

THE ROLE OF PHYSICS IN EDUCATION

By Leon M. Lederman

VIII INTERAMERICAN CONFERENCE ON PHYSICS EDUCATION

July 7-11, 2003, Havana, Cuba

"Science is an adventure of the whole human race to learn to live in and perhaps to love the universe in which they are. To be a part of it is to understand, to understand oneself, to begin to feel that there is a capacity within man far beyond what he felt he had, of an infinite extension of human possibilities . . .

I propose that science be taught at whatever level, from the lowest to the highest, in the humanistic way. It should be taught with a certain historical understanding, with a certain philosophical understanding, with a social understanding and a human understanding in the sense of the biography, the nature of the people who made this construction, the triumphs, the trials, the tribulations."

I. I. Rabi

I. INTRODUCTION

The beginning of "modern" science, the science of Galileo and Newton, was physics. This unique blending of experiment and of synthesis, the introduction of mathematics as the cadence of the intimate joining of experiment and synthesis became a model for the development of all science. Experiment provided the clues, suggestions. Creative imagination, inspired guesses emerge from the synthesizers. Once upon a time, there was no division of labor between synthesis and experiment-but this changed as the skills of experiment and the skills of the theoretical physicists separated into distinct although overlapping qualities.

Physics addressed the whys of phenomena, in the beginning, mostly natural like the movement of planets and moons. This processed into simply constructed motions like the trajectory of a thrown object, falling bodies. Physics began the strict process of rigorous definitions and precisely stated laws, laws of motion, the law of gravity; concepts were precisely formulated, e.g. the concepts of energy and momentum, invented because they can each be represented by a number describing a system of any complication, going through whatever internal collisions, processes, but the number never changes. Obviously we tend to some chauvinism in taking full credit for the formulation of what philosophy professors call "the scientific method". The chemists provided the first proof of the existence of atoms and biologists gave the first proof of the law of conservation of energy.

The history of physics is an unfolding of the laws from their accurate accounting of a small domain of observations to bolder generalizations, which could account for a much larger domain. A prime example is Newtonian mechanics with its wide application to astronomical and terrestrial motions, the shapes of planets, its tides, and so many other phenomena.

The study of electricity and magnetism required new formulations -and as the phenomena became better known, the laws of Coulomb, Ampere, and Faraday were brilliantly assembled by Maxwell into a new synthesis, which initially lived very comfortably with Newtonian mechanics. Phenomena extending over a vast spectrum of wavelengths: from the long waves (many meters) of radio through the infra-red, the visible, the ultraviolet to microwaves and x-rays and gamma radiation of 10^{-15} m and smaller: all electromagnetic. Meanwhile, other workers studied heat, thermodynamics, the properties of dilute gases and this inspired the quantitative use of the concepts of "atoms and molecules". A huge extension of the domain of physics followed the experimental demonstration of disturbing behaviors at the microscopic level. Quantum theory was one of the most extraordinary revolutions of the 20th century. It extended the Newtonian domain to the structural units whose properties and interactions explained chemistry and biology.

Almost all of the phenomena of chemistry now found their explanation in the quantum physics of atoms. Physicists meaningfully translated the

chemical statement: "Two elements combine to make a compound" into "two atoms combine to make a molecule". A huge piece of chemistry then had to do with structure of molecules; not only their chemical composition, but their 3-dimensional structure. Physics gave the basic Schrödinger equation and began the process of understanding molecular structures -a subject quickly taken up by computational chemistry aided by the use of huge "atom smashers" invented by physicists, but designed especially as powerful (synchrotron) x-ray sources to deduce molecular structures.

James Watson and Francis Crick both read Erwin Schrödinger's (one of the founders of Quantum Mechanics) 1945 book "What is Life" and this suggestive title set them on the historic course, the discovery of the double helix study of DNA and consequently, the radical change of modern biology to molecular biology.

Astrophysics and biophysics suggests the range of subjects in which physics provides the instruments and fundamental laws. Wherever subjects can supply information from which the types of atoms and their location can be ascertained, that is where physics can help.

A recent NRC report states: "Because all essential biological mechanisms ultimately depend upon physical interactions between molecules, physics lies at the heart of the most profound insights into biology."

Somewhat later in this lecture we will touch on a quality of laws of nature, which provides, perhaps, a clue to a profound unity of the sciences. The key word is symmetry. It will turn out that ideas of symmetry underlie the most fundamental laws of nature. Symmetry, in art and architecture, describes a harmony of proportions. This is applied in organic and inorganic processes, in relativity and quantum theory. We will show that much of this can be conveyed to students in high school and early college. Its omission seems to be inexcusable.

II. COHERENCE IN HIGH SCHOOL SCIENCE EDUCATION

US high schools (ages 14-18, pre-university) do not teach enough science. We are slowly increasing the minimum amount of science taught from one to three years. In many other countries, many more science courses are required with the necessary mathematics. However, if we now focus, not on future physicists, or even on future scientists, but on future citizens, then we find that the public science literacy, i.e. the level of understanding of science by the populations of many nations, developed and developing, is far from what a 21st century world requires.

This indicates that in much of the world, education to age ~18 does not produce science literate citizens who can participate in the decisions of their communities and nations in the vast potentialities of 21st century science and technology. These educated citizens can also participate in decisions as to which technologies are beneficial to the long term future of nations and which may have adverse effects.

There are many similarities in the educational problems of nations. For example, universally, primary school teachers are poorly trained in science and mathematics. This can be disastrous when primary school education is all there is, or even in industrial nations, where the attitudes of students towards science is shaped by the early years of schooling.

Another (almost) universal defect is the failure of school systems to recognize the natural hierarchy of science as a combination of interrelated core disciplines. One can imagine a pyramid (see Appendix, Figure 1), the base of which is mathematics, a pure invention of the human mind (but with its own rationale for existence) that provides the logical structure and the language of the other disciplines. This is followed by physics, which require mathematics as its language and provides the logical underpinning for other disciplines.

Next in the hierarchy is chemistry. The assertion that all laws of chemistry are logically supported by physics is defensible; examples abound. The periodic table lists the elements in a regular order that follows from the quantum structure of atoms and the exchange symmetry (i.e. the Pauli exclusion principle) provided by the laws of physics.

It may suffice to establish the physics-chemistry ordering in the pyramid by simply noting that all of chemistry depends upon the (quantum) structure of atoms and the resultant forces between atoms. In the case of complex systems and the laws that emerge from complexity, again the understanding involves reduction to atoms or molecules. Some examples are Brownian motion, the concept of temperature and pressure, solutions of salts, and electrochemistry. (However, it should also be noted that many principles of chemistry are not usefully reduced to underlying atomic and molecular properties.)

In a similar manner, biology would appear in the pyramid above chemistry, noting that modern, molecular-based biology is underpinned by physics and chemistry. James Watson's (1968/1980) book, *The Double Helix* details how powerfully these subjects entered into the discovery of the structure of DNA. Again, new laws of complex systems emerge out of the complexity but can (usefully or not) be reduced to basic properties of atoms and molecules.

What are the implications of all this for science education? The hierarchy of complexity emerging from simplicity should therefore be the guide for pedagogy. It calls for the exposure of students, at the earliest possible age, to study of the atom, the key to this revolution in science and technology. It would be useful for middle school students to have a reasonable sense of how small atoms are and some idea of the role of atoms in our understanding of the world. This information should also come to them through powerful out-of-school sources such as museums, science magazines, TV science programs, and newspapers that treat science.

Many school systems in Europe and Asia cycle through the core disciplines—for example, teaching a month of physics, then chemistry, and then biology. This is an improvement on the US system, in which about 50% of students study biology for one year, followed by a year of chemistry, and then (for only 20 %) a year of physics. In this proposed hierarchy, a full year of physics must precede chemistry to make use of the logical connections -that everything is made of atoms and that physics *is* the clarification of atomic structure subject to the principles of quantum physics. There is no basis for the pessimistic view that ninth graders cannot grasp such abstract ideas. Because it takes about a billion atoms to occupy a speck the size of the period at the end of this sentence, atoms are indeed an abstract concept. However, with computer simulation, imaging, and modeling, students can be exposed to the idea of atoms long before ninth grade. Ideas of combinations of atoms into molecules whose curious mechanics should replace and explain concepts of combining elements to make compounds. These can be modeled, and even the idea of forces binding atoms together can be qualitatively taught before ninth grade. Descriptive aspects of science should be presented in middle school, where students should also develop mathematical skills such as graphing (slopes and intercepts) and be introduced to algebra, data tables, and mathematical problem solving. By ninth grade, students should have learned enough algebra to begin the formal study of conceptual physics.

Key Elements of a 21st Century Approach to Science Literacy

Again, there are many alternative approaches to a detailed, 3-year core curriculum. Here is a summary of the key elements in the present example of a 21st century approach to science literacy for **all** high school students:

- There must be a core sequence consisting of the three key disciplines, in the order of physics, chemistry, and biology. Mathematical applications should be continuous so that biology makes use of all the mathematics up to pre-college level.
- It is essential that there be excellent laboratory work that is closely synchronized with the course work. Here we stress that a key to successful laboratory work is “inquiry” rather than the traditional “cook book” method.
- Each discipline should spend 20 % to 30 % of the course year on process, including selected pieces of history, applications to societal problems, and the social, political and economic issues that entwine science and society. The importance of this recommendation cannot be exaggerated. The stories associated with the core disciplines will be remembered long after $E = mc^2$ is forgotten. The stories embedded in the content will help to create a “science way of thinking”.
- Teachers should have ample time to conference with each other. This is professional development in the best sense. If the nations are serious about the importance of science education, they must support the bold idea that teachers can profitably use one day a week of course time which may include tutorials on line or visiting scientists from local universities. The conversations referred to may be regular weekly meetings of the physics, chemistry, biology, and math teachers. The connections between disciplines should be celebrated by many examples.
- A crucial point is that the impact of atoms, atomic structure, and molecule formation be taught early so that students enter chemistry with a fair command of atomic properties. Ideally, the science, math, and computer learning curricula in primary and middle school should be reviewed to ensure a match with this dramatically new high school sequence.
- Students can be encouraged to select science electives in, say, the third and fourth years after the core P-C-B curriculum. Advanced Placement or honors physics would be especially profitable, as would a course in earth science, which would make use of the students’ knowledge of the core disciplines. Of course, many important subjects are not mentioned here that can be inserted into the core curriculum, offered as electives, or even scheduled as a required fourth year. One is astronomy; another is statistical theory, prediction, and probability. The detailed folding in of computers and educational technologies is also missing from this sample curriculum, but these are important as well.
- As teacher interactions become a serious component of life in high school, one can imagine inviting in the history teacher and the art and literature teacher to these meetings. Discussions could lead to projects, seminars that would illuminate the fundamental unity of knowledge. We would then be truly educating our young students for all possible futures.

III. TEACHING CONCEPTUAL PHYSICS

In the proposed “logical sequence” of introductory courses, we are strongly advocating that the first introduction to a full year study of a discipline be physics, followed by chemistry and then followed by (mostly) molecular-based biology. We expect that high schools would require these core curricula courses for all students, accompanied by mathematics through at least pre-calculus, with many students taking calculus. However, high schools should offer elective courses and, in a truly 21st century school, another full year of science should eventually be required. An ideal candidate for a fourth year requirement is earth science (Geology), which offers the advantage of making use of all three core disciplines. Environmental science is another excellent integrated science that will teach the student the full power of the combination of disciplines. Elective courses would generally include astronomy, a second and more advanced year of physics, chemistry and biology. Often these would serve as examples of university-level instruction.

The difficulty in the physics, chemistry and biology sequence is that ninth grade (age 14-15) students do not yet have the mathematics that usually accompanies physics taken by older students in 11th or 12th grade. Thus the course must emphasize the grasp of concepts. The instructor must be aware of the level of algebra that his students are comfortable with and this implies the ability of the mathematics and physics teachers to communicate with one another on a regular basis.

The pioneer in the US on this kind of physics is Paul Hewitt, who originally devised this type of course for students who were considered mathematically limited. However, as the new sequence of “physics first” started to evolve in the US, Hewitt’s approach was modified as were several other efforts at engaging students in the concepts of physics, its powers and processes, but with a dilution of the requirement for any more than an introductory course in algebra. One approach (ACTIVE PHYSICS, Eisenkraft) concentrated on the applications of physics to the immediate interests of these young students, e.g. sports, the kitchen, etc. Another (Concepts in Physics, Hobson) emphasized the modern physics ideas of quantum theory and relativity.

We are now converging on a course which must do several things:

- 1) Give the student a clear impression of the respective roles of experiment and theory.
- 2) In spite of the small amount of mathematics, students should be able to appreciate the incisiveness of a mathematical statement, its power to predict the future of a simple system

(e.g. a ball rolling on a smooth level surface) and eventually, the value of mathematics in its role of “the language of science”. Motivations like these must improve the learning of mathematics.

- 3) The student should learn some of the history of physics, its heroes and heroines, what kind of characters created the concepts, e.g. Galileo, Newton, Maxwell, Einstein, Feynman..., how physics builds on prior knowledge and how great advances come about. Included here are something of the moral dilemma’s which faced physicists (also chemists and biologists in the next courses) as citizens of their nations in contrast to their responsibilities towards the entire world out of which emerged the heritage to which each physicist is beholden.
- 4) Finally, the concepts enter. It is not often that high school physics instructors can clearly distinguish concepts from facts or parts of a physical system. For example, consider this list: coal, oil, wood, molecule, energy, waterfall, and battery. Which word doesn’t fit? It is surprising how many get it wrong. Of course “energy” doesn’t fit, it is a concept, invented (discovered?) by physicists in order to simplify the understanding of processes. All the other words describe objects that exist even without physicists. The awesome power of the energy concept is derived from the Law of Conservation of Energy, which tells us that if we precisely define energy in various systems, e.g. kinetic energy, potential energy, electrical energy, nuclear energy, etc., then in a complicated device in which balls are flying around, springs are compressed and released, wheels are turning every which way, Bunsen burners are lit, etc., etc., there is one number, the total energy, which stays constant forever!

