ℓ_i are its baryonic and different leptonic charges. Note that Equation (4.9) shows that the integer or half integer nature of spin form a superselection rule.

4.3. Irrep of $\overline{P}_0$

All irreps (= unitary linear irreducible representations) of $\overline{P}_0$ are known. In 1937, Wigner† showed, by extending Frobenius' methods for finite groups to $\overline{P}_0$, that irrep of $\overline{P}_0$ are characterized by an orbit of $\overline{L}_0$ on T' the dual of T and an irrep of the corresponding little group (= stabilizer). The non-degenerate $\overline{L}_0$ invariant symmetric bilinear form on T (= Minkowski pseudo-Euclidean scalar products) yields an isomorphism of $\overline{L}_0$ space between T and T'.

To be explicit, we denote by $\underline{a}$, $\underline{b}$, ... and A, B, ... respectively the elements of T and $SL(2,\mathbb{C}) = \overline{L}_0$. Let $(a^0,\vec{a}) = (a^0,a^1,a^2,a^3)$ the coordinates of $\underline{a}$ in a basis of T. Consider the isomorphism between $T = R^4$ and the additive group of 2×2 hermitian matrices

$$\underline{a} \longleftrightarrow \tilde{a} = \begin{pmatrix} a^0 + a^3 & a^1 - ia^2 \\ a^1 + ia^2 & a^0 - a^3 \end{pmatrix} \quad .$$

As we saw, the group $\overline{P}_0$ is the semi-direct product $T \square SL(2,C)$ with the $SL(2,C)$ action on T

$$A \in SL(2,C) \quad , \quad \underline{a} \longleftrightarrow \tilde{a} \overset{A}{\rightsquigarrow} A\tilde{a}A^* \longleftrightarrow A\underline{a} \quad .$$

The Minkowski pseudo-Euclidean scalar product is

$$(\underline{a},\underline{b}) = a^0b^0 - a^1b^1 - a^2b^2 - a^3b^3$$

and the Minkowski "length" of $\underline{a}$ is

$$\underline{a}^2 = (\underline{a},\underline{a}) = \text{determinant } \tilde{a} \quad .$$

We denote by $(\underline{a},A)$ the elements of $\overline{P}_0$ with $\underline{a} \rightsquigarrow (\underline{a},1)$ the canonical injection

† E. P. Wigner, *Ann. of Math.*, 40, 149 (1939) reproduced in F. J. Dyson, *Symmetry Groups in Nuclear and Particle Physics*, Benjamin, New York (1966). Wigner's paper was the first one giving a complete family of irreps of a non-compact non-semi-simple Lie group.

$T \to \overline{P}_0$ and $A \rightsquigarrow (0,A)$ an injection† of $\overline{L}_0 = SL(2,C) \rightsquigarrow \overline{P}_0$. The $\overline{P}_0$ group law is

$$(\underline{a},A)(\underline{b},B) = (\underline{a} + A\underline{b},AB) \tag{4.10}$$

We will use the same notation for elements of T and T'.

It is useful to introduce the notion of stratum. When a group G acts on a set M, all the points with conjugate stabilizers form a stratum: in other words, a stratum is the union of all orbits of the same type (i.e., isomorphic as G-homogeneous spaces). The action of $\overline{L}_0$, decomposes T or T' in four strata. See Figure 4.1.

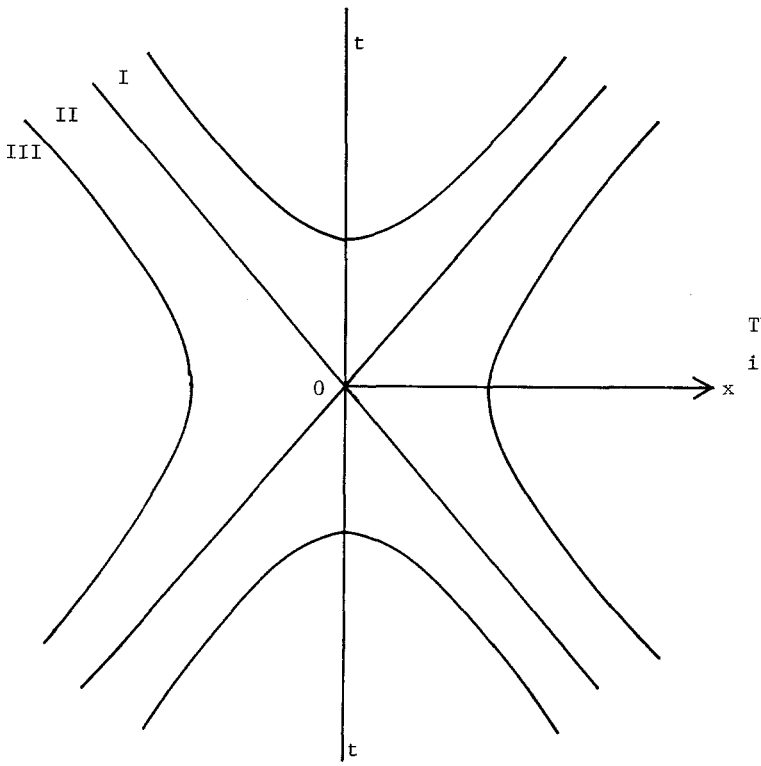


FIGURE 4.1. STRATA ON T BY THE ACTION OF $\overline{L}_0$

† This injection is unique up to a conjugation in $\overline{P}_0$. Indeed Wigner, in his paper of 1939, showed that for the abstract groups (with the action of $\overline{L}_0$ on T just defined), $H^1(\overline{L}_0,T) = 0$.

Table 4.1 lists these strata, the corresponding little groups (defined up to a conjugation in $\overline{L}_0$) and the corresponding irreps of $\overline{P}_0$.

TABLE 4.1. STRATUM OF T' DUE TO THE ACTION OF $\overline{L}_0$; IRREPS OF $\overline{P}_0$

Stratum	Little Group	Irrep of $\overline{P}_0$
I $\underline{p}^2 = m^2$	SU(2)	I_a $m > 0$ (i.e., $p^0 > 0$), 2j integer ≥ 0 I_b $m < 0$ (i.e., $p^0 < 0$), 2j integer ≥ 0
II $\underline{p}^2 = 0$	E(2)	II_a sign of p^0, 2λ integer
$\underline{p} \neq 0$	(2-dimensional Euclidean group)	II_b sign of p^0; Ξ positive number, $\omega = 1$ II_c sign of p^0; Ξ positive number, $\omega = -1$
III $\underline{p}^2 < 0$	$\widetilde{SL}(2,\mathbb{R}) = \widetilde{SO}(2,1)$ $\sim$ = double covering	III_a $m^2 < 0$, ascendant and descendant discrete series, $\pm j$ III_b $m^2 < 0$, principal series $i\sigma,\rho \in R$ III_c $m^2 < 0$, supplementary series, $0 \leq \sigma \leq \frac{1}{4}$
IV $\underline{p} = 0$	$SL(2,\mathbb{C}) = \overline{P}_0$	IV Irreps of $SL(2,\mathbb{C})$ (see Stein lectures) are irreps of $\overline{P}_0$ with T trivially represented.

ω is the non-trivial element of the center of $\overline{P}_0$; in I, ω is represented by $(-1)^{2j}$ and in II by $(-1)^{2\lambda}$. Wigner constructed the irreps of type I and II. Those of $\widetilde{SL}(2,\mathbb{R})$ needed for type III were given by Bargmann, *Ann. of Math.*, 48, 568 (1947) and those of $SL(2,\mathbb{C})$ (type IV for $\overline{P}_0$) were first given by I. M. Gelfand, M. A. Naimark, *Acad. Sci. USSR J. Phys.*, 10, 93 (1946) and *Isv. Akad. Nauk SSSR Ser. Mat.*, 11, 91 (1947).

Wigner's method is a particular case of Mackey's theory of induced representations. Since the (measurable) axis $t'0t$ and $0x$ are a set of representatives of the orbits, a theorem by Mackey† insures that this method yields all irreps of $\overline{P}_0$. As we will see, the only irreps of $\overline{P}_0$ which correspond to known particles are those of mass $m \geq 0$ (I_a and II_a in Table 4.1). Wigner, in his paper, has given the following realization:

An $\overline{L}_0$ invariant measure on the orbit Ω: $\underline{p}^2 = m^2 \geq 0$, $p^0 > 0$, is $d^3\vec{p}/p^0 = d\Omega$. Consider the functions f defined on Ω with value in a $2j + 1$

† G. W. Mackey has described his theory in a book for physicists: *Induced Representations of Groups and Quantum Mechanics*, Benjamin, New York (1968). The needed theorem is Theorem B, p. 43.

dimensional Hilbert K_j, such that

$$\|f\|^2 = \int_\Omega \langle f(\underline{p}), f(\underline{p})\rangle \frac{d\vec{p}}{p_0} < \infty \quad , \tag{4.11}$$

where $\langle f(\underline{p}), f(\underline{p})\rangle = \Sigma_\alpha f_\alpha(\underline{p}) f_\alpha(\underline{p})$ is the hermitian scalar product in K_j. Then

$$(U(\underline{a},A)f_\alpha)(\underline{p}) = \Sigma_\beta e^{i\underline{a}\cdot\underline{p}} Q_{\alpha\beta}(\underline{p},A) f_\beta(A^{-1}\underline{p}) \quad , \tag{4.12}$$

where the unitary $(2j+1) \times (2j+1)$ matrices Q satisfy

$$Q(\underline{p},A)Q(A^{-1}\underline{p},B) = Q(\underline{p},AB) \quad . \tag{4.13}$$

When $j = 0$ or when $m = 0$, $\dim K_0 = 1$, and the Q are complex numbers of unit module. In all cases, when $A, B \in L_p$, the little group of p, Equation (4.13) show that the Q form an irrep of L_p.

4.4. Particle States and Irrep of $\overline{P}_0$

What is a particle? This word is used very much by physicists. This word is attributed to the electron, the photon, and the 270 hadrons of Table 3.2 in 3.5, and also to nuclei (an "α-particle" for instance means a Helium nucleus) and even to atoms or ions. The meaning of this word is in full evolution; it was clear up to ten years ago. Let us try to define this word.

"A particle is a physical system which can be isolated and cannot be decomposed into subsystems without destroying it."

This concept is very clear for stable particles: electrons and positrons $e^\pm$, protons and antiprotons $p^\pm$, photons γ, neutrinos and antineutrinos ν, $\bar{\nu}$ and also stable nuclei (deuteron, α-particle, ${}_6C_{12}$) and the fundamental states of atoms or molecules formed with these nuclei.

Invariants which can be attributed to these particles are the $\overline{P}_0$ invariants; mass and spin, and the Hilbert space of the states of a particle is the carrier of the irrep of $\overline{P}_0$ of mass m, spin j (or for $m = 0$, of helicity λ).† For example, proton or antiproton $(m_p, 1/2)$, electron or positron $(m_e, 1/2)$, neutrinos $(0, -1/2)$, antineutrinos $(0, 1/2)$.

To good approximation this concept of particle can be extended to unstable particles whose lifetime is long enough to study them isolated ($\tau > 10^{-21}$ sec). Strictly speaking, because of Heisenberg's uncertainty relations between energy and time, the Hilbert space of states carries the rep $\int_{\Gamma(\Delta m)}^{\oplus} (m,j)dm$ where Γ is a curve

† The $\overline{P}_0$ invariant λ is called helicity by elementary particle physicists but was called circular polarization by its discoverer, Fresnel, in the 1820's. It just happens that the photon is the only known particle whose space of states is the carrier of a reducible representation $(m = 0, \lambda = 1) \oplus (m = 0, \lambda = -1)$ of P_0.

with a mean spread of Δm. For weak decays, $\Delta m/m < 10^{-14}$ and for electromagnetic decays, $\Delta m/m < 10^{-5}$.

It is not clear that this concept of particle can be usefully extended to the strongly unstable resonances (most of the "hadrons" of Table 3.2 in 3.5). Indeed their lifetime τ might be as short as 10^{-23} sec (and $\Delta m/m$ reach 0.15 for the ρ-meson) so they do not exist isolated since the range of strong interaction is 10^{-13} cm (of the same order as $3 \times 10^{10} \times 10^{-23} = \tau c$). However, we shall here call them particles.

We have used also the word particle as a collective name for different particles with the same spin and not very different masses and similar properties, so they can be considered approximately as identical. This was the case of "the nucleon" with the isospin degree of freedom corresponding to the two states p and n; the π-meson with the three states π^+, π^0, π^-.

There is another degeneracy for most particles. It has been found (or it is expected) that they come in pairs with the same representation (m,j) of $\overline{\mathcal{P}}_0$, but all charges are opposite within the pair. Such pairs are called charge-conjugate pairs, the two values of the corresponding degree of freedom are labeled "particle" and "antiparticle". Particles with all charges zero are called self-charge-conjugated, indeed there is no degeneracy under charge conjugation C for them (examples: $\gamma, \pi^0, \eta^0, \rho^0, \omega^0, \Phi^0, x^0$, etc.).

Elements of the enveloping algebra $E(\mathcal{P}_0)$ of the Lie algebra of $\mathcal{P}_0$ are observables which we will call the kinematical observables of the particle.