Other concepts such as momentum, velocity, force, acceleration, angular momentum, temperature and pressure, are also inventions out of the minds of physicists, not only to torture students, but also to enable the understanding of physical systems. It is out of the skillful use of these and other concepts that we can understand the pendulum clock (and design better clocks), the trajectory of a thrown object, the motion of vehicles, of ocean waves, of continents, the flight of artificial satellites, the motion of planets, stars and galaxies, the microstructure of matter, radioactivity, the production of radio waves, light from the sun and from other sources, the flow of electrons in wires, conductivity, the processes involving heat flow, etc., etc. In short, all of the devices that support our post-industrial societies and which, together with the expressions of human feelings and aspirations, define our civilization.

The teaching of conceptual physics is much more difficult than a senior high school (age 17-18)

physics course, even one requiring calculus. Instructors cannot “hide” behind the algebra but must emphasize the need for clarity in definition and the grasp of the concepts as they are introduced. The instructors must find sources (e.g. G. Holton’s recent reissue of his Project Physics textbook) for the process of how physics works the telling of stories and, very important, the instructor must choose his topics so that Conceptual Physics becomes a prerequisite for chemistry. This requires more communication between the instructors. The key to a smooth transition to chemistry is atomic structure: the forces between charged particles, circular motion, energy states and the teaching of that curious behavior of atoms: quanta with discrete energy states, shell structure, the concept of angular momentum but now quantized, e.g. spin. Here, we expect that the students’ knowledge of atomic physics will be enriched by the chemistry instructor who can review these concepts (most high school chemistry books compress a year’s worth of physics into their first few chapters).

In fact, the beauty of the physics, chemistry and biology sequence is just this use of physics in chemistry and of chemistry and physics in biology. However, the student in chemistry and then biology is more advanced mathematically and so concepts are now revised to reflect this mathematics.

In all of this, we emphasize that our design is for all students so that the science way of thinking penetrates into their minds and plays a role in decisions that these future citizens must make as participants in democratic society.

IV. WHAT IS SYMMETRY?

Much of science and mathematics has to do with understanding how change occurs in nature and in social and technological systems. Much of our technology has to do with creating and controlling change. Constancy, often in the midst of change, is also the subject of intense study in science.

The simplest account that can be given about a system is that it does not change. Because scientists are always looking for the simplest possible accounts, they are always interested, fascinated by any aspect of a system (thing) that doesn’t change even when many other aspects of the system do change.

Such are the conservation laws of energy, mass, electric charge, momentum or angular momentum. Just imagine a large number of molecules (or steel bee bees) moving rapidly, colliding elastically with one another and with the walls of a container and the assertion: the total momentum of this array of particles ($\vec{p} = \sum \vec{p}_i$, a vector sum!) remains constant through time! A simpler example is an explosion of a

steel sphere, which is designed to fragment into hundreds or thousands of pieces. Since the sphere and its explosives were initially at rest, the total momentum $\sum \vec{p}_i$ starts out at zero and therefore remains at zero as the fragments fly out in all directions. This remarkable simplicity may seem to be difficult to verify, yet all physicists will agree that it is a correct statement.

Symmetry is the operative concept of these “quantities that remain constant”, the 4-syllable word is “invariance”, e.g. the total momentum of a group of particles is invariant (does not change) with respect to time.

Often the symmetry refers to a pattern whose appearance does not change when we rotate (or translate), reflect (as in a mirror), or stretch in all dimensions.

A more abstract form of symmetry is when we compare different classes of particles as in this sentence: the properties of a particle, e.g. mass, lifetime, quantity of electric charge, and spin -do not change (are invariant) when we change the sign of the electric charge (i.e. change particle to antiparticle).

Symmetry is a crucial concept in mathematics, chemistry, and biology. Its definition is also applicable to art, music, architecture and the innumerable patterns designed by nature, in both animate and inanimate forms. In modern physics, however, symmetry may be the most crucial concept of all. Fundamental symmetry principles dictate the basic laws of physics, control structure of matter, and define the fundamental forces in nature. Some of the most famous mathematicians and physicists had this to say about symmetry:

"I aim at two things: On the one hand to clarify, step by step, the philosophic-mathematical significance of the idea of symmetry and, on the other, to display the great variety of applications of symmetry in the arts, in inorganic and organic nature."

Hermann Weyl (in his book "Symmetry")

"Special relativity emphasizes, in fact is built on, Lorentz symmetry or Lorentz invariance, which is one of the most crucial concepts in 20th Century Physics."

C. N. Yang (Nobel Laureate in Physics)

"Look at the symmetry of the laws, i.e., look at the way the laws can be transformed, and leave their form unchanged...." and, "Symmetry is fascinating to the human mind;"

everyone likes objects of patterns that are in some way symmetrical.... but we are most interested in the symmetries that exist in the basic laws themselves."

Richard P. Feynman (Nobel Laureate in Physics; in his "Lectures on Physics")

"I heave the basketball; I know it sails in a parabola, exhibiting perfect symmetry, which is interrupted by the basket. Its funny, but it is always interrupted by the basket."

Michael Jordan (former Chicago Bull)

The most powerful microscopes humans have built are the great particle accelerators, such as the Tevatron at Fermilab, in Batavia, Illinois. The Tevatron accelerates protons and antiprotons in opposite directions in a great circle, to energies of one trillion electron volts (as though you had a one trillion volt battery hooked up to a vacuum tube). These particles then collide head-on. The quarks and anti-quarks, inside of the protons and antiprotons, themselves collide. By reconstructing the debris from a collision of this kind physicists get a kind-of "photograph" of the structure of matter at the shortest distance scales ever seen, distances as small in comparison to Michael Jordan's basketball, as the basketball is small in comparison to the radius of the orbit of Pluto.

By studying physics at these tiny distance scales we can see that the forces of nature begin to share a common property, which is unseen at lower "magnification," at the larger distant scales. Today we understand that all of the fundamental forces in nature are unified, in a sense, by one elegant symmetry principle. This principle is subtle, and therefore it has a fancy name: it is called "local gauge invariance." Later on we'll try to explain it to you, but for now please accept this as a statement of fact.

The discovery of this unifying symmetry principle has allowed us to leap conceptually to distance scales one thousand trillion times smaller than can be seen with our most powerful particle accelerators. This has allowed us to conceive of what the Universe was like in the first one billionth of one billionth of one billionth of a second! At such short distances quantum gravity is active and rejects our normal notions of space and time. There we must use the symmetry principles (and related topological ideas) to imagine theoretically the complete unification of all forces. This leads to new ideas, to something called the "superstring" and an arcane mathematical system called M-theory that no one yet understands (we really don't even know what "M" stands for). Nevertheless, this is, perhaps, the most symmetry-pregnant logical system ever conceived by the human mind.

Preface from a book on Symmetry by C. Hill and L. M. Lederman

The purpose of this book is to provide a modern perspective on the subject of physics, the basic science of the laws and principles of space, time, and matter. This is intended for use in an introductory high school or college level physics course. It does not replace the standard course material, but rather, supplements it. This is also intended to be informative and entertaining for the beginner or interested reader who, at any stage of life, may seek a deeper understanding of what humans know about the physical world around and within us.

It is not possible to omit the climbing of the traditional ladder if one's goal is the mastery of the subject of physics. Physics is a very vertical subject. Every physicist must learn what $\vec{F} = m\vec{a}$ means, and how to use it. This requires the natural language of physics, which is mathematics, and we all know that Differential and Integral Calculus was invented by physicists! Every physicist must know that energy and momentum are conserved; that these are fundamental conservation laws that come from the basic defining principles of physics. Yet, remarkably, not every physicist knows that these conservation laws and even the fundamental dynamical principles of physics all come from something much deeper... fundamental symmetries of the basic structure of space, time, and matter.

The remarkable connection between the dynamics, i.e., the conservation laws and the forces of nature, etc., and symmetry, is a modern concept that was invented in the 20th century. It begins with the philosophical approach that Einstein brought with his reasoning about nature, but it is most evidenced in the discovery of "Noether's Theorem", which intertwines dynamics together with symmetry. When viewed from this perspective, the subject of physics leaps well into the modern era.

We believe that fundamental ideas about the connection of symmetry to physics are not far removed from the beginning student. It seems to us patently obvious that the best way to make physics as interesting and relevant to beginning students as it truly should be, is to introduce some of the ideas of symmetry principles as early as possible. This should be done with the rest of the curriculum. Basic mathematical ideas about symmetry, e.g., "group theory", can also be introduced into the high school math curriculum, perhaps with an example clearly and completely worked out such as the symmetry of the equilateral triangle. Many examples can then be given of the interplay of symmetry with physics, either in little self-contained problems or in discussing the large framework of conservation laws and the basic forces in nature.

So why doesn't Johnnie know Symmetry? The crucial role of symmetry in physics seems to us to be at odds with the practice of the complete omission of this subject in the teaching of beginning students, not only in the high school curriculum, but in the standard first year college calculus-based physics course. Symmetry does not appear in the US Standards. The absence of symmetry discourse in our teaching of physics represents a throwback to a purely nineteenth century perspective, which seems to permeate the curriculum.

Now, students are always attracted to the modern and "sexy", highly visible end-products of modern physics, e.g., semiconductors, lasers, nuclear and atomic processes, superconductors, superfluids, the formation of galaxies and black holes, the Big Bang, quarks and strings. The process of really learning about these things takes some six to eight years of undergraduate and graduate physics courses. Only then, if the student chooses a very abstract field of specialization, such as theoretical physics, will she begin to see the fundamental role of symmetry emerging in the basic laws of physics. Indeed, even today many practicing physicists have no idea about the concept of, e.g., gauge invariance, which is the basic symmetry principle underlying all known forces in nature!

It is possible, nonetheless, to incorporate elements of these underlying ideas of symmetry and its relationship to nature into the beginning courses in physics and mathematics, at the high school and early college level. They really are not that difficult. When the elementary courses are spiced with these ideas, they begin to take on some of the dimensions of a humanities or fine arts study: *Symmetry is one of the most beautiful concepts, and its expression in nature is perhaps the most stunning aspect of our physical world.* We believe that symmetry will prove to be a vehicle for stimulating and maintaining the student's interest in physics at the outset and a connection to the deeper aspects of the physical world. We have time only to present a few examples.

Application of Symmetry: Space and Time

The operation we chose first to test an aspect of the space provided to us by our universe is to move a laboratory along, say, the x-axis by a number of units, e.g., N meters. This is a space translation say from position A to position B, and the question we ask is: do the experiments performed in the laboratory at position A give results which are, in any sense, different from experiments performed when the laboratory is at position B? One might perhaps expect that the amount of motion, N meters would appear in the results of some measurements at B compared to A. But in fact, we find that there is no effect of the translation, the results of all experiments at A and B are identical. The distance N cancels. Generalizing: the laws of physics are invariant to a translation in space. All points in space are equivalent.

Translations in time give a similar result. The laws of nature are invariant to the absolute time at which these laws are tested.

Rotations of the laboratory in space are another test of a symmetry of space. Thinking of a satellite (so as not to be influenced by our awareness of gravity) way out in free space, its orientation does not influence the experiments carried out in the satellite.

We could apply this to a laboratory on the surface of the earth, but then the rotation would have to include the earth as part of the laboratory. The conclusion then is: The laws of physics are invariant to translation and to rotations in space and to translations in time.

Now we introduce the mathematical physicist Emmy Noether, who in 1913, proved a spectacularly important theorem: for every symmetry there exists a corresponding conservation law. This is not easy to prove, but the consequences can be easily stated.

- The symmetry of space translations corresponds to the law of conservation of momentum.
- The symmetry of space relative to rotations corresponds to the law of conservation of angular momentum.
- The symmetry with respect to time corresponds to the conservation of energy.

Thus, these well confirmed conservation laws teach us about the structure of space and time. These are three examples of continuous symmetry principles. Continuous because the transformations of displacement in space can be any displacement from a nanometer to a light year; the transformations in time can be any amount of time and the space rotations can be any number of degrees about any axis. There are some important discrete symmetries. I will list, without much discussion, the three important ones that emerge from the quantum theory.

1. Mirror Symmetry: The transformation is reflection in a mirror. The invariance is the statement that everything seen in a mirror corresponds to a real physical system. One way of stating this reflection symmetry is that all experiments and laws of nature derived from these experiments hold true in the mirror world. The official name for this symmetry is "parity" with symbol "P". To illustrate briefly, if one turns a screw with a screwdriver into a block of wood, the clockwise rotation of the screw advances it deeper into the wood. This is, by convention, called a "right handed screw" (There is a right hand rule in which the thumb points towards the advance of the screw and the fingers naturally curl to indicate clockwise rotation of the screws). But the mirror image of this profound activity indicates a counterclockwise rotation of the mirror screw as it advances into the mirror block of wood. Does this violate a law of nature? No, because the screws in our shop are, by

convention, right handed. However, it is easy to ask the shop to make left handed screws. Such screws exist. No problem! For those of you that are more mathematical, if we have a coordinate system with the z axis pointing to the mirror (the mirror is in an xy plane), the reflection transformation is simply replacing all z values in the laws of physics by $-z$. We would then state it in grown-up language as the laws of physics are invariant to replacing z everywhere it appears in the equations of physics by $-z$. In the Noether theorem, it is called the Law of Conservation of Parity.

2. **Another discrete symmetry is the direction in the flow of time: past to future.** So, the laws of physics are invariant to a change such that t (time) is everywhere replaced by $-t$. Many readers would here bridle and object because the experience is that we get older, hair turns gray, a dropped and hence smashed egg rarely puts itself together and jumps up whole, into the hands of the cook. However, the law finds its validity in microscopic physics where, say, two particles collide and perhaps change their states so that $A + B \rightarrow C + D$. In "normal time" A collides with B, and C and D go off in opposite directions at some angle with respect to the AB line. "T" or time reversal ($t \rightarrow -t$) invariance says that C + D coming in towards each other (along the previous CD line) will always produce A + B going out.

The examples out of our experience cited above are derived from complex systems and their origins and their relation to T invariance is outside the scope of this discussion.

3. The final discrete symmetry here is a new twist, not related to space and time, but to electric charge. It says: the laws of physics are invariant to a change in the sign of electric charge, i.e. electron $\rightarrow$ positron. However, it is more general

and the better statement is: the laws of physics are invariant to changing matter to antimatter. Its fancy name is "charge conjugation" and the symbol "C". A mythical world composed of anti-people does not obey the same laws of physics.

In summary, out of a plethora of symmetries, we have described six: three "classical" symmetries and three discrete (quantum) symmetries, P, T, and C. We have omitted many important ones, one of which, as we said, merely establishes the theory of relativity and has, as a consequence, the mysterious equation $E = mc^2$. We will close with two important statements.

- 1) The classical symmetries, to the best of our current knowledge, are always true symmetries in the sense of our definition. The P, C, and T symmetries turn out to be imperfect in the sense that for certain types of quantum processes, they fail! The laws of nature, for certain forces of nature, are **not** invariant to the processes of P (mirror reflection), T (time reversal), and C (matter -anti-matter interchange). The study of this failure is one of the hottest subjects in modern physics.
- 2) There is a symmetry that is crucial to the understanding of chemistry in the famous chart: The Periodic Table of the Elements. It is called exchange symmetry. Two identical particles, say, electrons, or hydrogen atoms, are so really identical that if we interchange the two particles, the system looks exactly the same (definition of symmetry). We have published the implications of this in *The Science Teacher* (February 2001, page 33). It is nothing less than the "Pauli Exclusion Principle" which guides the structure of the chemical elements!

These examples are designed to give you a glimpse of the importance and the power of the concept of symmetry.

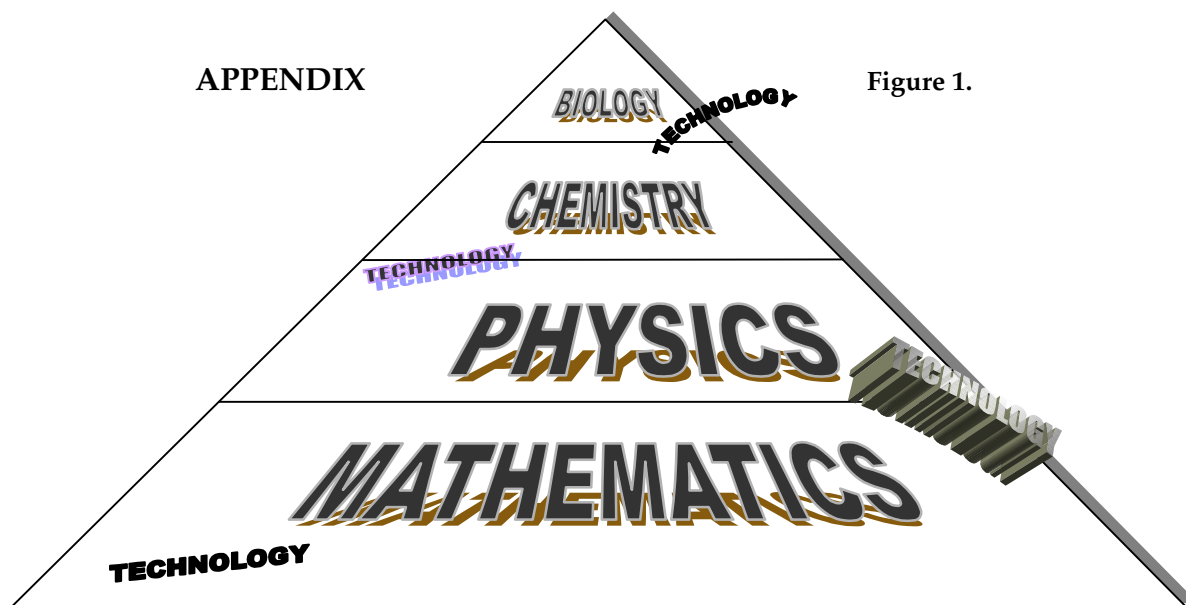


Figure 1.

THE EFFECT OF TRAP ANHARMONICITY ON THE CRITICAL TEMPERATURE FOR BOSE-EINSTEIN CONDENSATION

Augusto González*, Instituto de Cibernética, Matemática y Física, Ciudad de La Habana, Cuba

ABSTRACT

The anharmonicity of a magnetic atomic trap at long distances from its center (forty times the radius of the first atomic orbit along the elongated axis, i.e. hundreds of microns) is shown to increase up to 45 % the temperature for Bose-Einstein condensation. This effect is perhaps small in the traditional traps, but should certainly be taken into account in magnetic microtraps, which characteristic dimensions are of order 1 mm .

PACS numbers: 03.75.Fi

RESUMEN

El anharmonismo de una trampa atómica magnética a las distancias largas de su centro (cuarenta veces el radio de la primera órbita atómica a lo largo del eje largo, es decir ciento de micras) se muestra para aumentar por encima de un 45 % la temperatura para la condensación de Bose-Einstein. Este efecto es quizás pequeño en las trampas tradicionales, pero debe tenerse en cuenta ciertamente en microtraps magnético que las dimensiones características son de orden 1 mm .

Unlike the situation in liquid Helium, atomic vapors undergoing Bose-Einstein condensation (BEC) in magnetic traps are very rarefied¹. As a consequence, the free-boson model works extremely well. Interaction effects, computed in mean-field approximation, are shown to decrease the critical temperature for BEC in a few percent². Still lower corrections come from the approximation of the one-particle discrete spectrum by a continuum of states, the so-called "finite-N" corrections², which also decrease the critical temperature.

In the present paper, we show that a relatively important increase of the critical temperature (up to 45 %) could be related to trap anharmonicities at "long" distances from the center. By long distances, we mean around forty times the radius of the first atomic orbit along the elongated direction in the trap potential, i.e. around 200 μm . In the commonly used traps, which characteristic dimensions are a few centimeters, the belief is that anharmonic effects should be very weak at distances of 200 μm . Nevertheless, we present an example of a recent experiment³, in which the measured values of the total number of atoms in the trap and the critical temperature are not consistent unless anharmonicity (or other) effects raising T_c are included. On the other hand, the trap potential should certainly be anharmonic at 200 μm in the recently developed atomic microtraps⁴, which characteristic dimensions are of order 1 mm.

In the continuum free-boson model, the critical temperature is simply estimated from the conditions $N_0 = 0$, $\mu = 0$ in:

$$N - N_0 = \int_0^\infty \frac{g(\epsilon) d\epsilon}{\exp\left(\frac{\epsilon - \mu}{k_B T}\right) - 1} \quad (1)$$

N_0 is the number of atoms in the condensate, and $\epsilon \geq 0$ is the one atom excitation energy. The chemical potential takes values $\mu \leq 0$. g is the density of states.

Magnetic traps with cylindrical symmetry are very common. The trap potential, coming from the interaction of an hyperfine atomic species with the magnetic field, is proportional to the latter which, near the trap center, has a minimum and is written as:

$$B(\rho, z) = B_0 + b_\rho \rho^2 + b_z z^2. \quad (2)$$

The corresponding harmonic oscillator frequencies are ω_ρ and ω_z . It is useful to define $\omega_0 = (\omega_\rho^2 \omega_z)^{1/3}$, in terms of which the level density is written $g(\epsilon) = \frac{1}{2} \epsilon^{2/3} / (\pi \omega_0)^3$, and the critical temperature becomes:

$$k_B T_c^0 = \left(\frac{N}{\zeta(3)} \right)^{1/3} \pi \omega_0 = 0.94 N^{1/3} \pi \omega_0. \quad (3)$$

where $\zeta(x) = \sum_{n=1}^\infty 1/n^x$ is the Riemann Zeta function.

E-mail: *agonzale@icmf.inf.cu

The corrections to T_c due to interaction or finite-N effects take, respectively, the form^{2,5}:

$$\frac{T_c - T_c^0}{T_c^0} = -1.33 \frac{a}{r_0} N^{1/6}, \quad (4)$$

$$\frac{T_c - T_c^0}{T_c^0} = -0.73 \frac{\bar{\omega}}{\omega_0} N^{-1/3}, \quad (5)$$

where $r_0 = \sqrt{n/(m\omega_0)}$, a is the s-wave scattering length, and $\bar{\omega} = (\omega_z + 2\omega_p)/3$. Typical values $N \sim 10^6$, $a/r_0 \sim 10^{-2}$ lead to small numbers in the r.h.s. of Equations (4,5). As mentioned, the corrections (4) are computed within mean-field theory. More elaborated variational calculations based on diffusion Monte-Carlo⁶ or hypernetted chain theory⁷ support the validity of mean-field approximations at the relevant particle densities. Notice that both corrections (4) and (5) are small and make T_c decrease. In a general power-law trap, the corrections may be of any sign and their relative magnitudes may be still greater⁵. However, for usual magnetic traps one expects a harmonic potential coming from (2), at least in the vicinity of the trap center.

Let us suppose that at long distances, $r \gg r_0$, anharmonic terms are relevant in (2). In fact, one expects a saturation of the magnitude of B , which should approach the bias field value as the distance increases. In a trap potential that saturates at long distances, the density of energy levels decreases at high energies⁸. It means that the occupation of the ground state is increased, and thus the critical temperature is raised.

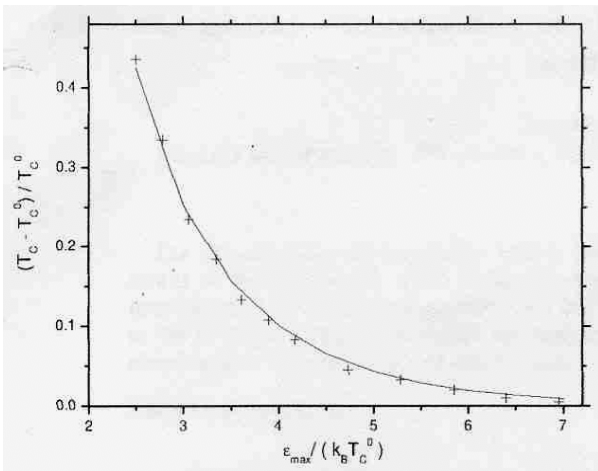


Figure 1. Shift in the critical temperature as a function of the energy cutoff. Solid line: corrections coming from Equation (6), crosses: finite-N calculations for the trap used in Reference 3.

We will model the density of levels of the (saturating) anharmonic potential in the simplest way: by truncating the one-particle spectrum. It means

that an upper integration limit, $\epsilon_{\max}$, is set in (1). The critical temperature is thus determined from:

$$t^3 \zeta(3) = \zeta(3) + \text{Li}_3(e^{-tx}) - tx \text{Li}_2(e^{-tx}) + \frac{(tx)^2}{2} \ln(1 - e^{-tx}), \quad (6)$$

where we have defined $t = T_c^0 / T_c$, $x = \epsilon_{\max} / (k_B T_c^0)$,

and $\text{Li}_k(z) = \sum_{n=1}^{\infty} z^n / n^k$ is the polylogarithm function.

The corresponding correction to T_c^0 , i. e. $1/t - 1$, is shown in Figure 1 as a function of x . It is around 10 % when $\epsilon_{\max} \approx 4k_B T_c^0$. As $k_B T_c^0 \sim 10^3 \hbar \omega_z$ in many of the experiments, energies of order $4k_B T_c^0$ involve particle orbits of radius around $60 r_{0z}$. On the other hand, when $\epsilon_{\max}/k_B \approx 2.5 T_c^0$, corrections to T_c^0 reach 40 %, and the cutoff distance is around $40 r_{0z}$. A fit to the long- x tail of the curve leads to:

$$\frac{T_c - T_c^0}{T_c^0} = 22 \left(\frac{k_B T_c^0}{\epsilon_{\max}} \right)^4. \quad (7)$$

In quality of example, let us consider the trap used in Reference 3, in which BEC of spin-polarized Helium was achieved. The example is quite interesting because independent measurements of N and T_c are reported. The authors-provided frequencies are $\nu_z = 115$ Hz, $\nu_p = 1090$ Hz. The critical temperature, number of atoms in the trap and the scattering length are estimated as $T_c = 4.7 \pm 0.5$ μ K, $N = (5 \pm 2.5) \times 10^6$, $a \approx 16$ nm, respectively. Turning back to Equations (3-5), we obtain that to a temperature $T_c = 4.7$ μ K, corresponds to a number of atoms $N = 1.4 \times 10^7$, well above the error bars. On the other hand, if we take the lowest value $T_c = 4.2$ μ K, then the number of atoms is $N = 8 \times 10^6$, a value closer but still outside error bars. The reason for this apparent discrepancy may be either the rough estimation of N , or the anharmonicity corrections, which have not been included.