The $\mathcal{P}_0$ Lie algebra is (with $M_{\mu\nu} = -M_{\nu\mu}$)

$$[iP^\mu, iP^\lambda] = 0, \quad [iP^\lambda, iM^{\mu\nu}] = iP^\mu g^{\lambda\nu} - iP^\nu g^{\lambda\mu} \quad . \tag{4.14}$$

$$[iM^{\mu\nu}, iM^{\rho\sigma}] = iM^{\mu\rho} g^{\nu\sigma} + iM^{\nu\sigma} g^{\mu\rho} - iM^{\mu\sigma} g^{\nu\rho} - iM^{\nu\rho} g^{\mu\sigma} \quad . \tag{4.14'}$$

The P^λ, $M^{\mu\nu}$ are the self-adjoint operators on $\mathcal{H}$ representing energy momentum and relativistic angular momentum. Pauli (unpublished) and Bargmann and Wigner (*Proc. Nat. Acad. Sci.*, (1967)) considered in $E(\mathcal{P}_0)$:

$$W_\lambda = -\frac{1}{2} \varepsilon_{\mu\lambda\nu\rho} P^\mu M^{\nu\rho} = -\frac{1}{2} \varepsilon_{\lambda\mu\nu\rho} M^{\mu\nu} P^\rho = (^*M \cdot P)_\lambda \tag{4.15}$$

which satisfies

$$[W^\lambda, P^\mu] = 0, \quad [W^\lambda, W^\mu] = i\varepsilon^{\lambda\mu\nu\rho} P_\nu W_\rho \quad , \tag{4.16}$$

$$[W^\lambda, M^{\mu\nu}] = iW^\mu g^{\lambda\nu} - iW^\nu g^{\lambda\mu} \quad . \tag{4.16'}$$

Note that

$$\underline{P} \cdot \underline{W} = P^\lambda W_\lambda = 0 \quad . \tag{4.17}$$

The center of $E(P_0)$ is generated by $\underline{P}^2 = P^\lambda W_\lambda$ and $\underline{W}^2 = W^\lambda W_\lambda$. For irrep of $\overline{P}_0$ these operators are the following multiples of the unit:

$$m > 0, \qquad \underline{P}^2 = m^2 \mathbb{1}, \ \underline{W}^2 = -m^2 j(j+\underline{1})\mathbb{1}$$

$$m = 0 \quad \text{IIa} \qquad \underline{P}^2 = 0, \quad \underline{W}^2 = 0 \quad \underline{W} = \lambda \underline{P}, \ \lambda \text{ helicity}$$

$$m = 0 \quad \text{IIb or IIc} \qquad \underline{P}^2 = 0, \quad W^2 = -\Xi < 0 \ .$$

4.5. Particle Polarization

In most experiments, the energy and momenta of the particles are measured (i.e.: monokinetic beam, target at rest, bubble chamber track curved in a magnetic field). The additional observables to be measured in order to have a complete knowledge of the particle state are called "the polarization". Since $\underline{W}$ commutes with $\underline{P}$, it is the polarization operator. However, the $\underline{W}$ components do not commute with each other. Equations (4.14), (4.14′), (4.15), (4.16), (4.16′) show that

$$P_0, P_1, P_2, P_3, W_3, W^2 \tag{4.18}$$

generate a maximal abelian subalgebra of $E(\overline{P}_0)$. It is easy to interpret† the $\underline{W}$ observables of a particle (m,j) when one remarks that $\mathcal{H}(m,j)$, the one-particle space of states, is a direct integral

$$\mathcal{H}(m,j) = \int_\Omega^\oplus K_j(\underline{p}) \frac{d^3\vec{p}}{p^0} \ , \tag{4.19}$$

of $2j+1$ dimensional Hilbert spaces K_j. Given $\underline{p}$, introduce an orthonormal tetrad of vectors in the Minkowski space (i.e., in T')

$$\underline{n}^{(0)} = \underline{p} m^{-1}, \underline{n}^{(i)} (i = 1,2,3, \alpha,\beta = 0,1,2,3) \ , \tag{4.20}$$

Minkowski scalar product

$$\underline{n}^{(\alpha)}, \underline{n}^{(\beta)} = g^{\alpha\beta} \ , \tag{4.20′}$$

right hand orientation

$$\varepsilon^{\lambda\mu\nu\rho} n_\lambda^{(\alpha)} n_\mu^{(\beta)} n_\nu^{(\nu)} n_\rho^{(\delta)} = -\varepsilon^{\alpha\beta\gamma\delta} \ . \tag{4.20″}$$

Introduce then the self-adjoint operators on $K_j(\underline{p})$

$$S^i = -\frac{1}{m} \underline{n}^{(i)} \cdot \underline{W}(\underline{p}) \tag{4.21}$$

---

† See for instance L. Michel, *N. Cim. Suppl.*, 14, 99 (1959) for more details and the treatment of the case $m = 0$.

where $\underline{W}(\underline{p})$ is the integrand of

$$\underline{W} = \int_{\Omega}^{\oplus} \underline{W}(\underline{p}) \frac{d^{3\rightarrow}p}{p^0} \quad . \tag{4.21'}$$

Of course, Equation (4.17) implies

$$S^0 = -\frac{1}{m}\underline{n}^{(0)} \cdot \underline{W}(\underline{p}) = 0 \quad . \tag{4.21''}$$

The $S^{(i)}$ satisfies the commutation relations

$$[S^{(i)}, S^{(j)}] = i\varepsilon_{ijk}S^{(k)} \quad , \tag{4.22}$$

so they generate a SU(2) Lie algebra, that of the little group of $\underline{p}$. For $\underline{p} = (m,0)$ (particle at rest), $S^{(i)} = M^{0i}(\underline{p})$. This shows the relation between angular momentum and polarization.

For a particle of energy momentum $\underline{p}$, the polarization density matrix $R(\underline{p})$ is a $TrR(\underline{p}) = 1$, hermitian $R^*(\underline{p}) = R(\underline{p}) \geq 0$, $(2j + 1) \times (2j + i)$ matrix which is transformed by $L \in SU(2)$ (defined by Equation (4.22)) as

$$L\underline{p} = \underline{p}, \; R(\underline{p}) \rightsquigarrow Q(\underline{p},L)R(\underline{p})Q^*(\underline{p},L) \quad , \tag{4.23}$$

Let

$$R(\underline{p}) = (2j + 1)^{-1}\mathbb{1} + \sum_{\ell=1}^{2j} R^{(\ell)}(\underline{p}) \quad , \tag{4.24}$$

be the decomposition of $R(\underline{p})$ into a sum of irreducible SU(2)-tensor operators. The $R^{(\ell)}(\underline{p})$ are called the multipoles of the polarization matrix. Note that $R(\underline{p})$ and the $R^{(\ell)}(\underline{p})$ can be given a completely covariant form.

$$R^{(1)}(\underline{p}) = -\frac{1}{2} s_{\alpha} \frac{W^{\alpha}}{m}, R^{(\ell)}(\underline{p}) = \frac{(-1)^{\ell}}{m^{\ell}} s_{\alpha_1\alpha_2\ldots\alpha_\ell} W^{\alpha_1}W^{\alpha_2}\ldots W^{\alpha_\ell} \quad . \tag{4.25}$$

Where $s_{\alpha_1\ldots\alpha_\ell}$ is a completely symmetrical ℓ^{th} order tensor satisfying

$$(\text{partial trace}) = s^{\alpha}_{\alpha\beta\gamma\ldots} = 0, \; p_{\alpha}s^{\alpha}_{\beta\gamma\ldots} = \underline{p} \cdot s = 0 \quad . \tag{4.26}$$

This is obtained† from the equivalent form of relation (4.21)

$$W(\underline{p}) = m\Sigma_i S^{(i)}\underline{n}^{(i)} \quad . \tag{4.27}$$

From now on, we shall consider only the dipole polarization (which is the only one existing for a spin 1/2 particle). Its evolution is given in a macroscopic

---

† For more details see C. Henry and E. De Rafael, *Ann. Inst. H. Poincaré*, 2A, 87 (1965).

($\sim$ slowly variable in space time) electromagnetic field F (with $F^i = E^i, F^{ij} = \varepsilon_{ijk}B^k$ and $*F$ the polar tensor of F) by the equation†

$$\dot{\underline{s}} = \frac{e}{m} M \cdot \underline{s}, \text{ with } M = F + P\left(\frac{(g-2)}{2} F + g' *F\right)P \quad , \tag{4.28}$$

and $P = \mathbb{1} - \underline{u} \otimes \underline{u} = P_{\underline{u}^\perp}$ where $\underline{u} = \underline{p}/m$, the quadri-velocity, which satisfies the Lorentz equation:

$$\dot{\underline{u}} \sim \frac{e}{m} F \cdot \underline{u} \quad . \tag{4.28'}$$

The $\cdot$ means the proper time derivative; M and F are skew symmetric tensors so (4.28) and (4.29) correspond to infinitesimal Lorentz transformations which of course preserve the Minkowski products:

$$\underline{u}^2 = 1, \ \underline{u} \cdot \underline{s} = 0, \ 0 \leq \delta = (-\underline{s}^2)^{1/2} \leq 1 \quad , \tag{4.29}$$

where δ is called the degree of (dipole = vector) polarization. The magnetic moment of the particle is $\mu = (g/2)e/m(\hbar = c = 1)$ and $\mu' = g' e/m$ is its electric dipole moment. (For neutral particles, write directly μ and μ'.) Note that $g = 2$ is a remarkable value of g which simplifies Equation (4.28). This, as first noted by Thomas in 1926, is characteristic of the Poincaré group and does not happen for Galilean invariance (as we commented in 2.6 and 2.9).

Consider a reaction between particles $A + B \to C + D + \ldots$ where all energy momenta are exactly known. Then the transition probability between pure polarization states is $\lambda = |\langle C;D,\ldots|S(p_A,p_B,p_C,p_D\ldots)|A,B\rangle|^2$ where the isometry

$$S(p_A,p_A,p_C,p_D\ldots) \in \mathrm{Hom}\ (\mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{H}_C \otimes \mathcal{H}_D \otimes \ldots) \quad . \tag{4.30}$$

More precisely, S is the restriction of a unitary operator, defined on $\mathcal{H}$ the Hilbert space of physics and called S-matrix in the physics literature. In the general case of given polarization density matrices

$$\lambda(p_A,p_B;p_C,p_D,p\ldots) = \mathrm{Tr} R_{C,D\ldots}(p_C,p_{D\ldots}) S R_{A,B}(p_A,p_B) S^* \quad , \tag{4.30'}$$

where $R_{A,B}(p_A,p_B)$ (resp., $R_{C,D}(p_C,p_{D\ldots})$ are hermitian operators†† on $K_{j_A}(p_A) \otimes K_{j_B}(p_B)$ (resp., $K_{j_C}(p_C) \otimes K_{j_D}(p_D) \otimes \ldots$) which reduce to $\mathbb{1}(2j_A + 1)^{-1}(2j_B + 1)^{-1}$ (etc.) when no polarization is observed. If one observes the polarization of only one of the particles, Equations (4.24), (4.25) and (4.30') show that $\lambda(p_A,p_B;p_C,p_D\ldots)$

---

† V. Bargmann, L. Michel, V. Telegdi, *Phys. Rev. Lett.*, 2, 435 (1959).

†† Practically, for all experiments, there is no correlation between the states of initial particles so $R_{A,B}(p_A,p_B) = R_A(p_A) \otimes R_B(p_B)$.

depends linearly on the different polarization tensors of this particle $(S_\alpha, S_{\alpha\beta}, S_{\alpha\beta\gamma}, \ldots)$.†

4.6. Invariance Under $P \times Z_2(C)$; PCT Theorem

If a physical theory is invariant under a group, say P_0, one can transform the theory by an automorphism $\alpha \in \mathrm{Aut}\, P_0$, (replace everywhere $g \in P_0$ by $\alpha(g)$). If α is an inner automorphism, by definition of P_0 invariance, the transformed theory is equivalent. If α is not an inner automorphism, the transformed theory might not be physically equivalent. If it is, then one can enlarge the invariance group, in order to include this automorphism. It is obvious that dilations are not an active invariance of physical theories (except when only zero masses occur). What can be said about P, T (and their product PT)? We will assume invariance under P and T and also under C, the charge conjugation, and in the next section, see if these invariances are respected in nature.

It seems a reasonable assumption that P does not act on A', the commutant of A, the representation on $\mathcal{H}$ of the algebra of observables (see Equation (4.5)). We do know the action of T on A', because T has to be represented by an antiunitary operator (see 1.2), i.e., by $U(T) = V(T)K$ where $V(T)$ is a unitary operator and K is a complex conjugation (whose choice cannot be canonical). K (as well as $U(T)$) induces an anti-linear automorphism on the algebra $L(\mathcal{H})$ i.e.,

$$K\lambda AK = \bar{\lambda} KAK,\ KABK = KAK\,KBK,\ K(A+B)K = KAK + KBK\ , \tag{4.31}$$

since $K^2 = \mathbb{1}$. Note that if $U = (U^{-1})^*$ is unitary, so is KUK. We assume that T leaves $U(A)$ and $U(Z)$ globally invariant, but acts as an anti-linear automorphism. Finally, we can introduce $U(C)$, the charge conjugation operator on $\mathcal{H}$. By definition C acts trivally on P_0 and anti-commute with all charges. More generally, physical properties of C tell us how it must act on A' which corresponds essentially to internal symmetry. Let

$$D = Z_2(P) \times Z_2(T) \times Z_2(C)\ , \tag{4.32}$$

$$P_c = P \times Z_2(C) = P_0 \rtimes D \tag{4.33}$$

In LM I, I gave the proof kindly tailor made by J. P. Serre for us physicists, (Theorem 1, p. 183).

$$H^2(P_c, U(Z)) = H^2(D, U(Z)) \oplus {}_2U(Z)^D\ , \tag{4.34}$$

(see a similar theorem in Moore's lecture), where ${}_2U(Z)^D$ is the group of the

† If the polarization of more than a final particle is observed one has also to introduce polarization correlations.

square roots of the unit of $U(Z)$ invariant under every element of D. We check that $e^{(i\pi(B+\Sigma_i L_i))} = e^{(-i\pi(B+\Sigma_i L_i))}$ is such an element so relation (4.9) is preserved.

What is the extension in (4.34) chosen by nature? Probably none, as we will see in the next section because P, C, PC (and probably T) are not automorphisms of the physical laws of nature. However, we can consider for D in (4.34), a subgroup of that of (4.32).

Let us first consider parity. Irreps of $\mathcal{P}^\uparrow = \mathcal{P}_0 \sqcup Z_2(P)$ are easily deduced from those of $\mathcal{P}_0$. For $m > 0$, and $m = 0$, $\lambda = 0$, there are two irreps of $\mathcal{P}^\uparrow$, $(m,j,\pm)$ or $(0,0,\pm)$ with opposite parity (eigenvalue of $U(P)$) whose restriction to $\mathcal{P}_0$ is irreducible. For mass zero, $\lambda \neq 0$ irrep of $\mathcal{P}^\uparrow$ are denoted by $(0,|\lambda|)$ because their restriction to $\mathcal{P}_0$ reduces to

$$(0,|\lambda|)_{\mathcal{P}_0} = (0,|\lambda|) \oplus (0,-|\lambda|) \quad . \tag{4.35}$$

Note that, as projective representations of $\mathcal{P}_0$, $(m,s,+)$ and $(m,s,-)$ are equivalent. More generally, since $g \to g^2$ is a surjective homomorphism of $U(Z)$, $H^2(Z_2(P),U(Z)) = 0$. So to speak of the parity of a state is not a canonical statement; only relative parity can be defined for states in the same superselection sector. By convention, the parity of the vacuum is taken +1, as well as that of the proton, the neutron, the electron, the Λ^0.