In conclusion, we have shown that the saturation of the magnetic field at distances around 200 - 300 μ m from the trap center leads to a decrease of the density of energy levels, and thus to an increase of T_c . We have modeled the decrease of the level density simply by means of an energy cutoff in the single-particle spectrum. The corresponding correction to T_c^0 partially cancels out with interaction or finite-N effects, both of which lower the critical temperature. Although expected to be small, detailed calculations of the magnetic field configurations in the commonly used traps are needed in order to

evaluate the role of anharmonicity corrections. With regard to the atomic microtraps, which characteristic dimensions are of order 1 mm, anharmonicities are expected to be very important and should be taken into account.

ACKNOWLEDGMENTS

Part of this work was carried out at the Abdus Salam ICTP. The author acknowledges the ICTP Associate and Federation Schemes for support.

REFERENCES

- ¹LEGGET, A.J. (2001): **Rev. Mod. Phys.** 73, 307.
- ²DALFOVO, F. **et al.** (1999): **Rev. Mod. Phys.** 71, 463.
- ³PEREIRA DOS SANTOS, F. **et al.** (2001): **Phys. Rev. Lett.** 86, 3459.
- ⁴FEENSTRA, L.; L.M. ANDERSSON and J. SCHMIEDMAYER: condmat/0302059.
- ⁵SALASNICH, L. (2002): **Int. J. Mod. Phys. B** 16, 2185.
- ⁶GIORGINI, S.; J. BORONAT and J. CASULLERAS (1999): **Phys. Rev. A** 60, 5129.
- ⁷FANTONI, S. **et al.** (2002): **Phys. Rev. A** 66, 033604.
- ⁸We assume that the confinement potential decays abruptly above r_0 . The density of energy levels refers to states in which the particle motion is confined inside the trap volume, i.e. the discrete spectrum. The contribution of states in the continuum spectrum (in which the particle is free to move in all the space) to the trap thermodynamic properties is proportional to the probability of finding the particle inside the trap, i.e. equal to zero.

RELACIONES DE DISPERSIÓN EN FLUIDOS ROTATORIOS COMPRESIBLES

José Marín Antuña, Departamento de Física Teórica, Universidad de La Habana, Cuba

RESUMEN

En este trabajo se obtienen las relaciones de dispersión de ondas planas armónicas bidimensionales en fluidos rotatorios compresibles y se analizan los diferentes casos relacionados con el ángulo entre el vector de onda y el eje de rotación del fluido. Se obtiene la expresión de la velocidad de grupo de las ondas y se busca la solución de las oscilaciones de una pared inclinada sumergida en el fluido.

ABSTRACT

In this paper dispersion relations are obtained for two-dimensional harmonic plane waves in a rotating compressible fluid. The different cases are analyzed in dependence of the angle between the wave vector and the rotation axis of the fluid. An expression for the group velocity of the waves is obtained and the solution of the oscillations of an oblique wall immersed in the fluid is founded.

1. INTRODUCCIÓN

En trabajos anteriores [1]-[2] fue obtenida una ecuación diferencial de cuarto orden para la descripción de los procesos físicos bidimensionales en fluidos rotatorios compresibles y se obtuvo la posibilidad de excitación y propagación de ondas periódicas armónicas provocadas por oscilaciones de paredes verticales y horizontales sumergidas en el fluido mediante la obtención explícita de las soluciones analíticas de los problemas de frontera correspondientes.

En el presente trabajo se obtiene una relación de dispersión para las ondas planas en el fluido rotatorio compresible, se discuten los resultados obtenidos y la relación entre la velocidad de grupo de las ondas y el vector de onda. Se da, además, una solución para el problema de las oscilaciones de una pared inclinada sumergida en el fluido.

2. DESARROLLO

1. Consideremos un fluido rotatorio compresible que ocupa todo el espacio. La ecuación que describe las oscilaciones pequeñas bidimensionales en este fluido es [1]:

$$L[u] \equiv \frac{\partial^2}{\partial t^2} \left[\frac{\partial^2 u}{\partial t^2} - \nabla_2^2 u + \omega^2 u \right] - \omega^2 \frac{\partial^2 u}{\partial z^2} = 0 \quad (1)$$

donde ∇_2^2 es el laplaciano en las coordenadas (x, z) , ω es el duplo de la velocidad angular de rotación del fluido que gira alrededor del eje Oz y $u(x, z, t)$ por su sentido físico es la presión dinámica del fluido, o cada una de las componentes v_x, v_y, v_z del vector

de velocidades de las partículas del fluido. En la obtención de (1) fue considerado que $\partial u / \partial y \equiv 0$, es decir, que los procesos son simétricos en el eje Oy . Para simplificar las notaciones hemos considerado la velocidad del sonido en el fluido $c = 1$.

Propongamos a la ecuación (1) una solución del tipo

$$u = u_0 e^{i(\vec{k}\vec{x} + \gamma t)} \quad (2)$$

donde $\vec{k} = (k_1, k_3)$ es el vector de onda, $\vec{x} = (x, z)$ y γ es la frecuencia de la onda. Sustituyendo (2) en (1), obtenemos:

$$\gamma^2(\gamma^2 - \omega^2) - \cos \theta \gamma^2 k^2 + \omega^2 k^2 \cos \theta = 0 \quad (3)$$

donde θ es el ángulo que forma el vector de onda $\vec{k}$ con el eje Oz , o sea, con la dirección del vector de velocidad angular del fluido. La expresión (3) es la relación de dispersión de las ondas planas en el fluido rotatorio compresible. De (3) se obtiene

$$k = |\gamma| \sqrt{\frac{\omega^2 - \gamma^2}{\omega^2 \cos^2 \theta - \gamma^2}} \quad (4)$$

La relación de dispersión (4) nos permite concluir la diferencia sustancial que existe entre la propagación de ondas planas armónicas en un fluido rotatorio y la propagación de dichas ondas en un fluido en reposo. Para $\omega \rightarrow 0$ de (4) se obtiene la recta $k = |\gamma|$; es decir, la relación conocida de las ondas planas. Además, se observa que, en dependencia del valor del ángulo θ (o sea, de la dirección del vector de onda $\vec{k}$), el cuadro cambia cualitativamente.

Efectivamente, en el caso en que el vector de onda $\vec{k}$ sea paralelo al eje de rotación del fluido, la relación de dispersión (4) se transforma en la expresión $k = |\gamma|$, lo que significa que en la dirección del eje Oz se propagan ondas planas armónicas con cualquier frecuencia γ , incluyendo la onda tipo escalón ($\gamma = 0$). Para las ondas planas armónicas con vector de onda $\vec{k}$ inclinado ($0 < \theta < \pi/2$) se obtiene que no para todo valor de γ pueden propagarse las ondas. Efectivamente, existe una zona prohibida para la frecuencia γ : $\omega \cos \theta < \gamma < \omega$.

Por último, en el caso en que el vector $\vec{k}$ sea perpendicular al eje de rotación del fluido ($\theta = \pi/2$), la relación (4) sólo tiene sentido real cuando $\gamma > \omega$. Esto significa que en ese caso se propagarán solamente aquellas ondas cuya frecuencia γ sea mayor que el duplo de la velocidad angular ω .

2. Calculemos la velocidad de grupo de estas ondas. Después de derivar (3) se obtiene para la velocidad de grupo la expresión

$$\vec{v}_g = \frac{\gamma}{2\gamma^2 - \omega^2 - k^2} \left\{ k_1, \frac{\gamma^2 - \omega^2}{\gamma^2} k_3 \right\} \quad (5)$$

De la expresión (5) para la velocidad de grupo pueden sacarse las siguientes conclusiones. En primer lugar, para el fluido en reposo ($\omega = 0$) y teniendo en cuenta que en este caso $k^2 = \gamma^2$, de (5) obtenemos

$$\vec{v}_g = \frac{1}{\gamma} \vec{k} \equiv \vec{v}_f \quad (6)$$

Es decir, la velocidad de grupo coincide con la velocidad de fase; no hay dispersión y la energía de las ondas se propaga en la dirección del vector $\vec{k}$.

En segundo lugar, en general para $\omega > 0$ la dirección de $\vec{v}_g$, es decir, la dirección de propagación de la energía de las ondas, no coincide con la dirección del vector de onda $\vec{k}$ y tiene lugar la dispersión de las ondas.

En el caso en que la dirección del vector $\vec{k}$ coincida con el eje de rotación Oz, es decir, para $\theta = 0$, como $k_1 = 0$, $k_3 = k$, $k = |\gamma|$, para la velocidad de grupo tendremos

$$\vec{v}_g = \frac{\gamma}{2\gamma^2 - \omega^2 - \gamma^2} \left\{ 0, \frac{\gamma^2 - \omega^2}{\gamma^2} k \right\} \equiv \frac{1}{\gamma} \vec{k} \equiv \vec{v}_f \quad (7)$$

Es decir, de nuevo las ondas se propagan sin dispersión.

Si $\theta = \pi/2$, bajo la condición $\gamma > \omega$ de propagación de las ondas y , en virtud de que de (4) en este caso $k = \sqrt{\gamma^2 - \omega^2}$, tendremos para el vector de la velocidad de grupo:

$$\vec{v}_g = \frac{\gamma}{2\gamma^2 - \omega^2 - \gamma^2 + \omega^2} \{k, 0\} \equiv \frac{1}{\gamma} \vec{k} \equiv \vec{v}_f \quad (8)$$

Es decir, nuevamente una propagación sin dispersión de las ondas.

Los resultados aquí obtenidos coinciden con los de los trabajos [1], [2] mencionados anteriormente en los que aparecen calculadas las expresiones analíticas de las soluciones de los problemas sobre la excitación de ondas planas por paredes oscilantes sumergidas en el fluido, colocadas respectivamente en posición horizontal y vertical.

3. Intentemos hallar una expresión analítica del problema de la excitación de ondas planas armónicas por una pared oscilante inclinada respecto al eje de rotación del fluido, sumergida en el fluido rotatorio compresible.

Supongamos que en el fluido rotatorio compresible se encuentra una pared Γ que oscila por la ley $\theta(t)$ sin γt y que rota junto con el fluido, donde φ es el ángulo de inclinación de la pared Γ con el eje de rotación del fluido. Entonces, el problema físico-matemático sobre la emisión de ondas provocadas por las oscilaciones de la pared Γ será:

$$L[u] = \frac{\partial^2}{\partial t^2} \left[\frac{\partial^2 u}{\partial t^2} - \nabla^2 u + \omega^2 u \right] - \omega^2 \frac{\partial^2 u}{\partial z^2} = 0$$

$$\left. \frac{\partial^k u}{\partial t^k} \right|_{t \leq 0} = 0 \quad \forall k = 0, 1, 2, 3 \quad (9)$$

$$u(x, z, t) |_{(x, z) \in \Gamma} = \theta(t) \sin \gamma t$$

donde u tiene el sentido físico de la componente de la velocidad de las partículas del fluido perpendicular a la pared Γ . Introduzcamos un sistema nuevo de coordenadas:

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} \quad (10)$$

Entonces, en las nuevas coordenadas y en virtud de la invarianza del laplaciano, el problema (9) adopta la forma

$$\frac{\partial^2}{\partial t^2} \left[\frac{\partial^2 u}{\partial t^2} - \nabla^2 u + \omega^2 u \right] - \omega^2 \sin^2 \frac{\partial^2 u}{\partial \xi^2} = 0 \quad \forall \xi > 0, t > 0$$

$$\left. \frac{\partial^k u}{\partial t^k} \right|_{t \leq 0} = 0 \quad \forall k = 0, 1, 2, 3 \quad (11)$$

$$u(\xi, \eta, t) |_{\xi=0} = \theta(t) \sin \gamma t$$

donde, ahora, $\nabla_2^2 u \equiv u_{\xi\xi} + u_{\eta\eta}$. Consideraciones físicas nos permiten afirmar que la solución u no depende de η . Por ello, la ecuación en (11) adopta la forma

$$\frac{\partial^2}{\partial t^2} \left[\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial \xi^2} + \omega^2 u \right] - \omega^2 \cos^2 \theta \frac{\partial^2 u}{\partial \xi^2} = 0 \quad (12)$$

donde θ es el ángulo entre la dirección del vector de onda $\vec{k}$ y el eje Oz. Apliquemos a (12) la transformada de Laplace; como resultado obtenemos:

$$p^2(p^2 + \omega^2)U - (p^2 + \omega^2 \cos^2 \theta)U_{\xi\xi} = 0 \quad (13)$$

de donde, para la función $U(\xi, p)$, transformada de Laplace de $u(\xi, t)$, obtenemos el problema de Cauchy

$$U_{\xi\xi} - \frac{p^2(p^2 + \omega^2)}{p^2 + \omega^2 \cos^2 \theta} U = 0 \quad (14)$$

$$U(\xi, 0) = \frac{\gamma}{p^2 + \gamma^2}$$

La solución acotada del problema (14) tiene la forma

$$U(\xi, p) = \frac{\gamma}{p^2 + \gamma^2} e^{-\xi p \sqrt{\frac{p^2 + \omega^2}{p^2 + \omega^2 \cos^2 \theta}}} \quad (15)$$

Sobre la base de las relaciones de dispersión obtenidas en el punto 1 del presente trabajo, la solución del problema (11) deberá tener la forma

$$u(\xi, t) = \sin \gamma t e^{i\gamma \sqrt{\frac{\omega^2 - \gamma^2}{\omega^2 \cos^2 \theta - \gamma^2}} \xi} \quad (16)$$

la que para $\gamma < \omega \cos \theta$ y $\gamma > \omega$ tiene carácter oscilante y para $\omega \cos \theta < \gamma < \omega$ tiene carácter amortiguado. O sea, que en el primer caso las ondas se propagan y en el segundo no. Como puede apreciarse, de (16), para $\theta = 0$ y para $\theta = \pi/2$ se obtienen los resultados de [2] para las paredes oscilantes sumergidas horizontal y verticalmente en el fluido.