Wigner in his Istanbul lectures in 1962 (same reference as LM I) has studied the projective irreps of $\mathcal{P}$ (and even $\mathcal{P}_c$). This study can be easily transferred to the study of the extension of $\mathcal{P}$ by $U(Z)$ (and then by $U(\mathcal{A}')$, from general results of group extension by a non-abelian kernel, as explained in LM I). This is not the case for $\mathcal{P}_c$ because $U(C)$, as unitary operator, does not act on the phase of the projective representation, but C as charge conjugation acts non-trivially on Z. We just give here the following results: $U(T)^2$, $U(CPT)^2$, $U(PT)^2$ are canonical (since $U(Z)$ is divisible and $U(Z) \ni g \to g^2$ is surjective) and are $\in {}_2U(Z)$. For non-zero mass states, a choice different from

$$U(T)^2 = U(PT)^2 = U(CPT)^2 = (-1)^{2j} \quad , \tag{4.36}$$

will require that irrep of $\mathcal{P}_c$ restricted to $\mathcal{P}^\uparrow \sqcup Z_2(C)$ are not irreducible. This would correspond to a new degree of freedom for particles which is not observed in nature.†

In usual quantum field theories, relations (4.36) are always satisfied. This is related also to the two following theorems:

---

† See Wigner discussion in his Istanbul notes and for a recent review see H. Goldberg, *N. Cim.*, 60, 509 (1969).

Theorem 3.

The good connection between spin and statistics† is a consequence of the Wightman axioms:†† covariance under $\mathcal{P}_0$ of finite component quantum fields, existence of vacuum, positivity of energy and "locality".

Theorem 4.

These axioms also imply invariance under CPT.†††

4.7. How to Observe Violation

4.7.a. Action of P, T, C on Observables

Let us summarize in Table 4.2 the action of the automorphisms P, T, PT of $\mathcal{P}_0$, on invariants of this group. C acts trivially on them, but exchanges particles and antiparticles. The self conjugated particles are eigenstates of C. For instance, consider quantum electrodynamics; C is an automorphism of this theory. The electromagnetic interaction Hamiltonian is:

$$H_{em} = \int j^{\mu}(x)A_{\mu}(x)d^{3}\vec{x} \quad . \tag{4.37}$$

By definition of C,

$$U(C)j^{\mu}(x)U(C)^{-1} = -j^{\mu}(x) \quad , \tag{4.38}$$

i.e., the electromagnetic current changes sign. So H_{em} is invariant under C if also

$$U(C)A_{\mu}(x)U(C)^{-1} = -A_{\mu}(x) \quad . \tag{4.39}$$

---

† i.e., integral (resp. half integral) spin fields describe particles which satisfy Bose = Completely symmetrical (resp. Fermi = antisymmetrical) statistics. This was proven by Pauli; his last publication in the subject is "Exclusion principle, Lorentz group and Reflection of space time and charge", p. 30 in *Niels Bohr and the Development of Physics*, Pauli editor, Pergamon, New York (1955). There he also proves the CPT theorem, first proven by Lüders and Schwinger.

†† See R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics and All That*, Benjamin, New York (1964); R. Jost, *General Theory of Quantized Fields*, American Mathematical Society, Providence (1965).

††† From weaker axioms (Haag-Araki theory of local observables), H. Epstein, *J. Math. Phys.*, 8, 750 (1967), has proven the CPT invariance of the S matrix. For infinite component fields, neither the connection between spin and statistics, nor the CPT invariance are implied by $\mathcal{P}_0$ invariance. For a counter example, see e.g., I. Todorov, *8th Nobel Symposium*, Wiley (1968).

By definition, $U(C)|0\rangle = |0\rangle$ where $|0\rangle$ is the vacuum. So

$$U(C)A_\mu(x)|0\rangle = -A_\mu(x)|0\rangle \quad , \tag{4.40}$$

i.e., a photon has charged conjugation -1. We have added in Table 4.2 the transformation of the electromagnetic field

$$F_{\mu\nu}(x) = (\partial_\wedge A(x))_{\mu\nu} = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) \quad . \tag{4.41}$$

For T, time reversal, the space part $\vec{j}$ of the e.m. current $j^\mu(x)$ changes sign (as a velocity) while the time component (whose space integral is the electric charge) does not. Hence the time reversal property of A_μ, of H_{em} (invariant), of $F^{0i} = E^i$ (electric field) and $F^{ij} = \varepsilon_{ijk}B^k$ (magnetic field).

Consider Equation (4.28). The quadrivector $mF \cdot u = (-\vec{B}\cdot\vec{p}, -p^0\vec{E} - \vec{p}\times\vec{B})$ transforms under t as d/dt mu. Hence, except for the term in g' (electric dipole) Equation (4.28) is invariant under P, T, PT. The term in g' is incompatible with both P and T.

TABLE 4.2. COVARIANCE UNDER P,T,C, OF THE INVARIANTS OF $\mathcal{P}_0$ AND THE ELECTROMAGNETIC FIELD

Physical Observable	P	T	PT	C	CPT
$\underline{p}_i \cdot \underline{p}_j$, $\underline{s}_i \cdot \underline{s}_j$	+	+	+	+	+
$\underline{p}_i \cdot \underline{s}_j$	-	+	-	+	-
$(\underline{p}_i,\underline{p}_j,\underline{p}_k,\underline{p}_\ell)$, $(\underline{p}_i,\underline{p}_j,\underline{s}_k,\underline{s}_\ell)$	-	-	+	+	+
$(\underline{p}_i,\underline{p}_j,\underline{p}_k,\underline{s}_\ell)$, $(\underline{p}_i,\underline{s}_j,\underline{s}_k,\underline{s}_\ell)$	+	-	-	+	-
helicity λ	-	+	-	+	-
$\vec{E}$ (electric field)	-	+	-	-	+
$\vec{B}$ (magnetic field)	+	-	-	-	+

(a,b,c,d) means determinant of the four components of four vectors.

4.7.b. Parity Violation

The consequence of invariance under P is called parity conservation. Consider two states S_1, S_1' of a physical system corresponding to each other through an "active" plane symmetry Σ, and S_2, S_2' two other states of the same

system also symmetric to each other through Σ. Let $\lambda_{12} = \mathrm{tr}R_1R_2$ and $\lambda'_{12} = \mathrm{tr}R'_1R'_2$ be the respective probabilities of transitions $1 \to 2$.

$$\text{Parity conservation} \Rightarrow \lambda_{12} = \lambda'_{12} \quad . \tag{4.42}$$

If an experiment yields $\lambda_{12} \neq \lambda'_{12}$, it proves parity violation. Since $\Sigma\lambda_{12} = \lambda'_{12}$, $\Sigma\lambda'_{12} = \lambda_{12}$, it means that $\lambda_{12} = a + b$, $\lambda'_{12} = a - b$, where

$$a = \frac{1}{2}(\lambda_{12} + \lambda'_{12}) \text{ is a scalar, } b = \frac{1}{2}(\lambda_{12} - \lambda'_{12}) \text{ is a pseudoscalar.} \tag{4.43}$$

So in a two particle decay of a <u>polarized</u> particle $\underline{p} \to \underline{p}_1 + \underline{p}_2$, (or more generally in a decay where only two energy momenta are observed) P conservation $\Rightarrow$ the angular distribution of decay products depends only on the even polarization multipoles $s_{\alpha\beta}$, $s_{\alpha\beta\gamma\sigma}$, ...†

In 1957, the following experiment was performed. Co^{60} nuclei at rest $(\underline{p} = m,\vec{0})$ were polarized in a magnetic field $\vec{B}$; this gives them a dipole polarization only: $\underline{s} = (0,\lambda\vec{B})$. So P is a symmetry of Co^{60} state $P(m,\vec{0}) = (m,\vec{0})$, $P(0,\lambda\vec{B}) = (0,\lambda\vec{B})$. Those nuclei decay spontaneously (β^- radioactivity) emitting electrons of energy momentum $\underline{q}(q^0,\vec{q})$ with an angular dependence proportional to $\underline{s} \cdot \underline{q} = -\lambda\vec{B} \cdot \vec{q} = -\lambda Bq \cos\theta$. This decay proved parity violation.

Similarly, in the spontaneous decay of zero spin π mesons $(\underline{p}_\pi^2 = m_\pi^2)$:

$$\pi^\pm \to \mu^\pm + \nu_{\mu^{\mp}} \quad ,$$

into a spin 1/2 μ-lepton and a massless $\nu_\pm$ (- for neutrinos, + for antineutrinos). The μ-lepton has a polarization s_μ (which can depend only on the observed quantitatives $\underline{p}_\pi = \underline{p}_\mu + \underline{p}_\nu$, $p_\nu^2 = 0$; remember $s_\mu \cdot p_\mu = 0$; see LM II).

$$\underline{s}_\mu = \mp\left(\frac{m_\pi^2 + m_\mu^2}{m_\pi^2 - m_\mu^2}\frac{\underline{p}_\mu}{m_\mu} - \frac{2m_\mu}{m_\pi^2 - m_\mu^2}\underline{p}_\pi\right) \quad , \tag{4.44}$$

where $\mp$ depends on the sign of the (electric charge of) $\mu^\pm$. This proves C and also P violation (by observation of a pseudoscalar $p_\pi \cdot s_\mu$ in the decay).

By the same type of argumentation we verify that those experiments are compatible with CP invariance. Note that in π-decay, the μ-polarization $\underline{s}_\mu$ (given by (4.44′)) satisfies $\underline{s}_\mu^2 = -1$ (complete polarization). Then P_0 invariance (through angular momentum conservation) requires that the accompanying $\nu_{\mp}$ is emitted in a pure helicity state $\lambda = \mp 1$. All observation on neutrinos helicity suggest that $\nu_{\mp}$ has helicity $\mp$ for both ν_μ and ν_e.

---

† See Equation (4.25) and, at the end of 4.5, the property for λ to be <u>linear</u> in s_α, $s_{\alpha\beta}$, $s_{\alpha\beta\gamma}$... .

This shows that the set of neutrino states in not invariant under P or C, and it implies that all reactions with neutrinos violate P and C. But neutrino-less (in fact, non-leptonic) decays of hyperons also violate parity. Example: $\Lambda^0 \to p^+ + \pi^-$, the angular distribution depends on $\underline{s}_\Lambda \cdot \underline{p}_p = -\underline{s}_\Lambda \cdot \underline{p}_\pi$ (since $\underline{s}_\Lambda \cdot \underline{p}_\Lambda = 0$).

4.7.c. Time Reversal Invariance

It would be better to call it "velocity reversal" since $T(\vec{p}/p^0 = \vec{v}) = -\vec{v}$.

Let S_1 and S_2 be two states of a physical system and S_1^T, S_2^T the corresponding states obtained by a T active transformation ($p^0 \rightsquigarrow p^0$, $\vec{p} \rightsquigarrow -\vec{p}$, $s^0 \rightsquigarrow s^0$, $\vec{s} \rightsquigarrow -\vec{s}$, $\lambda \rightsquigarrow \lambda$, etc.). Then

$$T \Rightarrow \lambda_{12} = \lambda_{21}^T \quad . \tag{4.45}$$

Note the reversal of time ordering for the two transitions. A precise experiment comparing the cross section of the two inverse reactions $\to$ and $\leftarrow$

$$\gamma + d^+ \overset{\to}{\leftarrow} p^+ + n \quad , \tag{4.46}$$

is in progress. (The rates are equal for pure states; since polarization is not observed one has to divide the rate by the dimension of the polarization space K_j for the particles $\to (2\ 1/2 + 1)^2 = 4$, $\leftarrow 2 \times 3 = 6$.)

Consider an elastic process (same initial and final particles) such as $\pi^- + p^+ \to \pi^- + p^+$ and compare the final polarization $\underline{s}'_p$ of the proton with the initial polarization of the proton target in another experiment. We must have

$$\lambda(\vec{p}_\pi, \vec{p}_p \to \vec{p}'_\pi, \vec{p}'_p, \vec{s}'_p) = \lambda(-\vec{p}'_\pi, -\vec{p}'_p, -\vec{s}'_p \to -\vec{p}_\pi, -\vec{p}_p) \quad . \tag{4.47}$$

There is an approximate condition of T invariance, in perturbation theory which is based on the following expansion of the "S-matrix".

$$S = I + iH + O(H^2) \quad , \tag{4.48}$$

where H has to be a self-adjoint operator (write $SS^* = S^*S = 1$, in first order in H). In this form, we have for orthogonal states (i.e., $R_i R_j = 0$)

$$\mathrm{Tr} R_j S R_i S^* \sim \mathrm{Tr} R_j H R_i H = \mathrm{Tr} R_i H R_j H \quad , \tag{4.49}$$

i.e., <u>in this approximation</u>

$$\lambda_{ij} = \lambda_{ji} \quad . \tag{4.50}$$

Then in this approximation, (4.45) reads $\lambda_{12} = \lambda_{12}^T$. Even in this approximation there is no positive evidence of violation of time reversal in physics, with perhaps the exception of K^0-decay (next section).

Note that in Equation (4.28), using Table 4.2, the term in g' (electric dipole) is not compatible with time reversal invariance (or with P invariance). So the existence of an electric dipole for an elementary particle would prove violation of both P and T. Experimentally g'_{neutron} is known to be $\leq 10^{-9}$ and $g'_{\text{electron}} \leq 10^{-12}$.

Note that PT invariance has a simple formulation. For example

$$\lambda(p_1 + p_2 \to p'_1 + p'_2) = \lambda(p'_1 + p'_2 \to p_1 + p_2) \quad ,$$

for spinless particles or for pure states (then change $\underline{s} \to -\underline{s}$, $s_{\alpha_1 \cdots \alpha_k} \to (-1)^k s_{\alpha_1 \cdots \alpha_k}$). This is known in physical literature as the "principle of detailed balancing".