3. CONCLUSIONES

Los resultados arriba expuestos tienen interés para el ulterior desarrollo de la teoría de la excitación y propagación de ondas bidimensionales en fluidos rotatorios compresibles y están en plena correspondencia con resultados anteriormente obtenidos. Sin embargo, la obtención de una expresión analítica para la solución del problema (11) se dificulta en grado extremo, debido a que la expresión (15) es la transformada de Laplace de un original muy difícil de calcular. La aplicación del método de fase estacionaria expuesto por Copson [3] permitiría obtener la expresión asintótica para $t \rightarrow \infty$ del original de la función (15), la cual coincide con (16).

REFERENCIAS

- [1] MARÍN, J. (1984): **Revista Cubana de Física** 4(3), 63.
- [2] _____ (1985): **Revista de Ciencias Matemáticas** 4(1), 105.
- [3] COPSON, E.T. (1965): **Asymptotic Expansions** (University Press, Cambridge).

MÉTODO SEMIABSOLUTO DE MEDICIÓN DE SECCIONES MACROSCÓPICAS DE ABSORCIÓN DE NEUTRONES TÉRMICOS EMPLEANDO FUENTES ESTACIONARIAS

Y. Borrego, E. Herrera, L. García-Parra y O. Díaz

Instituto Superior de Ciencias y Tecnología Nucleares, La Habana, Cuba

RESUMEN

En el presente trabajo se expone un método semi-absoluto para obtener la sección macroscópica de absorción de neutrones térmicos dada una muestra de un material desconocido. Otros parámetros, no menos importante, pueden ser igualmente obtenidos. Tres factores fundamentales y necesarios se combinan en estas mediciones: facilidades experimentales existente, sencillez a la hora de ejecutar los procedimientos de medición y uso de cantidades mínimas de material. Se someten a consideración los resultados obtenidos, aplicando la metodología novedosa para un prisma de grafito que contiene una fuente de Pu-Be cuya intensidad es de 10^7 n/s, efectuando las mediciones en muestras de agua ligera envenenada con ácido bórico (~150c.c) y utilizando un detector gaseoso de BF_3 .

ABSTRACT

A semi-absolute method to measure the macroscopic absorption cross-section for thermal neutrons using a sample of an unknown material is exposed on present work. Other parameters, less important no, can be equally obtained. Three fundamental and necessary factors combine in this technique: available experimental facilities, simplicity executing the measure procedures and demanding minimum quantities of bulk material. Measurements of light water poisoned with boric acid samples (~150c.c), placed deep in a graphite pile are performed. We have used a Pu-Be neutron source which strength is 10^7 n/s and a BF_3 gassy detector. Results applying the novel methodology will be discussed here.

INTRODUCCIÓN

Para la geofísica nuclear son de gran interés los macroparámetros que caracterizan el movimiento y la interacción de los neutrones térmicos con la sustancia, en especial, la sección macroscópica de absorción de neutrones térmicos Σ_a , ya sea de las rocas matrices, esquistos o de los fluidos constituyentes. El conocimiento de la misma puede mejorar significativamente la correcta interpretación de cualquier registro litológico. Los valores de Σ_a en diferentes rocas y salmueras son factores importantes en el análisis cuantitativo del tiempo de sondeo neutrónico. Las curvas de calibración confeccionadas a partir de datos sobre la porosidad neutrónica, basadas en la detección de neutrones térmicos dependen de la sección macroscópica de absorción de las rocas, entre otros factores [1-5]. En la caracterización de yacimientos petrolíferos con generadores neutrónicos de pozo es indispensable contar con datos de la Σ_a de las rocas matrices.

La razón que obliga a medir solo mediante vías directas el valor de Σ_a en rocas matrices, es la presencia de elementos de trazas (minoritarios) altamente absorbentes en el material, cuya composición química exacta resulta bastante difícil conocer. Para que se tenga una idea: la variación en el contenido

de tierras raras o boro a niveles de unos pocos ppm (casi indetectable) dentro de la roca, puede contribuir de manera nada despreciable (por encima de un 50%) en el valor de la sección de absorción [1,4]. De manera general, se puede afirmar que la medición experimental de Σ_a es un procedimiento difícil porque presenta ciertos inconvenientes a los cuales no siempre pueden dárseles soluciones adecuadas. Por su parte, los experimentos que se realizan empleando fuentes radioisotópicas de neutrones, a pesar de las consabidas ventajas desde el punto de vista de la simplicidad experimental, tienen ciertos inconvenientes si se quiere establecer de alguna manera una sistemática y unicidad en la metodología a desarrollar.

El primero de los problemas se refiere a que en todos los experimentos para la medición de Σ_a , cuyos neutrones empleados provienen de una, o un conjunto de fuentes estacionarias; la variación de la respuesta del detector, colocado a cierta distancia de una muestra, está asociada tanto a efectos dispersivos como a las absorciones de los neutrones por la propia muestra en estudio. Las ecuaciones de la teoría del transporte para estos casos solo establecen de forma directa una relación de los flujos neutrónicos con las geometría de la instalación y un parámetro conocido como: longitud de difusión de los neutrones L , el cual tiene de manera

intrínseca una dependencia de Σ_a y a Σ_s . La búsqueda de configuraciones de medición adecuadas para tratar de minimizar o eliminar los efectos dispersivos, ya sea de manera teórica como a través de calibraciones es un problema real cuya solución dependerá de las posibilidades concretas a disponer en la instalación, o de la habilidad del experimentador. Una forma muy elegante e ingeniosa, basada en el principio del LPOM [6] resuelve esta dificultad y fue propuesta para el caso de una medición de Σ_a con una fuente neutrónica de ^{252}Cf y utilizando un detector diferencial de BF_3 . [7]

El segundo inconveniente radica en las características propias de los materiales, cuyo rango de valores de Σ_a es muy bajo en comparación con la sección de dispersión Σ_s . Para las rocas matrices, la sección macroscópica de absorción varía entre 4 y 25 unidades de captura ($1 \text{ c.u.} = 10^{-3} \text{ cm}^{-1}$) y es siempre menor que el valor de Σ_s entre 3 y 100 veces. [1].

El tercero y último es la imposibilidad de cambiar el tamaño de la muestra geológica manteniendo constante su densidad total (esto puede efectuarse solo en circunstancias muy particulares). Además, la cantidad disponible de material de roca es usualmente bastante limitada (del orden de pocos kilogramos) y la forma de la muestra puede ser regular solo cuando la roca es triturada y puesta en algún contenedor. [1]

Varios métodos han sido desarrollados, bajo el supuesto de obtener una curva de calibración, a partir de un conjunto de muestras de rocas estandarizadas cuyos valores Σ_a son ya conocidos. [8,9]. Ellos tienen su basamento en el hecho de que para cierta instalación existe una configuración fuente-muestra-detector con buena geometría donde se pueden minimizar el efecto que producen las dispersiones de los neutrones térmicos en la muestra. [10]

Nuestro trabajo está basado en el uso de las técnicas de envenenamiento, las cuales son empleadas indistintamente en muchos de estos experimentos. A diferencia, aquí será el soporte que permite establecer una dependencia lineal de la variación de la longitud de difusión de una muestra dada, cuando a ella se le adicionan pequeñas cantidades bien determinadas de un material absorbente de los neutrones cuya Σ_a es bien conocida [11, 12]. La muestra a estudiar está rodeada de un medio moderador cuya función es mantener confinado en un volumen del espacio el campo neutrónico provocado por una fuente puntual. De forma teórica y sin necesidad de una curva previa de calibración cuyo carácter casi siempre es empírico, se puede obtener el valor de la Σ_a de la muestra en estudio, entre otros parámetros.

MATERIALES Y MÉTODOS

La metodología se ha desarrollado para una instalación experimental donde es posible confinar un campo neutrónico en un régimen estacionario y en alguna zona del mismo existe una cantidad apreciable de neutrones bien termalizados. En nuestro caso se trata de un clásico Prisma de Grafito. Este dispositivo está compuesto por bloques monolíticos de grafito de alta pureza o de reactor ($\rho = 1.674 \text{ g/cm}^3$). Ellos se disponen de tal manera que conforman un cubo perfecto de 180 cm de arista como se aprecia en la Figura 1. El prisma dispone de

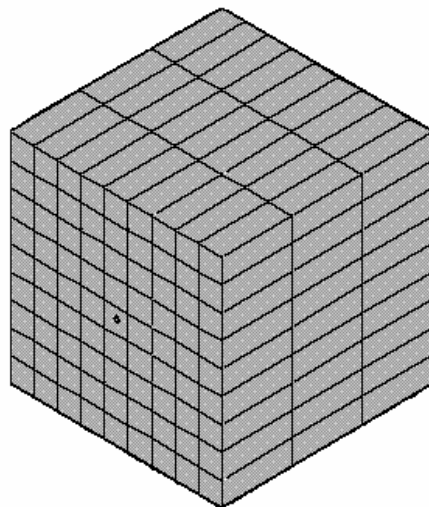


Figura 1. Esquema a escala del prisma que muestra la disposición de los bloques de grafito.

un canal central horizontal cuyo diámetro es de 43 mm para la introducción de la fuente, piezas cilíndricas de grafito, la muestra (120 ml) y el detector (ver Figura 2).

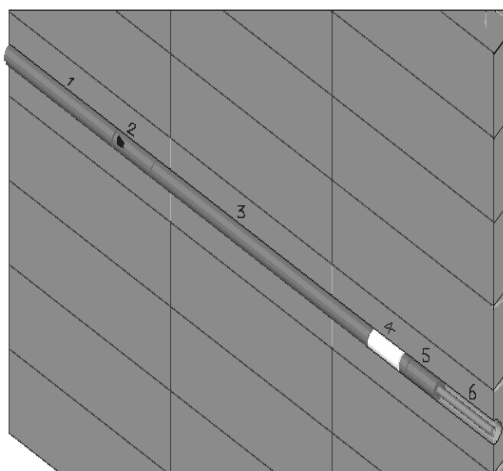


Figura 2. Disposición relativa de la fuente, la muestra y el detector para un experimento en el interior del prisma:

- 1,3,5. Cilindros o tacos de grafito,
2. Porta fuente (El punto negro representa la fuente de Pu-Be);
4. Pomo con agua;
6. Detector de BF_3 .

El cálculo teórico de la distribución de flujo neutrónico a lo largo del canal central, suponiendo que el prisma es lo suficientemente grande y el moderador elegido, en este caso grafito, permiten garantizar el establecimiento de un equilibrio térmico de los neutrones provenientes de la fuente con los átomos del medio; conduce a la solución de la ecuación de transporte neutrónico siguiente [11]:

$$\bar{D}(\nabla^2 \Phi(r)) - \bar{\Sigma}_a \cdot \Phi(r) + Q\delta(r - r_0) = 0 \quad (1)$$

donde se define la longitud de difusión de los neutrones dentro de un material como:

$$L = \sqrt{\bar{D}/\bar{\Sigma}_a} \quad (2)$$

Esta es la conocida aproximación P_1 de la difusión de neutrones térmicos, considerando una fuente puntual situada en la posición r_0 . Los autovalores de la ecuación (1) serán siempre dependientes de L y determinarán la forma de la función $\Phi(r)$ en el sistema de coordenadas seleccionado para resolver el problema, lo cual había sido comentado ya en el inicio. Debemos aclarar que para el caso de un medio moderador débilmente absorbente ($\Sigma_a \ll \Sigma_s$) es correcto definir los parámetros de la difusión a través del promedio sobre el espectro energético de equilibrio (maxwelliano). Esto es que el coeficiente de difusión D y la sección macroscópica de absorción Σ_a se pueden escribir como: [11]

$$\bar{D} = \int_0^\infty \left(\frac{E}{kT} \right) \frac{1}{3[\Sigma_{tr}(E) + \Sigma_a(E)]} e^{-E/kT} \frac{dE}{kT},$$

$$\bar{\Sigma}_a = \int_0^\infty \left(\frac{E}{kT} \right) \Sigma_a(E) e^{-E/kT} \frac{dE}{kT} \quad (3)$$

Ahora bien, constituye un hecho experimental que si se ha establecido un campo neutrónico, caracterizado por un flujo $\Phi(r)$, mediante una fuente isotópica en el interior de un moderador; al sustituir una pequeña porción del mismo por un material diferente; se producirán variaciones apreciables en el flujo neutrónico estacionario y ellas se pueden registrar colocando un detector detrás de la muestra.

La solución a este problema se obtiene aplicando la teoría de perturbaciones a la ecuación de la difusión (1) y teniendo en cuenta los cambios que aparecerán en la definición de las regiones interiores.

Es conocido además, que se establece una proporcionalidad directa entre la velocidad de conteos del detector y el flujo de neutrones térmicos en una posición determinada. Si esto es así, es posible encontrar una expresión para la longitud de difusión L' de la muestra desconocida. A la misma se arriba como resultado de escribir el flujo neutrónico a través una serie del tipo:

$$\Phi'(r) = \Phi^{(0)}(r) + \delta\Phi(r) + \delta^{(2)}\Phi(r) + \dots + \delta^{(n)}\Phi(r) \quad (4)$$

cuyos términos se pueden relacionar por una fórmula recurrente. [13]. El primer sumando de la serie $\Phi^{(0)}(r)$ no es otro que el flujo neutrónico no perturbado inicial $\Phi(r)$.