4.8. CP Violation

CP violation was first observed by I. H. Christenson, C. W. Cronin, V. L. Fitch and R. Turley, *Phys. Rev. Lett.*, 13, 138 (1964) in K^0-decay. Many experiments have confirmed it.

The state of a K^0 or $\overline{K}^0$ (= anti-K^0, Y = -1) can be described by the Hilbert space $\mathcal{H}^{(1)} = L_2(R^3,t) \otimes K_2$ where K_2 is the vector space of functions defined on the two element set (Y = 1, Y = -1). Then C is of the form $I \otimes C$ while P is of the form $P \otimes I$ so $PC = P \otimes C$. We assume that $P^2 = 1$, $C^2 = 1$, $PC = CP$ so $(PC)^2 = 1$ (as we have seen in 4.6, for spin 0, another assumption will increase the degree of freedom of K's). So we can write

$$\mathcal{H}^{(1)} = \mathcal{H}^{(1)}_+ \oplus \mathcal{H}^{(1)}_- \quad , \tag{4.51}$$

$$CP\mathcal{H}^{(1)}_\pm = \pm\mathcal{H}^{(1)}_\pm \quad . \tag{4.51'}$$

Now it is easy to deduce the action of CP on states of two π^0. These are two identical self-conjugated particles hence any state of $2\pi^0$ is eigenstate with value +1 for C. The tensor product of the representation (m,0) of $\mathcal{P}^\uparrow$, by itself yields

$$\overset{2}{\otimes}(m,0) = \sum_{\ell=0}^{\infty} \oplus \int_{2m}^{\infty} (m,\ell)dm \quad , \tag{5.52}$$

with symmetry

$$\square\square \text{ for even } \ell, \quad \begin{array}{|c|}\hline \\ \hline \end{array} \text{ for odd } \ell \quad . \tag{4.52'}$$

Only the $\square\square$ = symmetric states are allowed by Bose statistics. And (by an argument essentially similar to that yielding Equation (2.40), P acts in the space of Equation (4.51) by multiplication by $(-1)^\ell$ in each direct summand. So states of $2\pi^0$ are eigenstates of C, P, CP with eigenvalue +1.

For states of $\pi^+ + \pi^-$, one has to consider these two particles as identical in order to apply Bose statistics, but in the two different possible states of charge (+ and -). So states of $\pi^+\pi^-$ of total spin ℓ, are eigenstates of C, P, CP with eigenvalue $(-1)^\ell$, $(-1)^\ell$, 1. When CP was believed to be preserved, it was predicted that states of $\mathcal{K}_+^{(1)}$ in Equation (4.51) would decay into 2π while states of $\mathcal{K}_-^{(1)}$ would decay into 3π states which are eigenstates of CP with eigenvalue -1 (as e.g. all $3\pi^0$ states). This was exactly observed and the states of the two spaces $\mathcal{K}_+^{(1)}$ and $\mathcal{K}_-^{(1)}$ were also called "short" and "long" because the 2π-decay is faster.

In 1964, the above quoted experiment proved that the long lived meson also decays into 2π (with a rate $\sim 10^6$ slower than the short lived).

We do know that the universe around our galaxy is not CP invariant, but the influence of this asymmetry (which could depend on the relative velocity of the K-meson with respect to the galaxy, or the earth) seems to be ruled out by more precise experiments.

Must we conclude that there is a small violation ($(10^{-6})^{1/2} = 10^{-3}$ in amplitude) of CP in the transition $K \to 2\pi$? Another possibility could be that CP is conserved in this transition but that the two observed meson with exponential decay: short-lived K_S and long-lived K_L are non-orthogonal states with respectively a large c_S and a small c_L component in $\mathcal{K}_+^{(1)}$. Then the branching ratio

$$b_S = \frac{K_S \to 2\pi^0}{(K_S \to 2\pi^+ + \pi^-)} \quad \text{and} \quad b_\ell = \frac{K_L \to 2\pi^2}{K_L \to \pi^+ + \pi^-} \quad ,$$

should be equal, since they would be the branching ratio of all the states in $\mathcal{K}_+^{(1)}$. The value of b_S is $\sim 1/2$ (as predicted by the selection rule $\Delta\vec{T} = 1/2$, see 3.6). The first measured values of b_ℓ were around 10 to 12, but a value zero appears in another experiment. The present experimental evidence is still an incompatible set but "optimists" say it is compatible with $b_L \sim b_S \sim 1/2$.

So it is possible that CP violation is due to a still undetected interaction, to which no particle transition or spontaneous decay can be attributed, and which has to be superweak.

CP violation has also been observed in $K_L \to \pi^\pm + \ell^\mp + \nu$ (where $\nu\ell = \mu$ or e) decay; there is a relative difference of 3.10^{-3} in the two C or CP conjugated rates. But CP violation has not yet been observed anywhere else.

Of course physicists have proposed many theories (about thirty not yet ruled out by the meager experimental data) to explain CP violation. There is no possibility to give more details here.

To conclude, let us just remark that there is no evidence against CPT violation and there is one fact which suggests that CPT is a "much better" invariance than CP: a small upper limit of the $K^0 - \bar{K}^0$ mass difference is well

known. It is $m_{K^0} - m_{\overline{K}^0} < 10^{-14} m_K$. Such a perfect equality cannot be due to chance and suggests an invariance in nature which contains C. However, we have seen that C, CT, CP are ruled out, so CPT is the likely candidate in agreement with the CPT theorem 4.6.

Remark on Galilean Invariance

We dealt in Section 4 with relativistic Poincaré invariance only. Although we sometimes spoke in Chapters 2 and 3 of Galilean invariance, such invariance was not thoroughly used in atomic and nuclear physics. E. Inönü and E. P. Wigner characterized the irreps of $\overline{G}$, the covering of the Galilee group, in 1952 (*N. Cim.*, 9, 705).

Their results did not fit with physics. V. Bargmann (*Ann. of Math.*, 59, 1 (1954)) showed that for central extensions of the G Lie algebra g, $H^2(g,R) = R$. For each irrep of $\overline{G}$, this yields a family of projective irreps depending on one parameter m which corresponds to the mass of the particle.

See also O'Raifeartaigh's lectures where it is shown that projective irreps of an invariance group also appear in classical mechanics.

5. THE INTERNAL SYMMETRIES OF HADRONS

5.1. SU(3) Symmetry

5.1.a. The Octets

Table 3.2 of "elementary particles" in 3.5 is reminiscent of similar tables of atomic and nuclear spectra.

So, before a dozen of baryons and as many mesons were known, physicists were searching for a larger symmetry than that of U_2 (isospin and hypercharge) which we have described in 3.5. There is no point and no time to tell here about the ill-fated choices except to mention that of Sakata, with a U(3) group whose fundamental representation was spanned by p, n, Λ, the first three known baryons. (S. Sakata, *Prog. Theor. Phys.*, 16, 686 (1956).)

Just as Heisenberg proposed to consider neutron and proton as two states of the same spin 1/2 particle, the nucleon, by neglecting their very small mass difference (or more precisely attributing it to an electromagnetic self-mass effect), the eight known spin $\frac{1}{2}^+$ baryons p, n, Λ^0, Σ^-, Σ^0, Σ^+, Ξ^-, Ξ^0 could be considered as eight states of the "same" particle although the mass difference is of the order of 15 percent instead of 0.15 percent.

By 1961, seven pseudoscalar mesons (0^-) were known, with the same grouping in isospin and hypercharge $y = 1$, $t = 1/2$, K^+K^0; $y = -1$, $t = 1/2$, K^-K^0; $y = 0$, $t = 1$, $\pi^+\pi^0\pi^-$ but the spread in mass was much larger.

M. Gell-Mann and Y. Ne'emann independently proposed to use $SU(3)$ as a classifying group; the eight $\frac{1}{2}^+$ baryons and, predicting a $y = 0$, $t = 0$ pseudoscalar meson which was discovered a few months later and called η^0, the eight 0^- mesons form two octets = eight dimension space E_8 of the adjoint representation of $SU(3)$ ($\square\!\square$ in Young diagram notation). For instance the Hilbert space of states of one baryon is the tensor product $L(m, \frac{1}{2}^+) \otimes K(\square\!\square)$ where $L(m, \frac{1}{2}^+)$ is the space of the irrep $(m, \frac{1}{2}^+)$ of P, the Poincaré group and $K(\square\!\square)$ the octet space E_8. $SU(3)$ is an exact symmetry when the baryon mass differences are neglected. We can say that strong interactions will be decomposed into two parts: a strong SU-3 invariant part and a semi-strong part invariant under the subgroup $U_2(T,Y)$ only. This fits the reduction†

$$\square \text{ of } SU(3)|_{U(2)} = \square \otimes \cdots \otimes \square\!\square \otimes \blacksquare = (1,\tfrac{1}{2}) \quad (0,0) \quad (0,1) \quad (-1,\tfrac{1}{2}) \quad (y,t) \qquad (5.1)$$

$$\dim \quad 8 \quad = \quad 2 \quad + \quad 1 \quad + \quad 3 \quad + \quad 2 \quad .$$

But, would it be possible to consider the $SU(3)$ breaking semi-strong interaction as a perturbation of the very strong interaction? Surely, if you are an optimist. After all 15 percent (effect in baryon mass) is small compared to 1.

Let us now study the mass splitting within the $SU(3)$ multiplet.

5.1.b. The Mass Operator

The simple hypothesis for the mass operator M is that it can be decomposed into

$$M = M_0 + M'(y) \quad , \qquad (5.2)$$

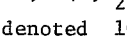
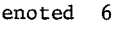
where M_0 is a "scalar" tensor operator and $M'(y)$ is the image of y (of the Lie algebra of $SU(3)$) by an octet = E_8-tensor operator. Let E be the space of an irreducible representation of $SU(3)$. Because $SU(3)$ is of rank two, or equivalently has two zero roots (which are zero weights for $\square$)††

$$\dim \mathrm{Hom}(E \otimes E_8, E)^{SU(3)} \leq 2 \quad . \qquad (5.3)$$

---

† For $u \in U(2)$, the black column means $(\det u)^{-1}$ while $\square$ means $(\det u)$.

†† If $\lambda_1 \geq \lambda_2 \geq 0$ are the number of squares in the first and second line of the Young diagram of an irrep of $SU(3)$, one also uses the notation $(\lambda_1 - \lambda_2, \lambda_2)$ for the irrep of $SU(3)$. The contragredient of (p,q) is (q,p), so (p,p) is self-contragredient, as in $\square$ $= (1,1)$, while $\square\!\square\!\square$ $= (3,0)$ of dimension 10, has for contragredient $\square\!\square\!\square$ denoted $\overline{10}$ by the physicists.

More precisely, it is 2, except for the trivial irrep, for which it is zero, and it is 1 for the irreps whose Young diagram has only 1 line (i.e., $(\lambda_1,0)$), or two equal lines $\lambda_1 = \lambda_2$, (i.e., $(0,\lambda_2)$): for example,  and its contragredient  which are also denoted 10 and $\overline{10}$ because they are of dimension 10. This is also true for = (1,0) and  = (0,1) denoted 3 and $\overline{3}$; and  = (2,0),  = (0,2) denoted 6 and $\overline{6}$.

Another way to interpret (5.3) is to say that on the Hilbert space K of an irrep of SU(3) there are at most two linearly independent octet-tensor operators. Thus, in the approximation where U(2) is an exact symmetry (i.e., neglect of electromagnetic and weak interactions) the particle masses in a multiplet depend on three parameters (one, the expectation value of M_0, and two at most for $M_1(y)$). From 1.5, we know that we can take for each E, as linearly independent octet-tensor operators F and $D = F \vee F$, where $x \rightsquigarrow F(x)$ is the representation (up to the factor i) of the SU(3) Lie algebra on $\mathcal{H}$, the Hilbert space of hadronic states; it satisfies $F \wedge F = iF$. Explicitly, for any $\underset{\sim}{p}_1$, $\underset{\sim}{p}_2 \in K$ the space of an irrep of SU(3) in $\mathcal{H}$ and for any octet-tensor operator T

$$\langle \underset{\sim}{p}_1, T(x) \underset{\sim}{p}_2 \rangle = \langle \underset{\sim}{p}_1, (\alpha F(x) + \beta D(x)) \underset{\sim}{p}_2 \rangle \quad . \tag{5.4}$$

In the physics literature α/β is called the F/D ratio. If the octet part $M'(y)$ (see Equation (5.2)) of the mass operator has no matrix elements between two subspaces of $\mathcal{H}$ carriers of inequivalent SU(3) irreps,† this implies that

$$M = M_0 + M_1 F(y) + M_2 D(y) \quad , \tag{5.5}$$

where M_0, M_1, M_2 are SU(3) scalar operators. The operators $F(y)$ and $D(y)$ commute and their common eigenspaces are U(2) multiplets, so they are functions of Y and $T(T+1)$, the generators of the center of the enveloping algebra of $U_y(2)$. By definition $F(y)$ is proportional to Y, the hypercharge operator, and by computation one finds

$$D(y) = \vec{T}^2 - \frac{1}{4} Y^2 - \frac{1}{3} K \quad , \tag{5.6}$$

where K is the (quadratic) Casimir operator of SU(3). So with a convenient change in the definition of the scalar operators, on a SU(3) multiplet the mass of a state of hypercharge y, isospin t

$$m = m_0' + m_1' + m_2 (t(t+1) - \frac{1}{4} y^2 \quad . \tag{5.7}$$

Applied to the octet of Baryons N, Λ, Σ, Ξ this yields a relation between their four masses

$$\frac{1}{2} (m_N + m_\Xi) = \frac{1}{4} (3m_\Lambda + m_\Sigma) \quad , \tag{5.8}$$

(Gell-Mann, Okubo mass relation) which is well verified within few MeV (for mass $> 10^3$ MeV!).

---

† There are exceptions to this rule: see 5.1.d, the vector mesons. To convey the main idea, we simplify here too much.

For mesons (zero baryonic charge) because of the charge conjugation between particles and antiparticles, M_1' must be zero. The Gell-Mann Okubo mass relation for pseudo-scalar-mesons

$$m_K = \frac{1}{4}(m_\pi + 3m_\eta) \quad , \tag{5.9}$$

is verified only within 50 MeV, about 1/10 of the K and η mass. Optimistic physicists have found good reasons why this relation should be better verified by m^2 (instead of m).