Un resultado relativamente simple se obtiene para el caso de una geometría cartesiana; pues cuando se exige la convergencia de la serie (4) se llega a una expresión muy sencilla, que se constituye en una de las fórmulas de trabajo:

$$\frac{1}{L'^2} = \frac{1}{L^2} \left(1 - \frac{Z' - Z_0}{F_1 \cdot Z'} \right), \quad (5)$$

donde: L es la longitud de difusión del moderador (grafito en nuestro caso),

Z_0 la velocidad de conteo medida cuando no se ha introducido la muestra (medición de calibración),

Z' la velocidad de conteo después de introducirla y colocado el detector en la misma posición de la medición anterior y

F_1 un parámetro geométrico adimensional, independiente del tipo de material y que se puede expresar de forma analítica para este caso. [13]

La expresión (5) es aún insuficiente para calcular el valor del parámetro Σ_a de la muestra. Para separar en la respuesta del detector las contribuciones de las absorciones y dispersiones de los neutrones acudimos a un envenenamiento gradual de la misma. Emplearemos para ello un elemento fuertemente absorbente (Boro) y en cantidades mínimas lo que garantiza evitar distorsiones significativas en el espectro energético de equilibrio (heating effect). Dichas consideraciones conducen a la siguiente expresión lineal:

$$\frac{1}{L'^2} = \frac{1}{D^*} \cdot \Sigma_a(\text{H}_3\text{BO}_3) + \frac{1}{L_c^2} \quad (6)$$

Aquí L' es la longitud de difusión de la mezcla (H_3BO_3 + sustancia inicial); $\Sigma_a(\text{H}_3\text{BO}_3)$ es la sección macroscópica de absorción del ácido bórico para neutrones térmicos presente en la muestra, promediada sobre el espectro de Maxwell; L_c es la longitud de difusión de la muestra original y D^* es un coeficiente de difusión característico donde se ha reflejado el hecho de que nuestro detector no está situado dentro de la propia muestra envenenada. Los experimentos de comprobación realizados con disoluciones de ácido bórico en agua deben conducir a:

$$\frac{1}{L_D^2} = \frac{1}{(D_G - D_W)} \cdot \Sigma_a(\text{H}_3\text{BO}_3) + \frac{1}{L_W^2} \quad (7)$$

$$\bar{\Sigma}_a(\text{H}_2\text{O}) = \frac{\bar{D}_W}{L_W^2}, \quad (8)$$

donde las magnitudes con subíndice (w) se refieren al agua pura o muestra inicial, con subíndice (D) a la solución y (G) al grafito. [13]

Los valores de secciones macroscópicas para el ácido bórico presente en la disolución fueron calculados a partir de los datos nucleares de secciones microscópicas σ_a de las bibliotecas ENDF [14] y JENDL [15] y utilizando el programa INTER [16,17]. Los valores finales de σ_a para los elementos químicos naturales y promediados sobre un espectro maxwelliano se calcularon con el programa Intermix [18].

RESULTADOS

Para la comprobación de la metodología descrita se realizó un experimento con una muestra de agua (120ml), una fuente isotópica de neutrones Pu-Be y un detector proporcional de BF_3 cuya disposición se refleja en la Figuras 2 y 3.

La fuente isotópica de Pu-Be utilizada tiene una potencia de $(1.6 \pm 0.7) \cdot 10^7$ n/s y es de forma cilíndrica ($\varnothing = 17.8$ mm, $L = 22.0$ mm). Ella se coloca dentro de un contenedor de grafito como muestra la Figura 3.

El sistema de conteo neutrónico utilizado consta de un detector proporcional gaseoso de BF_3 , modelo SNM-12 (fabricación soviética) cuyas dimensiones son las siguientes: 170 mm de longitud efectiva y 7 mm de diámetro exterior. El contador está conectado por medio de un cable coaxial de 10 m de longitud a un pre-amplificador CANBERRA modelo 2005. Después de 20 m de cable soterrado, el sistema de detección se conecta a un rack CANBERRA modelo PS 12/1 que contiene a la fuente de alimentación modelo 3106 B y un amplificador espectrométrico de investigaciones

modelo 2020. La parte final del sistema de medición se completa con un rack EUROCAR donde se ha dispuesto un analizador monocanal TSCA 301 y un scaler-timer 401. Las mediciones se realizaron para un voltaje de trabajo del detector de 536 V.

Un mapeo para la búsqueda de una posición óptima de irradiación, esto es que se garantiza un flujo bien termalizado donde son válidas las aproximaciones propuestas, fue realizado a priori. La Figura 4 presenta los resultados del estudio. Se observa, que para ciertas configuraciones, el valor calculado de la longitud de difusión para una muestra de agua se mantiene más o menos constante, al colocar el detector en diferentes posiciones detrás del contenedor, lo que fue predicho por la ecuación (4). Una vez elegida la posición óptima se realizaron las mediciones del flujo neutrónico colocando diferentes muestras de agua débilmente envenenadas con ácido bórico (0.3 % - 1 %), la Tabla 1 presenta los datos de las muestras utilizadas.

Tabla 1. Secciones macroscópicas de absorción promediadas para el ácido bórico añadido a la muestra. Fueron calculadas teniendo en cuenta las secciones microscópicas de los elementos y las densidades de las disoluciones.

Disolución	$\rho_{\text{Dis.}} \text{ (g/cm}^3\text{)}$	$\Sigma_a \text{ (ENDF)}$	$\Sigma_a \text{ (JENDL)}$
$\text{H}_3\text{BO}_3 - 0,30590 \%$	0.99380	0.0198615	0.0198496
$\text{H}_3\text{BO}_3 - 0,50222 \%$	0.99558	0.0326669	0.0326470
$\text{H}_3\text{BO}_3 - 0,70013 \%$	0.99981	0.0457337	0.0457056
$\text{H}_3\text{BO}_3 - 1,00573 \%$	0.99630	0.0654646	0.0654251

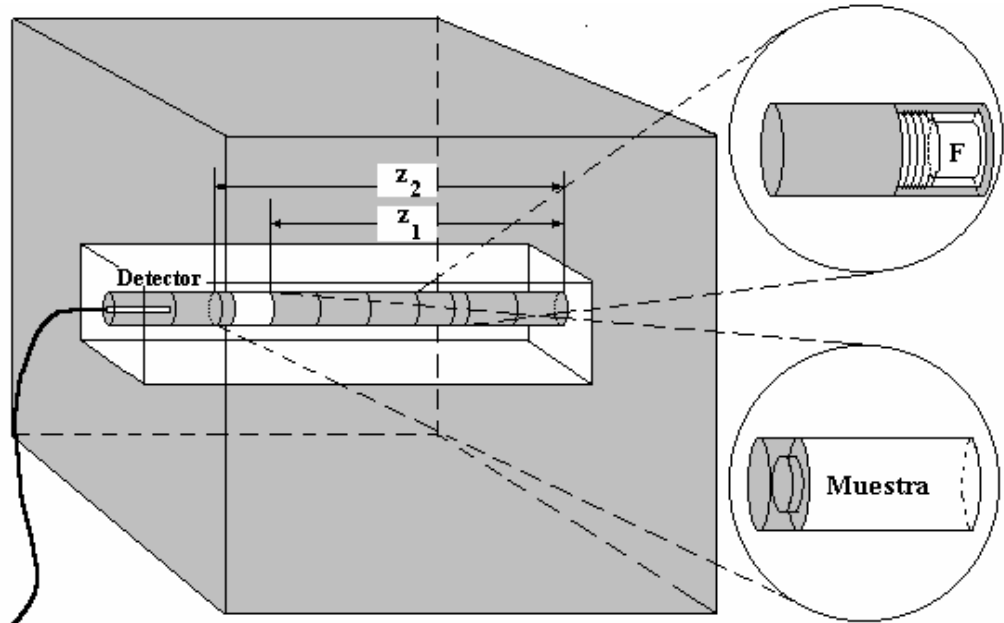


Figura 3. Esquema de las distancias geométricas necesarias y detalle de los contenedores para la fuente y la muestra.

Distribución de los valores de la L del Agua con una distancia F- M de 80 cm en el prisma de grafito

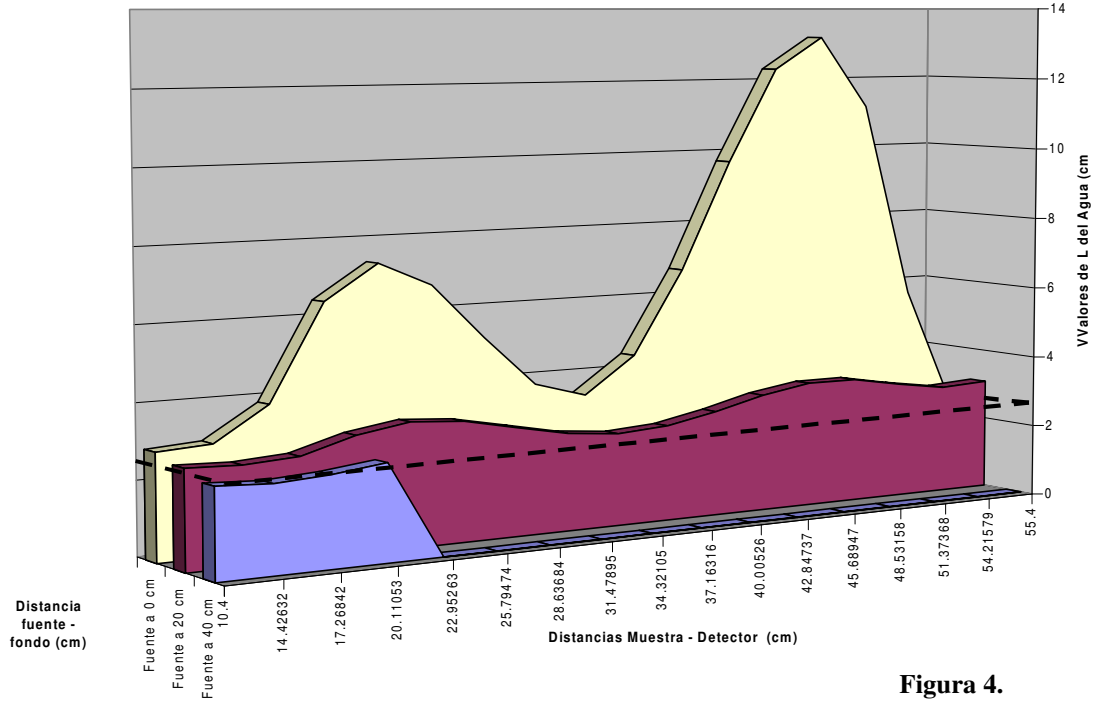


Figura 4.

Variación de la Longitud de Difusión de la Muestra de Agua medida para diferentes geometrías.

Los resultados experimentales aparecen en la Figura 5. Los ajustes lineales según la ecuación (5) tienen una buena correlación (99 %) y los errores de la pendiente y el intercepto fueron del 7.9 % y el 3 % respectivamente.

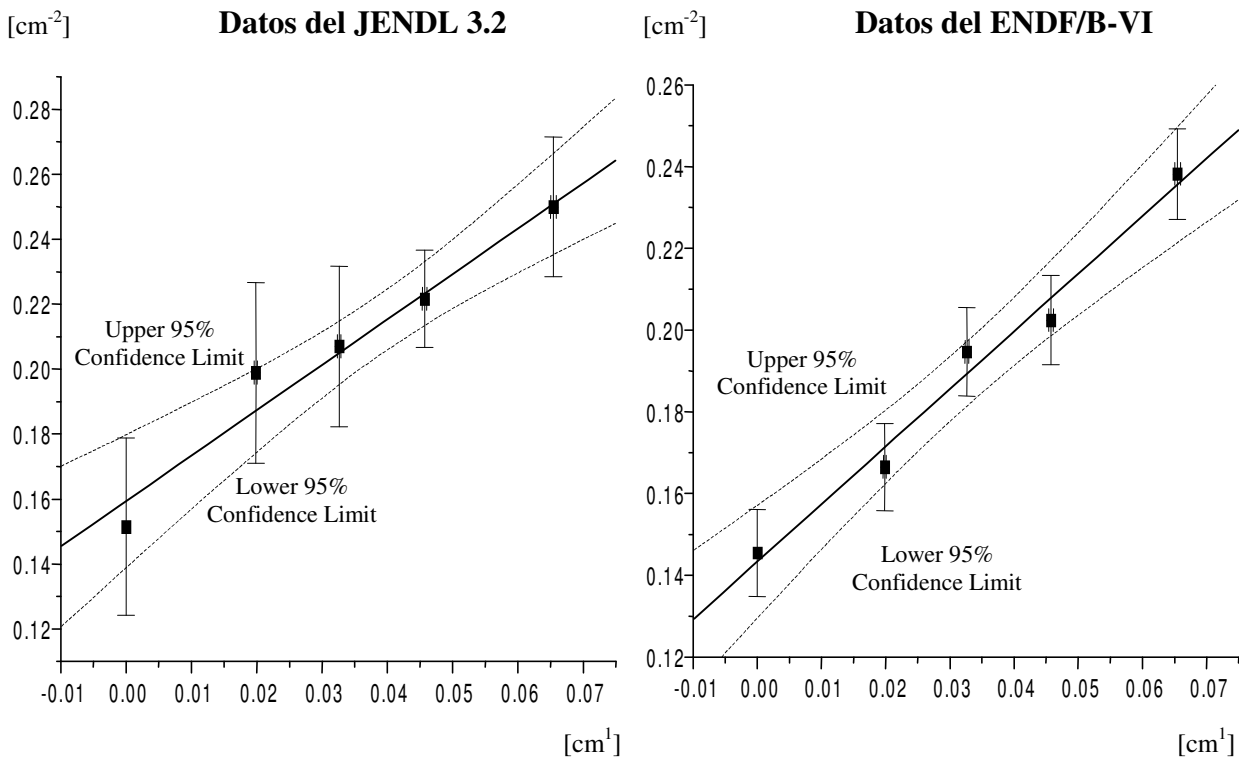


Figura 5. Dependencia de la Longitud de Difusión de la muestra ($1/L^2$) con la sección macroscópica de absorción del ácido bórico $\Sigma_a(\text{H}_3\text{BO}_3)$.