5.1.c. The First Baryon Decuplet

When SU(3) was proposed as symmetry group in 1961 only the first N and Σ excited states, $\Delta(j^p = \frac{3^+}{2}, t = 3/2)$, $\Sigma^*(j^p = \frac{3^+}{2}, t = 1)$† were known. Gell-Mann putting them in a 10 representation, predicted a Ξ^*, $(j^p = \frac{3^+}{2}, t = 1/2, y = -1$, excited state of Ξ) and finally a particle $\Omega(j^p = \frac{3^+}{2}, t = 0, y = -2$. As we have seen in the 10 (i.e., $\square\square\square$) irrep, the mass must depend linearly on two parameters (one for M_0 and only one for $M'(y)$), so in this decuplet the Gell-Mann and Okubo relation predicts for the mass m_y of the states of hypercharge y,

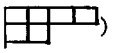
$$m_y = m_{\Sigma^*} - (m_{\Sigma^*} - m_\Delta)y \quad . \tag{5.10}$$

A few months later (in 1962) the predicted Ξ^* was found with a mass of 1530 MeV (to be compared to the predicted value (1385 + (1385 - 1236) = 1534 MeV!). It was later established that its spin is 3/2 and it has the same relative parity as Ξ. But the Ω^-, which should be stable against strong and electromagnetic decay, since it would be the lowest hadronic state with b = 1, y = -2, was frantically looked for and not found ... immediately. Many physicists had given up hope and given explanation why the Ω did not exist, before it was found in 1964 after two and a half years of feverish impatience. The Ω mass is 1672 MeV, (to be compared to the predicted 1677 MeV). Its spin has not yet been measured since less than a score of Ω particles have been observed up to now. If it had not been looked for where it was predicted, when would the Ω have been observed by chance?

5.1.d. Other SU(3) Multiplets

The known experimental data at a given date give a deformed view of the SU(3)-multiplets. In the baryon case for instance, no excited states of the Ω are known yet, although $\frac{5^+}{2}$ and other decuplets probably exist. Some octets have been tentatively identified, although too few excited Ξ states are yet known and their quantum numbers are not measured.

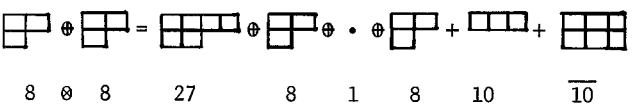
† Is often called also Y*. We denote by j^p the spin j and parity p.

The mesons seem to prefer to occur in nonets. Indeed a $q = 0$, $y = 0$, $t = 0$, 0^- meson is known in addition to the octet of 0^-. A nonet of 1^-: ρ, ω, ϕ, K^*, $\bar{K}^*$, is very well known. The mass formula could not apply to the known "octet" and the ϕ was predicted. The ω and ϕ are orthogonal states of "mixed configuration" $q = y = t = 0$, $\omega = |1\rangle \cos\alpha + |8\rangle \sin\alpha$, $\phi = |8\rangle \cos\alpha - |1\rangle \sin\alpha$, where $|1\rangle$ is a SU(3) singlet and $|8\rangle$ is the octet vector $q = y = t = 0$. A nonet of 2^+ is also well established and an octet of 1^+ is likely. There is some possibility of a 27-plet () of baryons (not indicated in the Table 2.3, for the experimental data are still preliminary). It is to be noted that only irreps of the adjoint group $SU(3)/Z_3$ do appear.

5.1.e. Cross-Sections and Decays of Resonsances

SU(3) invariance implies ratios of resonances decay rates (measured by the natural width and the different branching ratios) into lighter hadrons. This yields remarkably good predictions and explains strange facts such as the small branching ratio for the decay of ϕ into 2π.

For two octet-particle reactions $A + B \to C + D$, one can deduce that the scattering amplitude belongs to the representation



$$8 \otimes 8 \quad 27 \quad 8 \quad 1 \quad 8 \quad 10 \quad \overline{10} \qquad (5.11)$$

symmetric antisymmetric

which yields seven arbitrary parameters. There are less in $8 \otimes 8 \to 8 \otimes 10$. The way to correct for the mass difference is not obvious and the predictions are not spectacular.

An anthology of original papers in SU(3) has been published by Gell-Mann and Ne'emann, *The Eightfold Way*, Benjamin, New York (1964). There is also a book on this subject by M.Gourdin, *Unitary Symmetry*, North-Holland, Amsterdam (1967).

5.2. Geometry on the SU(3)-Octet†

We give here some geometrical properties of the adjoint representation of SU(3).

We have defined in (1.18), (1.19), and (1.19′) the SU(3) invariant scalar product (x,y), the Lie algebra product $x \wedge y$, and the symmetric algebra product

† Full proofs and more results are given in a preprint of L. Michel and L. Radicati, with this title. It also contains some generalizations to SU(n).

$x \vee y$ for any pair of elements $x, y \in E_{n^2-1}$, the real vector space of the adjoint irrep of $SU(n)$. We restrict ourselves here to $n = 3$ and call E_8 the octet space. Its elements can be realized as 3×3 traceless hermitian matrices. They satisfy the equation

$$x^3 - (x,x)x - \mathbb{1} \det x = 0 \quad , \tag{5.13}$$

whose coefficients obey the relation

$$4(x,x)^3 \geq 27(\det x)^2 \quad . \tag{5.14}$$

We find that

$$\det x = \frac{2}{3}(x, x \vee x) \quad , \tag{5.15}$$

so (5.14) can also be written

$$(x,x)^3 \geq 3(x, x \vee x)^2 \quad . \tag{5.16}$$

Orbits of $SU(3)$ on E_8 are in a bijective correspondence with the pairs of real numbers (x,x), $(x, x \vee x)$ satisfying (5.16). When $(x,x)^3 > 3(x, x \vee x)^2$, x is called a regular element of E_8 and its isotropy group G_x is $U(1) \times U(1)$. Its Lie algebra is a Cartan subalgebra and it is generated by x, and $x \vee x$. When $(x,x)^3 = 3(x, x \vee x)^2$, x is called an exceptional element and its isotropy group is $U(2)$. We will also call such x a q-vector or a pseudo-root. We will use from now on only normalized vectors : $(x,x) = 1$. Those vectors r satisfying $(r \vee r, r) = 0$ are the root-vectors. Every pseudo-root vector is of the form

$$q = \pm\sqrt{3}\, r \vee r \quad , \tag{5.17}$$

and also satisfies

$$\sqrt{3}\, q \vee q = \mp q \quad . \tag{5.18}$$

We call it positive or negative (normalized) q-vector. We denote by f_x, d_x the linear mappings $a \overset{f_x}{\rightsquigarrow} x \wedge a$, $a \overset{d_x}{\rightsquigarrow} x \vee a$. Then

$$[f_a, f_b] = f_{a \wedge b}, \quad [f_a, d_b] = d_{a \wedge b} \quad , \tag{5.18}$$

so for $\forall a, b$ of a Cartan subalgebra C_x, the $f(a)$ can be diagonalized simultaneously on a basis z_k of the complexified E_8. Since C_x is left stable by f_a and d_a, we decompose $f_a = f_a'' \oplus f_a^{\perp}$, $d_a = d_a'' \oplus d_a^{\perp}$ on $C_x \oplus C_x^{\perp}$. Then

$$f_a'' = 0, \quad f_a^{\perp} z_k = i\,(r_k, a) z_k, \quad k = 1, \ldots, 6 \tag{5.20}$$

$$d_a^{\perp} z_k = (r_k \vee r_k, a) z_k = \frac{1}{\sqrt{3}} (q_k, a) z_k, \quad k = 1, \ldots, 6 \tag{5.20'}$$

where r_k are the six unit roots of C_x and $q_k = \sqrt{3}\, r_k \vee r_k$ are the three positive unit pseudo roots of C_x.

The two eigenvalues of d_a'' are $\pm 1/\sqrt{3}$.

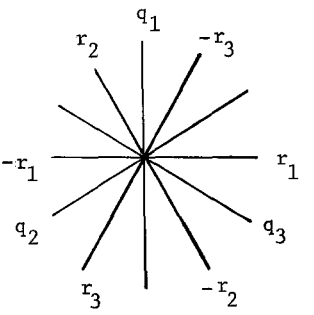


FIGURE 5.1. Roots $\pm r_i$ and pseudo-roots $q_i = \sqrt{3}\, r_i \vee r_i$ of a Cartan subalgebra. The SU(3)-Weyl group $S(3)$ permutes the three q_i.

Lemma

Every two-plane of E_8 contains at least a root. Indeed, the continuous odd function $(x, x \vee x)$ of x on the unit circle $(x,x) = 1$ of the two-plane has at least a zero. There are linear manifolds of root vectors.

For example: given a pseudo-root q, and using the same notation for a Lie subalgebra of $SU(3)$ and its vector space (subspace of E_8)

$$E_8 = U_q(1) \oplus SU_q(2) \oplus U_2(q)^{\perp} \quad ,$$

where the three- and four-dimensional $SU_q(2)$ and $U_2(q)^{\perp}$ spaces contain only root-vectors. An octet of particles form an orthonormal basis of the complexified E_8, which diagonalizes the f_a for all $a \in C_{(y,q)}$ the Cartan algebra generated by the hypercharge and the electric charge directions since Y, Q are generators of $U_y(2) \subset SU(3)$. The Gell-Mann-Nishijima relation

$$Q = T_3 + \frac{1}{2} Y \quad , \tag{5.21}$$

among generators of $U(2) \subset SU(3)$ is translated in the octet geometry; y, $-q$ are unit positive pseudo-roots, $Q = -2/\sqrt{3}\ F(q)$, $Y = 2/\sqrt{3}\ F(y)$,† t_3 is a root, $T_3 = F(t_3)$. We give in Figure 5.2 the corresponding roots of the two lowest octets of particles and also the weight of the lowest decuplet of baryons.

† The factors $2/\sqrt{3}$ are found from the condition that the spectra of Q and Y are the set of integers. Equation (5.21) implies that q and y are normalized pseudo-roots of opposite sign. The choice of sign here $+y$, $-q$ is conventional and corresponds to Figure 5.2.

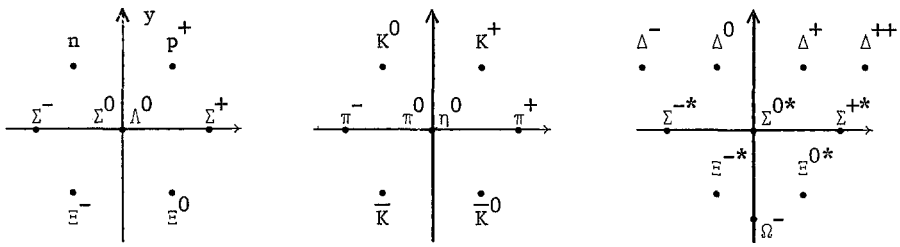


FIGURE 5.2. ROOTS OF OCTETS OF PARTICLES AND WEIGHT OF THE DECUPLET $\Sigma_0 = t_3 = \pi_0 \Lambda_0 = y = \eta_0$, corresponds to the two zero roots.

5.3. Electromagnetic and Weak Interactions in SU(3)

5.3.a. Electromagnetic Interaction

As shown by Equation (5.21), the electric charge operator Q is a generator of $U(2) \subset SU(3)$, so it is also a generator of $SU(3)$ and as we have seen

$$Q = \frac{2}{\sqrt{3}} F(-q) \quad , \tag{5.26}$$

where, as we have seen, q is a pseudo-root. The $SU_q(2)$ is called the U-spin group in the literature, and we can speak of U-spin multiplets, which have the same electric charge $u = 1/2$, p^+, Σ^+ and also Ξ^-, Σ^-, $u = 1$; η, Ξ^0, $1/2\ \Sigma^0 + \sqrt{3}/2\ \Lambda^0$; $u = 0$, $\sqrt{3}/2\ \Sigma^0 - 1/2\ \Lambda^0$. The electric charge is the integral of the time component of the electromagnetic current

$$Q = e\int j^0(x)d\vec{x} \quad , \tag{5.27}$$

and $\partial/\partial^\mu\ j^\mu(x) = 0 \Rightarrow Q$ is a constant (more generally P invariant) operator. Of course $j^0(x)$ could have any $SU(3)$ covariance, with the condition that the integral of the non-octet part vanish. The simplest hypothesis is to assume that the electromagnetic current $j^\mu(x)$ is the image in the direction $-q$ of an octet-tensor operator,

$$e\,\frac{2}{\sqrt{3}}\, j^\mu(x;-q) \tag{5.28}$$

(compare with Equation (5.26)). This allows us to draw many conclusions. The magnetic moment of the particle of a multiplet is given by the expectation value of an octet-tensor operator in the direction $-q$. It thus depends on two constants only for an octet (one for a decuplet) and the particles of the same u-spin multiplet have the same magnetic moment. For example one predicts $\mu_{\Sigma^+} = \mu_{p^+}$ which is well-verified. Measurements of μ_{Λ^0}, μ_{ε^+} and μ_{Ξ^-} are in progress, as well as the rate of $\Sigma^0 \to \Lambda^0 + \gamma$ which is related (as a "magnetic dipole" transition) to the values of μ in this octet. The ratio of rates of electromagnetic decay can be predicted. For example:

$$\frac{\text{rate } \pi^0 \to 2\gamma}{\text{rate } \eta^0 \to 2\gamma} = \frac{(t_3,q)^2}{(y,q)^2} \times \text{ ratio of phase-space} = 3 \times \text{ ratio of phase-space} \quad , \tag{5.29}$$

(using (5.26) and $(y,t_3) = 0$). The observed ratio $\phi \to \mu^+ + \mu^-$, $\omega \to \mu^+ + \mu^-$ is a good confirmation of the mixing angle. Finally ratios of photo production cross-sections can also be predicted successfully.

The mass differences inside a $U_y(2)$ multiplet are thought to be of electromagnetic origin. They are quadratic in $j^\mu(x;-q)$ but to a good approximation it seems that only the scalar and octet part are important, so to a good approximation the mass operator (5.5) can be written, when one adds electromagnetic effects,

$$M = M_0 + M_1 \frac{2}{\sqrt{3}} F(y) + M_2 D(y) + M_3 \frac{2}{\sqrt{3}} F(-q) + M_4 D(-q)$$

and inside an SU(3)-multiplet the values of the masses are given by

$$m = m'_0 + m'_1 y + m'_2 (t(t+1) - \tfrac{1}{4} y^2) + m'_3 q + m'_4 (u(u+1) - \tfrac{1}{4} q^2) \tag{5.30}$$

which is well verified for baryons.

5.3.b. Weak Interaction

Cabibbo generalized to SU(3) the Gell-Mann Feynmann hypothesis on the vector part of the weak current $v_\mu^\pm(x)$ coupled to the leptonic current $\ell^{\mp\mu}(x)$ (see 2) by the assumption that $v_\mu^\pm(x)$ and $j_\mu(x)$ are images by the same octet-tensor operator (that we shall denote $v_\mu(x)$) of three different directions: $-q, c_\pm$. Explicitly

$$\begin{aligned} \text{em current} &= \frac{2}{\sqrt{3}} e v_\mu(x;-q) \quad , \\ \text{weak current} &= \frac{G}{\sqrt{2}} v_\mu(x;c_\pm) \quad , \end{aligned} \tag{5.31}$$

(where G is the Fermi constant). The second Cabibbo assumption is that the axial-vector parts of the weak current $a_\mu^\pm(x)$ are images of another octet-tensor operator, in the same direction $c_\pm$. The total weak current

$$h_\mu^\pm(x;c_\pm) = v_\mu^\pm(x;c_\pm) - a_\mu^\pm(x;c_\pm) \quad , \tag{5.32}$$

is thus also image by an octet-tensor operator. See Cabibbo's original paper (*Phys. Rev. Lett.*, 10, 531 (1963)) in *The Eightfold Way* anthology (p.207) for the predictions.

The $\pm$ subscript corresponds to the electric charge of the current, i.e.,

$$[Q, h_\mu^\pm(x)] = \pm h_\mu^\pm(x) \quad , \tag{5.33}$$

and using the fact that Q is an SU(3) generator, $Q = 2/\sqrt{3}\, F(-q)$, we can write this equation in the form (1.9)

$$-\frac{2}{\sqrt{3}}\,[F(q),h_\mu(x,c^\pm)] = -\frac{2}{\sqrt{3}}\,h_\mu(x,q\wedge c^\pm) = \pm h_\mu(x,c_\pm) \quad , \tag{5.34}$$

from (5.34) we get

$$q\wedge c_\pm = \mp\frac{\sqrt{3}}{2}\,c_\pm \tag{5.34'}$$

which means that $c_\pm$ are eigenvectors of $F(q)$. Writing $c_\pm = 1/\sqrt{2}(c_1 \pm ic_2)$ Equations (5.31, 5.34') imply that c_1 and c_2 are unit vectors $\in U_q(2)$, so they are root-vectors, as we have seen in (5.21). Equation (5.34') is equivalent to $q\wedge c_1 = c_2$, $q\wedge c_2 = -c_1$ which in turn implies

$$\sqrt{3}\,c_1\vee c_1 = \sqrt{3}\,c_2\vee c_2 = \sqrt{3}\,c_3\vee c_3 = c \quad , \tag{5.35}$$

where

$$c_3 = c_1\wedge c_2 \quad . \tag{5.35'}$$

This means that c, c_1, c_2, c_3 span $U_c(2)$; note also that c, $c_3 \in U_q(2)$. The pseudo root c is called the "weak hypercharge" or "Cabibbo hypercharge". It is a conserved quantity for weak interactions. It commutes with q, $c\wedge q = 0 \Rightarrow (c,q) = -1/2$. However, it does not commute with y; indeed, there are weak transitions violating hypercharge conservation. This lack of commutation is expressed by the $\neq -1/2$ value of

$$(y,c) = 1 - \frac{3}{2}\sin^2\theta \quad , \tag{5.36}$$

where θ is the Cabibbo angle. As we have seen in 3.6, its experimental value is 15 degrees and it is rather well verified that $v_\mu^\pm$ and $a_\mu^\pm$ define the same direction c of weak hypercharge.† The value of this angle is empirically given by

$$\mathrm{tg}\theta = m_\pi/m_k \quad ,$$

Cabibbo's theory not only explained the relative slower rate†† (by $\mathrm{tg}^2\theta$) of the weak transition violating the hypercharge y, but also explained that the super allowed $\Delta T = 0$ nuclear β-decay were slower than the $\mu \to e + \nu + \bar{\nu}$ decay by a ratio $\cos^2\theta$.

The "computation" of this angle θ is one of the challenging present problems of physics. It is worth while to point out a purely algebraical relation, giving q as function of y and c.

Given two non-commuting (normalized positive) pseudo-roots y and c, there is always a unique pseudo-root which commutes with both of them

$$\lambda q = \sqrt{3}\,y\vee c + \tfrac{1}{2}(y+c) \quad , \tag{5.37}$$

---

† To be more precise, the angle of c_v and c_a with y is the same but c_v and c_a could be at a small angle between each other and this has been exploited as a possible explanation of CP violation.

†† To be accurate, it is not the rate but the probability transition = rate/phase space volume, since the phase space volumes, which should be equal in an exact SU(3)-symmetry, are in fact unequal.

where

$$\lambda = -(1 - (y,c)) \quad . \tag{5.37'}$$

The most commonly proposed form of non-leptonic weak interaction is

$$H_{N.L.} = \frac{G}{\sqrt{2}} \sum_{\varepsilon=\pm 1} \int h^{\mu}(x,c_{\varepsilon}) h_{\mu}(x,c_{-\varepsilon}) d^{3\vec{x}} \quad , \tag{5.38}$$

with the drawback that $H_{N.L.}$ is the image of a reducible tensor operator with some component in the "27" irrep of SU(3). The $\vec{\Delta T} = 1/2$ rule when $|\Delta Y| = 1$ for those weak transitions suggests that this 27 component is negligible compared to the octet component. The proposal of Radicati[†]

$$H_{N.L.} = \frac{G}{\sqrt{2}} \int (h^{\mu}(x) \vee h_{\mu}(x))(c) d^{3\vec{x}} \quad , \tag{5.39}$$

makes $H_{N.L.}$ the component along the weak hypercharge c of an irreducible octet-tensor operator. It is compatible with the known experimental data.

5.4. Critical Orbits of a G-Invariant Function on a Manifold M[††]

Given a group G acting on a set M, the set of all points of M which have conjugated little groups is called a stratum. So a stratum is the union of all orbits of the same type. Inclusion gives a partial ordering of all subgroups, modulo a conjugation, of a group. It corresponds to an (inverse) ordering on the strata. The set of fixed points form the minimal stratum (maximal isotropy group = G). If in the action of G on M there are no fixed points, there might be several minimal strata.

For example, in 5.2 we have seen that in the action of SU(3) on the unit sphere S_j of the octet space, there is the open dense general stratum $|(x \vee x,x)| < 1/\sqrt{3}$, containing a one parameter family of six-dimensional orbits (little group $U(1) \times U(1)$ and a minimal stratum made of two four-dimensional orbit $(x \vee x,x) = \pm 1/\sqrt{3}$. In this paragraph we want to consider

a) the smooth[†††] action of a compact Lie group G on a smooth manifold M. This action is given by the smooth mapping (= manifold morphism)

$$G \times M \xrightarrow{\phi} M \quad \text{with} \quad \phi(g_1,\phi(g_2,m)) = \phi(g_1 g_2,m) \quad ,$$

b) a real smooth function $M \xrightarrow{f} R$ which is G invariant, that is, the function is constant on every G orbit of M

---

† L. Radicati in *Old and New Problems in Elementary Particle Physics*, Academic Press, New York (1968).

†† This part is entirely a common work with Radicati, partly published in *Coral Gables Conferences 1968*, partly circulated in a preprint.

††† We use the word smooth for infinitely differentiable.

$g \in G$, $m \in M$, $f(\phi(g,m)) = f(m)$.

The differential at $m_1 \in M_1$ of a smooth mapping $M_1 \overset{\Psi}{\to} M_2$ is denoted $d\Psi_{m_1}$; it is a linear mapping (with $m_2 = \Psi(m_1)$).

$$T_{m_1}(M_1) \xrightarrow{d\Psi_{m_1}} T_{m_2}(M_2) \tag{5.40}$$

where $T_{m_i}(M_i)$ is the tangent plane of M_i at m_i. So $df_p \in T'_p(m)$ the dual vector space of $T_p(m)$. We call critical point, the $p \in M$ such that $df_p = 0$.

The stabilizer (= little group = isotropy group) G_m in $m \in M$ is a closed and therefore compact subgroup of the compact group G. As is well known,† one can choose local coordinates in a neighborhood V_p of p such that the action of G_p is linear. Let $E_p(M)$ be the vector space corresponding to this linear representation of G_p; so $V_p \subset E_p(M)$. Since G_p is compact and M real, this linear action can be made orthogonal so $E_p(M)$ is a euclidean space. We can then identify df_p with a vector of $E_p(M)$ that we shall call $(\text{grad } f)_p$. The G-orbit of p, $G(p)$, is the image of $g \overset{\phi(p)}{\rightsquigarrow} \phi(g,p)$; it is a submanifold of M; its tangent plane in p, denoted $T_p(G(p))$, is the image of $d\phi_e^{(p)}$ where e is the unit of G. The isotropy group G_p transforms $G(p)$ into itself. Similarly $T_p(G(p))$ is an invariant subspace of $E_p(M)$. The orthogonal subspace $N_p(G(p)) = T_p(G(p))^{\perp} \subset E_p(M)$ is also invariant and it is called the "slice" at p. Note that $(\text{grad } f)_p \in N_p$. Indeed, by definition, for $x \in T_p(M)$, $((\text{grad } f)_p, x) = \lim_{\alpha \to 0}[(f(p + \alpha x) - f(p))]\alpha^{-1}$. The bracket is 0 when $p + \alpha x \in G(p)$, the orbit of p, so it stays zero at the limit, when $x \in T_p(G(p))$.

Note also that $(\text{grad } f)_p$ is invariant by G_p. Let $g \in G_p$; $(g \cdot (\text{grad } f)_p, x) = ((\text{grad } f)_p, g^{-1} \cdot x) = \lim_{\alpha \to 0} \alpha^{-1}(f(p + \alpha g^{-1} \cdot x) - f(x))$, and since $g^{-1} \cdot p = p$, $f(p + \alpha g^{-1} \cdot x) = f(g^{-1} \cdot (p + \alpha x)) = f(p + \alpha x)$, so $\forall x \in E_p(M)$, $(g \cdot (\text{grad } f)_p, x) = ((\text{grad } f)_p; x)$. If the slice $N_p(G(p))$ has no vectors invariant by G_p, then $(\text{grad } f)_p = 0$. We can summarize this by the:

Theorem 1

Let G be a compact Lie group acting smoothly on the smooth real manifold M. If for $p \in M$, the canonical linear representation of G_p on the slice N_p does not contain the trivial representation of G_p, then $G(p)$ is a critical orbit for

† Consider a Riemann metric on M; it is transformed by the action of G_p. By averaging with a G_p-invariant measure, one obtains a G_p-invariant Riemann metric and G_p transforms into each other the geodesics from the fixed point p. In the neighborhood V_p of p, take geodesic coordinates.

any real valued G-invariant smooth function on M (where here again we denote by the same symbol, e.g., SU(2), the vector space of the Lie subalgebra, and also the group!).

Example 1. We have studied the action of SU(3) on $S_7 \subset E_8$. Let q be a unit q-vector, $G_q = U_2(q)$, $T_q(M) = \{q\}^{\perp} \subset E_8$, $T_q(G(q)) = U_2(q)^{\perp}$, $N_q(G(q)) = SU_2(q)$ and $U_2(q)$ acts linearly on it, without fixed vectors.

Example 2. p is an isolated fixed point in M. So there is a neighborhood V_p of p with no other fixed points and $N_p = E_p(M)$ has no invariant $G = G_p$ vector.

This proves that p is a critical point for every G invariant function on M.

We shall now assume moreover, that M is compact. Then there is one stratum (called generic stratum) which is open dense in M; the minimal strata are closed and compact. Let C be a connected component of a minimal stratum; let $p \in C$, $F_p \subset E_p$ be the linear subspace of G_p fixed points. Because G_p is maximal, the points of $V_p \cap F_p$ have G_p as stabilizer so they belong to C. Given a G-invariant real valued smooth function f, let $n = (\text{grad } f)_p$. As we have seen $n \in F_p$ so for small enough $|\varepsilon|$, $p + \varepsilon n \in C$. We can write

$$(n,n) = \lim_{\varepsilon \to 0} \varepsilon^{-1}(f(p + \varepsilon n) - f(p)) \tag{5.41}$$

so if f is constant on C, every $p \in C$ is a critical point of f. If f is not constant on C it has at least an orbit of maxima and an orbit of a minima. Let p a point of such an orbit, and $n = (\text{grad } f)_p$. Then, in Equation (5.41),

$$f(p + \varepsilon n) - f(p) \begin{matrix} \geq 0 & \text{if } f \text{ is minimum} \\ \leq 0 & \text{if } f \text{ is maximum} \end{matrix} \text{ at } p \ ,$$

which means that (n,n) either has the sign of $\pm\varepsilon$ (+ at minimum, - at maximum) which is impossible, or must be zero.

Theorem 2.†

Let G be a compact Lie group, acting smoothly on the real compact manifold M, and let f be a real valued G-invariant, smooth function on M. Then f has at least a critical point for each connected component C of each minimal stratum.

† To prove this theorem, that Radicati and I conjectured, we received great help from A. Borel, C. Moore, and R. Thom.

We will now be interested in a particular function on a sphere: let G be a compact Lie group, E the real vector space carrier of a linear representation $g \rightsquigarrow R(g)$, irreducible over the reals. So R (up to an equivalence) is an orthogonal representation and it is self-contragredient. We denote by $(\underline{x},\underline{y})$ the invariant Euclidean scalar product in E. Let us assume that (with $\vee$ the symmetrical tensor product) $\dim(\mathrm{Hom}\ E \vee E,E)^G = 1$. As we have seen in 1.5, there is a unique (up to a constant factor) symmetrical algebra

$$x \otimes y \xrightarrow{\Psi} x_T y \quad \text{where} \quad \Psi \in \mathrm{Hom}(E \vee E,E)^G \tag{5.42}$$

with $x_T y = y_T x$.