La Tabla 2 presenta los resultados finales para los valores calculados de longitud de difusión y sección macroscópica de absorción de neutrones térmicos. Aquí se han comparado con el valor medio de una serie de puntos experimentales obtenidos por otros autores, los cuales fueron sometidos a un proceso de depuración estadística. También se realizó un análisis de control estadístico de los valores obtenidos por nosotros con vistas a reconocer la calidad de los mismos. [13, 19].

Tabla 2. Mediciones de L y Σ_a para el Agua Ligera.

	Este trabajo	Reportes
Longitud de difusión	$(2.64 \pm 0.08) \text{ cm}$	$(2.76 \pm 0.06) \text{ cm}$
Sección macroscópica de absorción de neutrones	$(2.1 \pm 0.2) \times 10^{-2} \text{ cm}^{-1}$	$(1.93 \pm 0.03) \times 10^{-2} \text{ cm}^{-1}$

CONCLUSIONES

En el trabajo se muestra una metodología cuyos resultados son de interés e indican posibilidades reales de realizar mediciones exitosas de secciones macroscópicas de absorción de neutrones térmicos a partir del empleo de fuentes isotópicas estacionarias usando el modelo propuesto. La sencillez de los procedimientos y su justificación teórica permiten extenderlo a la medición de los macroparámetros neutrónicos en otros materiales. El hecho de prescindir del uso de muestras estandarizadas o certificadas previamente, para obtener una curva de calibración experimental es una ventaja que reduce los costos del experimento. Se debe tener en consideración, que la implementación en otro tipo de muestras ej.: rocas matrices, requiere un análisis adicional en cuanto a la optimización de los procedimientos de envenenamiento. Algunos trabajos de medición de secciones macroscópicas en rocas matrices ya han hecho referencia a lo anterior [20]. El empleo de las técnicas de simulación por Montecarlo, pudiera ayudar en posteriores refinamientos de la metodología sometida a consideración.

REFERENCIAS

- [1] CZUBEK, J.A. (1985): **Measurement of Macroscopic Neutron Absorption Cross Section and Others Macroparameters of Rocks**, Raport No 1281/AP. Institute of Nuclear Physics, Kraków, Poland.
- [2] _____ (1990): "Rocks Neutron Macroparameters", **Acta Geophysica Polonica**, XXXVIII(4).
- [3] _____ (1990): **Int. J. Appl. Radiat. Instrum. Part E. Nucl. Geophys**, 4(3), 293-304.
- [4] _____ et al. (1996): **Progress in Nuclear Energy**, 30(3), 295-303.
- [5] KREFT, A.; A. BOLEWSKI Jr. and M. CIECHANOWSKI (1989): **Int. J. Appl. Radiat. Instrum. Part E. Nucl. Geophys**, 3(4), 367-372.
- [6] HOOVER, J.I. et al. (1948): **Physical Review** 74(8), 864-869.
- [7] KREFT, A.; A. BOLEWSKI Jr., M. CIECHANOWSKI and K. MORSTIN (1989): **Int. J. Appl. Radiat. Instrum. Part E. Nucl. Geophys**, 3(4), 361-366.
- [8] ANDERSEN, E. and G. LOVHOLDEN (1988): **J. Phys. D: Appl. Phys.**, 21, 1550-1555.
- [9] SOOD, A.; R.P. GARDNER and T.K. GRAY (2000): **Appl. Radiat. Isot.**, 53, 603-616.
- [10] TITTLE, C.W. and G.C. CRAWFORD (1984): **SPWLA, 24 Annual Logging Symposium**, Paper I.
- [11] BECKURTS, K.H. and K. WITRZ (1964): "Neutron Physics", **Kernforschungszentrum Karlsruhe**. By Springer-Verlag OHG. Berlin. Göttingen. Heidelberg. New York.
- [12] KASH, S.W. and D.C. WOOD (1953): **Physical Review** 90, 564.
- [13] BORREGO MOREJÓN, Y. (2000): **Tesis de Diploma**, ISCTN, La Habana, Cuba.
- [14] ENDF /B-VI (1999): **The U.S. Evaluated Nuclear Data Library for Nuclear Reaction Data**. By the U.S. National Nuclear Data Center – 1990 including revisions up to Aug. 1999.

- [15] JENDL 3.2 (1994): **The Japanese Evaluated Nuclear Data Library**, By the JAERI Nuclear Data Center. Second revision, Aug.
- [16] DERMONT CULLEN, E. (1998): "The 1998 ENDF Pre-processing Codes". University of California, Lawrence Livermore National Laboratory, USA.
- [17] DUNFORD, C.L. (1998, April 10): "ENDF utility Codes Release 6.111".
- [18] LÓPEZ-ALDAMA, D. (2000): **Program Intermix.exe**, ISCTN, La Habana, Cuba. Comunicación privada.
- [19] HERRERA PERAZA, E. (1995): Tesis de Doctorado. ISCTN, La Habana, CUBA.
- [20] CZUBEK, J.A. **et al.** (1986): **Methodology of preparation of rock samples for measurement of the thermal neutron macroscopic absorption cross-section**, Raport No 1320/AP. Institute of Nuclear Physics, Kraków, Poland.

ESTUDIO TEÓRICO DE VARIAS SUCESIONES CUASIRREGULARES BINARIAS, TERNARIAS Y CUATERNARIAS

R. Betancourt Riera, W. Pardo Tamayo y R. Pérez-Álvarez
Departamento de Física Teórica, Universidad de La Habana, Cuba

RESUMEN

Analizamos varias sucesiones binarias, ternarias y cuaternarias, y estudiamos el carácter del espectro para problemas simples. Se demuestra el carácter Pisot, así como que los espectros de todos estos problemas son singulares continuos. Se reportan resultados numéricos para el caso de electrones en el marco de la Teoría de Masa Efectiva de una de estas sucesiones que resulta tener propiedades similares a la Rudin-Shapiro.

ABSTRACT

We analyze some binary, ternary and quaternary sequences and analyse the character of the spectrum for simple problems. We prove the Pisot character of these sequences. It is shown that the spectra of all these problems are singular continuous. A numerical test is given for the case of electrons in the framework of the Effective Mass Theory for one of these sequences and it is shown that this sequence is similar to the Rudin-Shapiro one.

INTRODUCCIÓN

Desde principios de la década de 1970 se estudian con gran interés los sistemas regulares a muchas capas [1-5] (Pozos Cuánticos, Superredes, etc.). En los últimos años se ha consolidado el estudio de las llamadas Heteroestructuras Quasiregulares (QH: *Quasiregular Heterostructures*) [6-12]. Estas consisten en una sucesión de capas construida de acuerdo con cierta regla de sustitución, de manera que resulta un sistema no periódico pero con un algoritmo preciso de secuenciación de las distintas capas (QS: *Quasiregular Sequences*). Entre las QSs más estudiadas están las de Fibonacci, Thue Morse, Circular, Rudin-Shapiro, etc. La de Fibonacci, por ejemplo, se forma según la regla de sustitución

$$\begin{aligned}x_{\text{FIBO}}(A) &= AB \\x_{\text{FIBO}}(B) &= A\end{aligned}\quad (1)$$

Una de las propiedades más interesantes de estas estructuras es que muchas excitaciones elementales poseen espectro singular continuo, a diferencia, por ejemplo, de los pozos cuánticos en que típicamente es discreto, más eventualmente un continuo, y de las superredes donde es a intervalos o bandas. Bajo condiciones bastante generales esta propiedad se ha demostrado en varias de estas QSs gracias a un Teorema debido a A Bovier y JM Ghez; la Rudin Shapiro no cumple las hipótesis del teorema de Bovier-Ghez [6, 11, 12].

Para estudiar estos sistemas se han introducido los conceptos de Alfabeto, Matriz Característica, Razón Áurea, y otros, que damos por conocidos y que se pueden ver en las referencias citadas, particularmente en [6, 11-15].

Con la vista fija en la posibilidad de obtener nuevas propiedades, en el presente trabajo analizamos varias QS binarias, ternarias y cuaternarias. La elección específica se debe a que no han sido estudiadas hasta el momento. Estudiamos su carácter Pisot. Seguidamente argumentamos teóricamente el hecho de que los espectros de problemas simples sobre tales estructuras son singulares continuos. Por último formulamos algunas conclusiones. Mención especial merece una de ellas, que llamaremos RBR, y que no satisface las hipótesis del Teorema de Bovier-Ghez [11,12] pero que sin embargo tiene espectro singular continuo. Reportamos entonces resultados numéricos para espectros de electrones de conducción que ratifican esta conclusión.

QS Binarias

Las sucesiones que nos interesan se definen por las siguientes reglas de sustitución:

$$\begin{aligned}x_1(A) &= ABA \\x_1(B) &= A\end{aligned}\quad (2)$$

$$\begin{aligned}x_2(A) &= AB \\x_2(B) &= ABA\end{aligned}\quad (3)$$

Ambas tienen por alfabeto a $\{A, B\}$. Las matrices características son, respectivamente

$$M_1 = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \quad (4)$$

$$M_2 = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \quad (5)$$

De aquí es inmediato que las dos sucesiones son primitivas. Además, en el caso de la regla x_1 los autovalores y autovectores son:

$$l_1 = 1 - \sqrt{2}; l_2 = 1 + \sqrt{2} \quad (6)$$

$$v_1 = \begin{bmatrix} q - \sqrt{2} \\ 1 \end{bmatrix}; v_2 = \begin{bmatrix} q + \sqrt{2} \\ 1 \end{bmatrix}$$

Por su parte, para la segunda de las sucesiones estudiadas, o sea x_2 :

$$l_1 = 1 - \sqrt{2}; l_2 = 1 + \sqrt{2} \quad (7)$$

$$v_1 = \begin{bmatrix} -\sqrt{2} \\ 1 \end{bmatrix}; v_2 = \begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix}$$

En ambos casos el segundo autovalor es la razón áurea y el correspondiente autovector da la relación entre las poblaciones de las letras en la sucesión. Además, este cálculo indica que ambas sucesiones tienen la propiedad Pisot por lo que son derivables por corte y proyección de una red de dimensión superior [9,10].

Introduzcamos las siguientes notaciones para las trazas de las Matrices de Transferencia de las Cadenas Canónicas [6]:

$$x_n = \text{Tr}[A_n]; y_n = \text{Tr}[B_n]; z_n = \text{Tr}[B_n A_n] \quad (8)$$

Entonces la primera de las sucesiones estudiadas (x_1) tiene asociado el mapa de trazas

$$\begin{aligned} x_{n+1} &= x_n z_n - y_n \\ y_{n+1} &= x_n \end{aligned} \quad (9)$$

$$z_{n+1} = x_n^2 z_n - x_n y_n - z_n$$

mientras que en el segundo caso (x_2)

$$\begin{aligned} x_{n+1} &= z_n \\ y_{n+1} &= x_n z_n - y_n \\ z_{n+1} &= x_n z_n^2 - y_n z_n - x_n \end{aligned} \quad (10)$$

No hace falta aburrir al lector con detalles insustanciales. De estos mapas de trazas se deduce sin mucho esfuerzo que en los dos casos se satisfacen las hipótesis del Teorema de Bovier-Ghez, con lo cual se puede asegurar sin más que las excitaciones elementales definidas sobre estas estructuras tienen espectro singular continuo.

Una QS Ternaria

Con idénticos objetivos de estudio veamos la siguiente QS ternaria

$$\begin{aligned} x_3(A) &= AB \\ x_3(B) &= C \\ x_3(C) &= CA \end{aligned} \quad (11)$$

cuya matriz característica es

$$M_3 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \quad (12)$$

El cubo de esta matriz es una matriz positiva. De aquí se desprende que la matriz y la QS son primitivas. Sus autovalores son

$$\begin{aligned} l_1 &= 1,75488 L \\ l_2 &= 0,12 L + i0,74 L \\ l_3 &= 0,12 L - i0,74 L \end{aligned} \quad (13)$$

por lo que esta QS tiene carácter Pisot. El primero de estos autovalores es obviamente la razón áurea. Los autovectores por su parte son:

$$\begin{aligned} v_1 &= \begin{bmatrix} 1,32L \\ 0,75L \\ 1 \end{bmatrix} \\ v_2 &= \begin{bmatrix} -0,66 L - i0,56L \\ -0,87 L + i0,74L \\ 1 \end{bmatrix} \\ v_3 &= \begin{bmatrix} -0,66 L + i0,56L \\ -0,87 L - i0,74L \\ 1 \end{bmatrix} \end{aligned} \quad (14)$$

De aquí se deduce que en la QS las letras están en la relación 1,32..., 0,75..., 1. Definiendo las trazas como

$$x_n = \text{Tr}[A_n]; y_n = \text{Tr}[B_n]; z_n = \text{Tr}[C_n]$$

$$l_2 = 1$$

$$w_n = \text{Tr}[B_n]; r_n = \text{Tr}[A_n C_n] \quad (15)$$

$$l_3 = -0,57 L + i0,36 L \quad (19)$$

$$s_n = \text{Tr}[B_n C_n]; v_n = \text{Tr}[C_n B_n A_n]$$

$$l_4 = 0,57 L - i0,36 L$$

el mapa de trazas queda

$$\begin{aligned} x_{n+1} &= w_n \\ y_{n+1} &= z_n \\ z_{n+1} &= r_n \\ w_{n+1} &= v_n \\ r_{n+1} &= x_n v_n - s_n \\ s_{n+1} &= z_n r_n - x_n \\ v_{n+1} &= x_n (z_n v_n - w_n) - z_n s_n + y_n \end{aligned} \quad (16)$$

y de aquí se deduce el carácter singular continuo del espectro de excitaciones elementales.