Since the representation is self-contragredient and the tensor product is associative

$$(x_T y,z) = (x,y_T z) = \{x,y,z\} \quad . \tag{5.43}$$

Hence, the invariant $\{x,y,z\}$ is a completely symmetrical G-invariant trilinear form on E. Let $f(\{x,y,z\})$ be a function on the unit sphere $S = \{x \in E,(x,x) = 1\}$.

Using λ as a Lagrange multiplier, critical points of f are given by the equation

$$\mathrm{grad}(f(\{x,x,x\}) + \lambda(1 + (x,x)) = 3f'x_T x - 2\lambda x = 0 \quad , \tag{5.44}$$

where f' is the derivative of the one variable function f; e.g., if $f = \{x,x,x\}$, $f' = 1$. In other words, critical points of f are given by solutions of

$$x_T x = \lambda x \quad ,$$

i.e., the idempotents (or nilpotents for $\lambda = 0$) of the symmetrical algebra.

5.5. SU(3) × SU(3) Symmetry

Physicists have considered symmetries higher than SU(3) for the hadronic world. Of course they are coarser, but still useful as we shall see. The SU(3) × SU(3) symmetry becomes an exact symmetry of the hadronic world when the masses of 0^- mesons are neglected. Note that it is not much more drastic to say that those masses are equal to zero than to say that they are equal as is already implied by SU(3). As a matter of fact, a much milder approximation than SU(3) is to neglect only the π-meson mass (only 140 MeV, and this is smaller than the 0^--meson mass differences). This corresponds to a SU(2) × SU(2) × U(1) subgroup of SU(3) × SU(3). We give in Figure 5.3, a scheme of the lattice of symmetry groups which have been considered for hadronic physics, but in this section we limit ourselves to SU(3) × SU(3) and its subgroup. (See also O'Raifeartaigh lectures for the higher symmetries.) At the level of the middle line of Figure 5.3, a new feature appears; a mixing of internal symmetry and relativity invariance. It is very mild for SU(3) × SU(3) since it concerns only the parity operator. The total symmetry group to consider is the semi-direct product by $Z_2(P)$

$$(P_0 \times SU_3 \times SU_3) \square Z_2(P) \quad , \tag{5.49}$$

which acts naturally on P_0 and exchanges the two SU_3 factors in $SU_3 \times SU_3$. To distinguish such SU(3) factors, let us denote them as $SU_3^{(+)} \times SU_3^{(-)}$; they are called in the physics literature the $\pm$ chirality group. The group (5.49) is a good frame for understanding the relation of P (parity operator) with the different interactions. This will become clear in the following. The diagonal subgroup $SU(3)^d \subset SU_3^{(+)} \times SU_3^{(-)}$ is the SU(3) group of invariance of 5.1, 5.2, and 5.3.

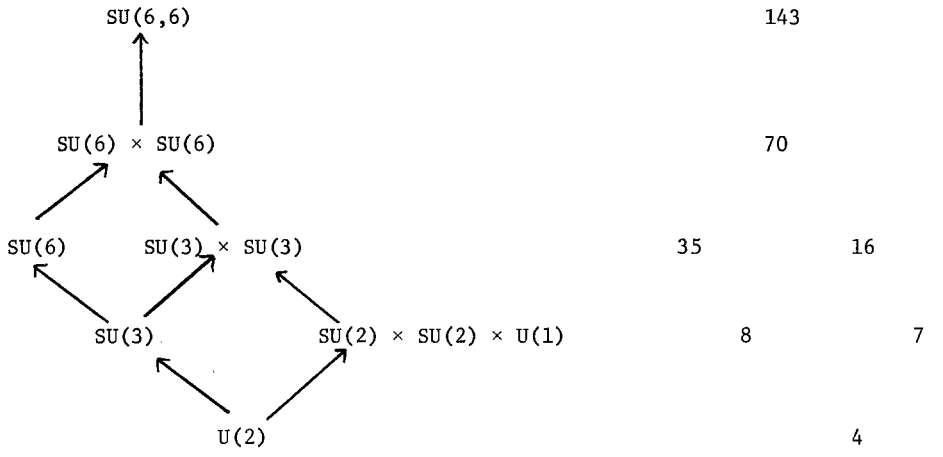


FIGURE 5.3. SYMMETRY GROUPS AND THEIR DIMENSIONS

Lattice of symmetry groups used in hadronic physics. $\rightarrow$ means injection as subgroup.

We will denote a vector of the 16-dimensional vector space E_{16} of the SU(3) × SU(3) Lie algebra by a direct sum of two vectors

$$\tilde{a} = a_+ \oplus a_- \quad , \tag{5.50}$$

$a_\pm$ belongs to the $SU_3^{(\pm)}$ octet.

The invariant Euclidean scalar product (given by the Cartan-Killing form) is, in terms of the octet scalar product

$$(\tilde{a},\tilde{b}) = (a_+ \oplus a_-, b_+ \oplus b_-) = \tfrac{1}{2}(a_+,b_+) + \tfrac{1}{2}(a_-,b_-) \quad . \tag{5.51}$$

The Lie algebra law is (we use $\wedge$ for it)

$$\tilde{a} \wedge \tilde{b} = (a_+ \wedge b_+) \oplus (a_- \wedge b_-) \quad ,$$

and since $\dim \mathrm{Hom}(E_{16} \vee E_{16}, E_{16})^{SU3 \times SU3} = 1$, there is a unique canonical symmetrical symmetrical algebra

$$\tilde{a} \vee \tilde{b} = (a_+ \vee b_+) \oplus (a_- \vee b_-) \quad . \tag{5.52}$$

The covariance property of the electromagnetic and weak interactions are most naturally extended to $SU(3) \times SU(3)$ by the following hypothesis: the electric current $j_\mu(x)$, the vector part $v_\mu^{(\varepsilon)}(x)$, the axial vector part $a_\mu^{(\varepsilon)}(x)$ of the (charged $\varepsilon = \pm 1$) weak hadronic current $h_\mu^{(\varepsilon)}(x) = v_\mu^{(\varepsilon)}(x) - a_\mu^{(\varepsilon)}(x)$ are images of the same E_{16} tensor operator, which we will denote $h_\mu(x;\tilde{a})$; the vector currents correspond to $SU(3)^d$ and the axial vector current to the anti-diagonal. The weak current has pure "-" chirality. Explicitly: electromagnetic current

$$\frac{2}{\sqrt{3}} e h_\mu(x;-(q \oplus q)) \quad , \tag{5.53}$$

(charged) weak currents

$$\frac{G}{\sqrt{2}}(h_\mu(x;0 \oplus c_1) \pm i h_\mu(x;0 \oplus c_2)) \tag{5.53'}$$

and the Radicati form of the (non leptonic) weak hadronic interaction is

$$\frac{G^2}{2}\int h_\mu(x) \vee h^\mu(x))(c)d^3\vec{x} = \frac{G^2}{2}\int (h_\mu(x) \vee h^\mu(x)(0 \oplus c)d^3x \quad . \tag{5.54}$$

The generators of $SU(3) \times SU(3)$ are the space integral of the current, i.e.,

$$a \rightsquigarrow F(\tilde{a}) = \int h_0(x;\tilde{a})d^3\vec{x} \quad , \tag{5.55}$$

is the representation (up to i) of the $SU(3) \times SU(3)$ Lie algebra on the Hilbert space of physics

$$[F(\tilde{a}),F(\tilde{b})] = iF(\tilde{a} \wedge \tilde{b}) \quad , \tag{5.56}$$

and for the particular case of the E_{16}-tensor operator $(h_\mu(x)\tilde{a})$

$$[F(\tilde{a}),h_\mu(x,\tilde{b})] = ih_\mu(x,\tilde{a} \wedge \tilde{b}) \tag{5.57}$$

as we saw in Equation (1.9).

In the approximation where $SU(3) \times SU(3)$ is an exact symmetry $\partial_\mu h^\mu(x,\tilde{a}) = 0$ and the $F(\tilde{a})$ are well defined. Since $SU(3) \times SU(3)$ is a broken symmetry, the usual difficulty to define the self-adjoint operator $F(\tilde{a})$ arises. (See O'Raifertaigh's lectures, this Volume.)

The equation[†]

$$\tilde{a} \vee \tilde{a} = \lambda \tilde{a} \quad , \tag{5.58}$$

for unit vectors $\in S_{15} \subset E_{16}$ has two sets of solutions. One is the set of $1/\sqrt{2}\ \tilde{a} = \pm c \oplus 0$ or $\pm 0 \oplus c$, where c is a (normalized positive) pseudo-root and $\lambda = \mp\sqrt{2/3}$. This set is made up of two minimal strata, each consisting of two pieces of one orbit each. So each of the four orbits is a critical orbit for <u>every</u> smooth $SU(3) \times SU(3)$ invariant function in S_{15} the unit sphere of E_{16}. The stabilizers are, up to a conjugation, $SU_3^{(+)} \times U_c^{(-)}(2)$ and $U_c^{(+)}(2) \times SU^{(-)}(3)$ for the two strata.

[†] See L. Michel and L. Radicati, preprint, *Breaking of the SU(3) × SU(3) Symmetry in Hadronic Physics.*

The other type of solution is the set of vectors,

$$\pm(q_1 \oplus q_2) \quad ,$$

which form two orbits of a four separated orbit stratum ($\pm q_1 \oplus \mp q_2$ for the two other orbits) whose stabilizer is $(U_{q_1}(2) \times U_{q_1}(2))_{\square} Z_2$. The pseudo-roots $\pm(q \oplus q)$ of the diagonal $SU^{(d)}(3)$ are on the orbits of solutions while those of the anti-diagonal $(\pm q \oplus \mp q)$ are not. This has a bearing on parity.

It seems to us remarkable that the electromagnetic charge direction $-(q \oplus q)$ and the weak hypercharge direction $(0 \oplus c)$ give two solutions, one of each type, of Equation (5.58).

$SU(3) \times SU(3)$ is not only broken by electromagnetic and weak interaction, but also by semi-strong U_2-invariant interaction. There are two different interesting intermediate approximations of symmetry of strong interactions between U_2 and $SU(3) \times SU(3)$; those of the fourth line of Figure 5.3, $SU(3)$ already studied, and $SU(2) \times SU(2) \times U_1$, which implies the Adler-Weissberger sum rule, and more recently emphasized by Gell-Mann, Oakes and Renner. In both cases H_{strong} is, to a good approximation, the sum

$$H_{strong} = H_0 + H_1(\underset{\sim}{m}) \quad , \tag{5.59}$$

of H_0 invariant under $SU(3) \times SU(3)$ and of $H_1(\underset{\sim}{m})$ which is the image of $\underset{\sim}{m}$ by a $SU(3) \times SU(3)$ tensor operator for the (irreducible over reals) $(3,\bar{3}) \oplus (\bar{3},3)$ representation. The two corresponding directions $\underset{\sim}{m}$ for these two approximations are again idempotent or nilpotents of the canonical symmetric algebra. I refer to my preprint with Radicati for details. This 18-dimensional irreducible real representation of $SU(3) \times SU(3)$ on E_{18} (which is the one which naturally arises in a quark model) is such that $\dim \mathrm{Hom}(E_{18} \vee E_{18}, E_{18})^{SU(3) \times SU(3)} = 1$ so there is a unique canonical symmetrical (real) algebra on E_{18} which has $SU(3) \times SU(3)$ as group of automorphisms. We denote this algebra law by $\underset{\sim}{m}_1 \top \underset{\sim}{m}_2$.

The equation

$$\underset{\sim}{m} \top \underset{\sim}{m} = \lambda m \quad , \tag{5.60}$$

has only two types of solutions (for vectors on the invariant unit sphere $S_{17} \subset E_{18}$) belonging to two minimal strata, the one for $|\lambda| = 2/3$ corresponds to $SU^{d}(3)$ as stabilizer. The other, for $\lambda = 0$ corresponds to $SU_y^{(+)}(2) \times SU_y^{-}(2) \times U_y^{d}(1)$.

Theorem 1 shows that this latter case $(\lambda = 0)$ corresponds to a critical orbit for all $SU(3) \times SU(3)$ invariant functions on S_{17} (unit vectors of the $3\bar{3} + \bar{3}3$ irrep); this orbit is also a minimal stratum of dimension 9. The stratum corresponding to $SU(3)^{d}$ is also minimal; it is a nine-dimensional connected submanifold (of S_{17}) made up of eight-dimensional orbits. So from Theorem 2, each invariant function has at least two critical orbits in this stratum. For all functions of $(x, x \top x)$ these two orbits are $x \top x = \pm 2/3\, x$.

Note Added After the Seattle Rencontres.† I do not understand why I have not used in Seattle, as emphasized by Equation (5.49), $(SU(3) \times SU(3))_{\square} Z_2$ instead of $SU(3) \times SU(3)$. Then, the two orbits on S_{17}, $x \mp x = \pm 2/3\, x$ are critical for all functions. Radicati and I also wonder why we have not considered before the groups $(SU(3) \times SU(3))_{\square}(Z_2 \times Z_2)$ where the discrete group $Z_2 \times Z_2 = \{I,P,C,PC\}$ is generated by the parity and the charge conjugation operators. Among the strata of S_{15} for this group, there are four which contain only one orbit. These orbits are the critical ones of S_{15}. Typical points ($\simeq$ unit vectors up to a sign) of these orbits are

$\pm\tilde{q} = \pm(q \oplus q)$	the direction of electromagnetic interaction,	(5.53)
$\pm\tilde{c}_i = 0 \oplus \pm c_i (i = 1,2)$	the Cabibbo direction of weak coupling,	(5.53′)
$\pm\tilde{c} = 0 \oplus \pm c$	the "weak hypercharge" direction proposed by Radicati,	(5.54)
$\pm\tilde{r} = \pm(r \oplus \varepsilon r)$, $\varepsilon = \pm 1$ (root vectors)	a direction which some authors (for instance M. L. Good, L. Michel, and E. de Rafael, *Phys. Rev.*, 151, 1199 (1966), have used in their proposed theory of the CP violating interaction.	

Radicati and I have also included Theorem 1 into a more complete:

Theorem 1′

Let G be a compact Lie group acting smoothly on the real manifold M, $p \in M$. The three following propositions are equivalent.

a) the orbit of p is critical (for every G-invariant real valued smooth function f on M, $df_p = 0$),

b) the orbit of p is isolated in its stratum, i.e., $\exists$ a neighborhood V_p of p such that $(x \in V_p$ and $x \notin G_p) \Rightarrow G_x$ is not conjugate to G_p,

c) the canonical linear representation of G_p on the slice N_p does not contain the trivial representation.

Theorem 1 is simply $c \Rightarrow a$.

5.6. SU(6), Quarks, Current Algebra, Boot-Strap, Etc.

The title of this section is a statistical sample of key words found these last years in papers on fundamental particle physics.†† This last section is not a

† After the Seattle Recontre, L. Radicati and I collected the above results to present them in a lecture on September 19, 1969 in Rome (see preprint, *Geometrical Properties of the Fundamental Interactions*). The following improvements were then obvious to us.

†† For the last year, the passwords are Veneziano and duality. It is a sociological fact that there are fads in fundamental particle physics.

conclusion but an open-end to the description of a very rapidly changing situation; the view that physics gives us of the hadron world.

SU(6) Symmetry. SU(6) Symmetry was introduced independently† by Gürsey and Radicati (*Phys. Rev. Lett.*, 13, 299 (1964)) and by B. Sakita (*Phys. Rev.*, 136 B, 1756 (1964), for mesons only). It was noticeable that mass-differences between SU(3) multiplets were not larger than those inside multiplets.

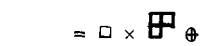
Both groups of authors, inspired by the SU(4) = supermultiplet Wigner theory for nuclei (3.3) extended it to fundamental particles by enlarging the SU(2) isospin to SU(3). So in the non-relativistic version, the space of the one particle hadron states is the tensor product

$$\mathcal{H}^{(1)} = L_2(R^3,t) \otimes K_\sigma \otimes K_\lambda \quad ;$$

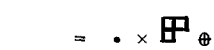
here K_σ and K_λ are respectively, two- and three-dimensional Hilbert spaces and the action of $\overline{G}$, the central extension of the Galilee group, and of SU(6) on $\mathcal{H}^{(1)}$ are respectively, (with $\overline{G} \overset{\Psi}{\to} SU(2)$ also (2.9) and Equations (2.57) and (2.58)††

$$\begin{aligned} \mathcal{H}^{(1)} &= L_2(R^3,t) \otimes \quad K_\sigma \otimes K_\lambda \\ &\overline{g} \in \overline{G} \to \pi(\overline{g}) \otimes \Psi(\overline{g}) \otimes I \\ &u \in SU(6) \to I \otimes u \end{aligned} \tag{5.61}$$

The lowest two multiplets of SU(6) are given in Figure 5.4. For the baryon, it belongs to the irrep  of dimension 56; for the meson, to the  , the 35-dimensional adjoint irrep of SU(6). The X^0 (not discovered in 1964!) is a singlet. We give here the decomposition of these irrep into $SU_2 \times SU_3$ irrep

 $\Big|_{SU_2 \times SU_3}$ = 

$2 \times 8 + 4 \times 10 = 56$

 $\Big|_{SU_2 \times SU_3}$ = 

$(1 \times 8) + (3 \times 8) + (3 \times 1) = 35$

The mass formula for each SU(6) multiplet becomes

$$m = m_0 + m_1 y + m_2(t(t+1) - \tfrac{1}{4} y^2) + m_3 j(j+1) + m_4 q + m_5(u(u+1) - \tfrac{1}{4} q^2) \quad .$$

† In fact, Gell-Mann in, *Physics*, 1, 63 (1964), page 74 (reproduced in, *The Eightfold Way*, anthology, p. 203), was the first to introduce SU(6) in the physics of elementary particles but, for once, he did not work out its physical applications.

†† For more details, see L. Michel, "The Problem of Group Extensions of the Poincaré Group and SU(6) Symmetry", p. 331; 2nd Coral Gables Conferences, *Symmetry Principles at High Energy*, Freeman and Co., San Francisco (1965).

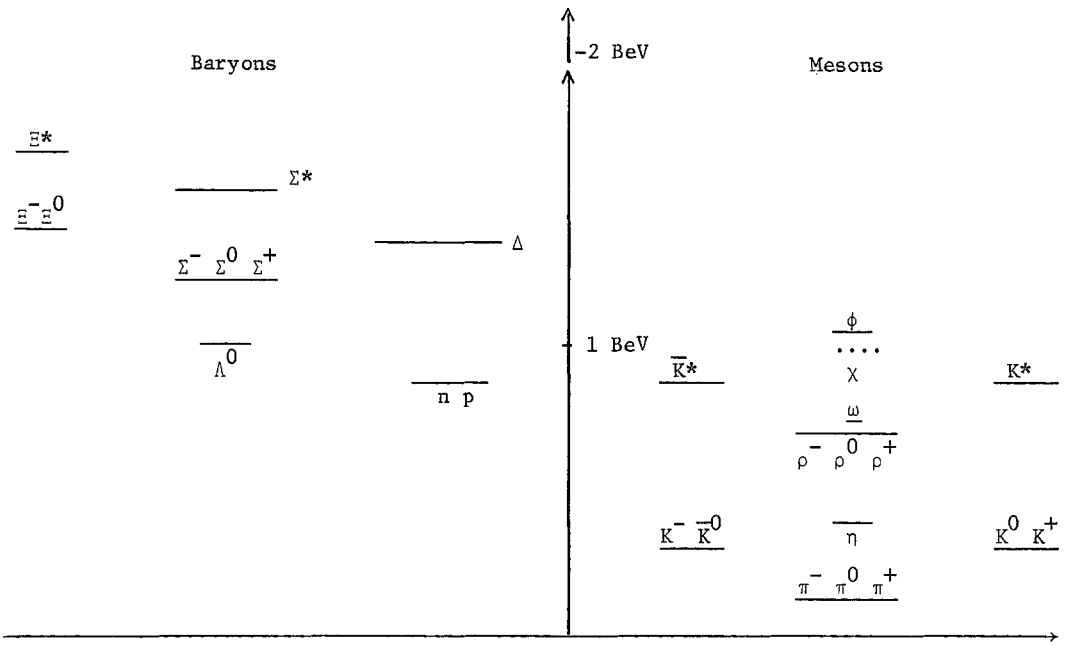


FIGURE 5.4. THE $(8 \times 2) + (10 \times 4) = 56$-PLET OF BARYONS AND THE $(8 \times 1) + (9 \times 3) = 35$-PLET OF MESONS IN THE SU-6 CLASSIFICATION OF HADRONS.

Neglecting the electromagnetic mass difference $(m_4 = m_5 = 0)$ the four-parameter formula predicts well the masses of the eight lowest $U_{(2)}$- multiplets of baryons. The magnetic moment of baryons depends on only one parameter μ_p so we have the relation

$$\mu_n = -\frac{2}{3}\mu_p \quad , \tag{5.62}$$

which is within 3 percent of reality (this is too good!).

How should one apply SU(6)' invariance to particle reactions? Some physics and empirical rules (e.g., so called $SU(6)_w$) have to be injected, and the symmetry is still useful.

However, the drawback is the difficulty in reconciling SU(6) with relativistic invariance.†

Quarks. It is a natural tendency in science to try to explain the universe by the smallest number of different types of building blocks, such as the four elements of the Greeks, which at the end of the XIXth century had reached nearly ninety chemical elements. From 1910 to 1929 (measurement of the spin and statistics

† This will be dealt with by O'Raifeartaigh, when he studies the two upper lines of the diagram of Figure 5.3.

of the N^{14} nucleus, see 2.10) only three particles p^+, e^-, γ, were known and needed to build the universe again. But one had to add ν in 1931, n and e^+ in 1932, etc., so now we have the Table 3.2 of 3.5 = spectroscopy of hadrons.

Is it possible to return to "simplicity"? The hoped for building blocks have been called quarks by Gell-Mann: the 3 spin 1/2 quarks for the multiplet 3 (= fundamental irrep $\square$) of SU(3) and 6 (= irrep $\square$) of SU(6). There are also 3 antiquarks belonging to the contragredient irrep

$$\overline{3} = \begin{array}{|c|}\hline\\\hline\\\hline\end{array} \text{ of SU(3)} \quad \text{or} \quad \overline{6} = \begin{array}{|c|}\hline\\\hline\\\hline\\\hline\\\hline\end{array} \text{ of SU(6)} \quad .$$

Mesons of Table 3.2 are formed of one quark and one antiquark $\overline{q}$. Lowest bound states of $q + \overline{q}$ yield all expected meson states. Baryons of the same table are made of 3 quarks, which are, for the lowest state, in the SU(6) state $\square\square\square$, so they must have a space symmetry $\begin{array}{|c|}\hline\\\hline\\\hline\end{array}$ to obey Fermi statistics; this from our experience acquired in Chapter 2 and 3 does not seem compatible with attractive forces. And how to explain the saturation by 3; why should 2-quark or 4- or 5-quark states not also be stable?†

Forgetting these difficulties one can search for quarks. (They should be very heavy, stable, have fractional quantum numbers $b = 1/3$, $q = 2/3$ or $-1/3$) and compute with them (good prediction of the "quarks model", e.g., by Dalitz, Lipkin.) They have not been found experimentally, and quarks can simply be looked at as the physicists' name for an orthonormal basis of the fundamental $\square$ irrep of SU(6), used in their computations!

Current Algebra. Let $\tilde{a} \rightsquigarrow D(\tilde{a})$ be the SU(3) × SU(3) Lie algebra adjoint irrep E_{16}. Any E-tensor operator function on space time $f(\vec{y},\underset{\sim}{m})$ will satisfy Equation (1.9) at any fixed time

$$[F(\tilde{a}),f(\vec{y},\underset{\sim}{m})] = if(\vec{y},D(\tilde{a})\underset{\sim}{m}) \quad , \tag{5.63}$$

where $\underset{\sim}{m} \in E$. Equation (5.57) is a particular case for $f(\vec{x},\underset{\sim}{m}) = h^\mu(\vec{x},\tilde{b})$.

Replace $F(\tilde{a})$ by its expression (5.55). After commuting the symbols $[$ and $\int$, Equation (5.63) reads (use $(1 = \int \delta(\vec{x} - \vec{y})d^3\vec{x})$),

$$\int d^3\vec{x}[h^0(\vec{x},\tilde{a}),f(\vec{y},\underset{\sim}{m})] = i\int d^3\vec{x}\delta(\vec{x} - \vec{y})f(\vec{y},D(\tilde{a})\underset{\sim}{m}) \quad ,$$

for any tensor operator function of $\vec{x}$. It is very suggestive to write the equality for the integrands

$$[h^0(\vec{x},\tilde{a}),f(\vec{y},\underset{\sim}{m})] = i\delta(\vec{x} - \vec{y})f(\vec{x},D(\tilde{a})\underset{\sim}{m}) \quad . \tag{5.64}$$

† There are several ways out of these difficulties, but the most efficient seems to me that of O. W. Greenberg and collaborators who have introduced three types of 3(q and $\overline{q}$). They obtain a remarkable hadronic spectrum.

Equation (5.56) is written in this local form

$$[h^0(\vec{x},\tilde{a}),h^\mu(\vec{y},\tilde{b})] = i\delta(\vec{x}-\vec{y})h^\mu(\vec{x},\tilde{a}\wedge\tilde{b}) \quad . \tag{5.65}$$

This is called current algebra in the literature. For the time component $\mu = 0$, one speaks of the current algebra of charges. For a space-component one has to introduce in the second member a distribution (usually called Schwinger terms, see O'Raifeartaigh lectures).

Very few physical results require the local form of current algebra and cannot be deduced from the form (5.63). However, physicists prefer to consider current algebra as an hypothesis. They like the analogy with quantum mechanics which is expressed by the algebra (= Lie algebra of the Heisenberg group) of p's and q's at a given time. Let us note also that in this frame B. W. Lee (*Phys. Rev. Lett.*, 17, 145 (1965)) has given a meaning to SU(6) symmetry. There is an anthology on "current algebra" physics (see below).

Boot-Strap. When there are so many particles, one hesitates to distinguish which ones are elementary. Boot-strap is a physical concept† which deals with particles on a more democratic basis. Boot-strap is expressed by non-linear (simply quadratic) equations, invariant under the hadronic symmetry group G (no larger group than SU(3) has been used). Such equations yield solutions which break the symmetry of G. Indeed, from the abstract point of view of group invariance, these equations are of the form

$$a \vee a = \lambda a \quad ,$$

and we have already shown how this yields the directions in nature which break the $SU(3) \times SU(3)$ symmetry.

For the readers who wish to read the physics literature, we recall the existence of the anthologies (with commentaries) of original papers, that we have already mentioned.

- *Quantum Theory of Angular Momentum*, Biedenharn, L. C., and van Dam, H., Academic Press, New York (1965).
- *Symmetry Group in Nuclear and Particle Physics*, Dyson, F. J., Benjamin, New York (1966) (which also contains three lectures by Dyson).
- *The Eightfold Way*, Gell-Mann, M., and Ne'eman, Y., Benjamin, New York (1964).
- *Current Algebras*, Adler, S. L., and Dashen, R. F., Benjamin, New York (1968).

---

† Although its father, G. F. Chew has written recently a paper entitled "Boot-strap, a scientific concept?", and given an ambiguous answer!

ACKNOWLEDGMENTS

It was very exciting to prepare these lectures, and discuss some points with my colleagues in or near Bures (Deligne, Fotiadi, Lascoux, Radicati, Stora, Thom, et al.). For the preparation of these notes, I benefited from many discussions with the participants of the Rencontres, among them B. Kostant, G. Mackey, C. Moore, L. O'Rafeartaigh, and more especially the Rencontres Director, V. Bargmann. His friendly advice helped me to improve many points of the original draft. By their careful readings, Dr. Abellanas and Professor Bargmann suppressed most of the misprints of the original draft.

Unhappily, these notes do not convey the lively interruptions during the lectures. They are incomplete (no time to deal with molecular and solid state physics!) and written much too hastily. I apologize to the reader, asking him to remember that he is not reading a book, but perishable lecture notes. I still hope they will incite some readers to better learn this fascinating part of physics.

I acknowledge the wonderful hospitality offered by the Battelle Memorial Institute, to the participants (and their families!), and the perfect organization of this fruitful Rencontres. The only sad point was the absence of E. P. Wigner, the most, and yet not enough, quoted scientist in these notes.