Otros muchos ejemplos de sucesiones binarias y ternarias siguen este mismo paradigma, o sea son primitivas, tienen carácter Pisot, sus matrices características dan la razón áurea y las proporciones de letras sin mayores dificultades y el análisis de los mapas de trazas da como resultado que se cumplen las hipótesis del Teorema de Bovier-Ghez lo que lleva sin más a que el espectro de muchas excitaciones elementales es singular continuo. En [15] hay una veintena de ejemplos de QS binarias y ternarias que agotan las QS de este tipo con índice de inflación pequeño.

Una QS cuaternaria que no satisface las hipótesis del Teorema de Bovier-Ghez

Estudiemos ahora la siguiente regla

$$\begin{aligned} x_4(A) &= AC \\ x_4(B) &= DB \\ x_4(C) &= ADB \\ x_4(D) &= AB \end{aligned} \quad (17)$$

cuya matriz es

$$M_4 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \quad (18)$$

y que llamaremos RBR por las iniciales de uno de nosotros. Es inmediato que esta QS es primitiva pues M^3 no tiene ceros. Su razón áurea es $l_1 = 2,1479 L$. Los restantes autovalores son

por lo que no tiene la propiedad Pisot. El autovector más relevante, o sea el asociado al autovalor de la razón áurea es

$$v_1 = \begin{bmatrix} 1,46L \\ 1,46L \\ 0,68L \\ 1 \end{bmatrix} \quad (20)$$

de donde se infiere que las letras A, B, C y D están en la relación 1,46..., 1,46..., 0,68..., y 1.

Definamos las trazas

$$\begin{aligned} x_n &= \text{Tr}[A_n]; y_n = \text{Tr}[B_n]; z_n = \text{Tr}[C_n] \\ w_n &= \text{Tr}[D_n]; r_n = \text{Tr}[A_n B_n] \\ s_n &= \text{Tr}[B_n C_n]; v_n = \text{Tr}[C_n A_n] \\ h_n &= \text{Tr}[B_n D_n]; g_n = \text{Tr}[D_n A_n] \\ q_n &= \text{Tr}[C_n D_n]; l_n = \text{Tr}[C_n B_n D_n] \\ p_n &= \text{Tr}[B_n D_n A_n]; t_n = \text{Tr}[C_n A_n B_n] \\ k_n &= \text{Tr}[A_n B_n D_n C_n] \end{aligned} \quad (21)$$

El mapa de trazas derivado es

$$\begin{aligned} x_{n+1} &= v_n \\ y_{n+1} &= h_n \\ z_{n+1} &= p_n \\ w_{n+1} &= r_n \\ r_{n+1} &= k_n \\ s_{n+1} &= h_n p_n - x_n \\ v_{n+1} &= v_n p_n + l_n - z_n h_n \\ h_{n+1} &= h_n r_n + g_n - w_n x_n \\ g_{n+1} &= v_n r_n + s_n - z_n y_n \\ q_{n+1} &= p_n r_n - w_n \\ l_{n+1} &= p_n (h_n r_n + g_n - w_n x_n) - y_n \\ p_{n+1} &= h_n (v_n r_n + s_n - z_n y_n) + v_n g_n + q_n - w_n x_n v_n \end{aligned}$$

$$t_{n+1} = h_n(v_n p_n + l_n - z_n h_n) - x_n v_n + z_n$$

$$k_{n+1} = (h_n r_n + g_n - w_n x_n)(v_n p_n + l_n - z_n h_n) + x_n(v_n r_n + s_n - z_n y_n) - t_n - r_n(x_n v_n - z_n)$$

La regla de sustitución auxiliar es

$$\begin{aligned} \phi(x) &= v; & \phi(y) &= v; & \phi(z) &= p \\ \phi(w) &= r; & \phi(r) &= k; & \phi(s) &= hp \\ \phi(v) &= vp; & \phi(h) &= hr \\ \phi(g) &= vr; & \phi(q) &= pr \\ \phi(l) &= phr; & \phi(p) &= hvr \\ \phi(t) &= hvp; & \phi(k) &= hrvp \end{aligned} \quad (22)$$

con el alfabeto

$$B = \{x, y, z, w, r, s, v, h, g, l, q, p, t, k\}$$

Las hipótesis del teorema de Bovier-Ghez son tres, a saber:

1. Que la QS sea primitiva.
2. Que la Regla de Sustitución Auxiliar f sea semiprimitiva en algún subconjunto C del Léxico.
3. Que $x^k(0)$ para $k < \infty$ que contenga el cuadrado de alguna letra o palabra $b \in B$.

Vemos que de estas tres condiciones sólo se satisface la primera, puesto que nuestra QS es primitiva. La segunda condición es imposible de cumplir ya que ϕ no es semiprimitiva. Si desarrollamos ϕ en cualquier subconjunto $C \in B$ vemos que aunque existe un k para el cual $\phi^k(C_i)$ contiene todas las letras de C aparecen nuevas letras de B que no pertenecen a C . Esto último nos indica que ϕ no será primitiva en C , ya que el hecho de aparecer una nueva letra cambia su alfabeto y en consecuencia el orden de la matriz característica asociada a ϕ en el subconjunto. Esto último nos demuestra que ϕ no es semiprimitiva y por tanto no se cumplen para esta regla de sustitución las proposiciones 3.1 y 3.2 dadas en [11,12], por lo que podemos concluir que el primer teorema de Bovier-Ghez no se cumple.

El incumplimiento de este teorema no implica que podamos llegar a conclusiones sobre la naturaleza del espectro de la QH. Las condiciones que impone el teorema 1 de Bovier-Ghez son sólo condiciones suficientes para que el espectro sea singular continuo, por lo que se hace necesario otro tipo de análisis para establecer la naturaleza del espectro.

Espectro de electrones en una QS de tipo RBR

Para discriminar el carácter del espectro de la QS cuaternaria estudiada, hagamos un experimento numérico. Usando el mapa de trazas o la regla de sustitución para las matrices de transferencia es relativamente fácil calcular el espectro [6]. En la Figura adjunta se presenta la IDOS (*Integrated Density of States*) de electrones de conducción para una octava generación de una QS RBR donde las capas A y C están constituidas por el material $Al_{0.35}Ga_{0.65}As$, mientras que B y D están formadas por GaAs sin mezclar. A y C tienen tres capas atómicas mientras que B y D tienen seis. Esto hace que las primeras tengan un ancho de 8,4 Å y las segundas 16,8 Å. Los datos de los materiales y demás parámetros del cálculo se pueden ver en [15].

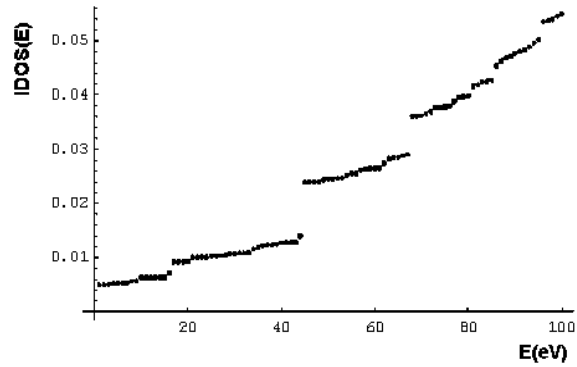


Figura 1. Densidad Integrada de Estados (IDOS) de una QS de tipo RBR.

Este espectro tiene aspecto de escalera del Diablo. Para comprobar su fractalidad calculamos su dimensión $D(q)$ para varios valores representativos del índice q . En la Tabla adjunta se pueden ver los resultados.

Tabla 1. Dimensión fractal $D(q)$ para algunos valores de q de una QS de tipo RBR.

q	$D(q)$
0.0	0.76
1.0	0.69
2.0	0.60
3.0	0.53

Estos números indican claramente el carácter multifractal de este espectro. En [15] se pueden ver más detalles de los análisis tanto teóricos como numéricos presentados aquí.

CONCLUSIONES

Hemos analizado varias QS binarias y ternarias relativamente simples. Hemos demostrado que gozan de la propiedad Pisot lo que las hace derivables de una red en una dimensión superior por el método de corte y proyección [6-10]. Reportamos sus mapas de trazas y los principales momentos de

la aplicación del teorema de Bovier-Ghez. De aquí se deriva sin más que el espectro de problemas simples sobre estas QS es singular continuo. Hasta el momento no hay estudios experimentales de estas sucesiones concretas aunque sí en sucesiones muy parecidas [6, 14] cuyo estudio teórico ya ha sido reportado [6, 14, 15].

Este mismo análisis en el ejemplo de la QS que hemos llamado RBR es distinto. Esta QS resulta primitiva, no Pisot, y no satisface las hipótesis del teorema de Bovier-Ghez. No obstante, el cálculo numérico de electrones en Teoría de Masa Efectiva no deja lugar a dudas acerca de la multifractalidad

del espectro. Esto último implica que el espectro no es exactamente autosimilar ya que tiene una distribución no constante de exponentes de escalado.

En [6] se reportan análisis semejantes a los llevados a cabo en el presente trabajo pero aplicados a la sucesión de Rudin-Shapiro. La comparación con los resultados del presente trabajo lleva a que ambas sucesiones tienen las mismas características. Entonces, presumiblemente estas dos QS son ejemplares de toda una clase de QS con propiedades peculiares.

REFERENCIAS

1. ESAKI, L. (1984): **J. Physique Colloques**, 45, C5-3.
2. BASTARD, G. (1989): "Wave mechanics applied to semiconductor heterostructures", Éditions de Physique, Paris.
3. VINTER, B. and C. WEISBUCH (1991): "Quantum Semiconductor Structures", Academic Press, San Diego.
4. GARCÍA-MOLINER, F. y otros (1994): "Electrones y Fonones en Pozos Cuánticos", editado por la Universidad Nacional de Educación a Distancia (UNED), Madrid, España.
5. GELLER, M.R. (2001): "Quantum phenomena in low-dimensional systems", <http://www.arXiv:cond-mat/0106256>.
6. PÉREZ-ÁLVAREZ, R. and F. GARCÍA-MOLINER (2001): "Quasiregular Heterostructures", capítulo invitado en "Contemporary Problems of the Condensed Matter Physics", Nova Science Publishers, ed. por S. Vlaev y M. Gaggero-Sager.
7. MACIA, E. and F. DOMÍNGUEZ-ADAME (2000): "Electrons, phonons and excitons in low dimensional aperiodic systems", Madrid, Editorial Complutense.
8. International Workshop on Aperiodic Crystals, March 11-20, (1986): Les Houches, France. Edited by D Gratias and L Michel. **Journal de Physique**, 47, Colloque C3, supplément au No 7 (Juillet 1986).
9. BOMBIERI, E. and J.E. TAYLOR (1986): "Which distributions of matter diffract? An initial investigation", **J. de Physique**, 47, Colloque C3, supplément au n° 7, C3-19-c3-28.
10. _____ (1987): "Quasicrystals, tilings, and algebraic number theory; some preliminary connections", **Contemporary Mathematics** 64, 241-264.
11. BOVIER, A. and J.M. GHEZ (1993): **Commun. Math. Phys.** 58(1), 45.
12. _____ (1995): **J. Phys. A: Math. Gen.** 28, 2313.
13. KOLAR, M. et al. (1993): **J. Phys. A: Math. Gen.** 26, 7343.
14. ZHU, Y-y and N-b Ming (1990): **Phys. Rev. B** 42:6, 3676.
15. BETANCOURT RIERA, R. (2003): Tesis de Maestría, Universidad de La Habana.

MATRIZ CARACTERÍSTICA Y RAZÓN ÁUREA DE SUCESIONES CUASIRREGULARES

R. Pérez-Álvarez, R. Betancourt Riera, y W. Pardo Tamayo

Departamento de Física Teórica, Universidad de La Habana, Cuba

RESUMEN

De una manera completamente general escribimos la relación de recurrencia entre las poblaciones de letras en dos generaciones sucesivas en términos de la Matriz Característica (CM). En el límite cuando la generación es infinito, obtenemos que la razón áurea es autovalor de la CM; las componentes del autovector correspondiente resultan ser las proporciones en las que se encuentran las diferentes letras en la sucesión.

ABSTRACT

In a general manner we write down the recurrence relation for the population of the different letters in a Quasiregular Sequence in terms of the Characteristic Matrix (CM). In the limit as the generation becomes infinite we obtain that the Golden Ratio is an eigenvalue of the CM. The components of the corresponding eigenvector are the proportions in which the different letters enter in the sequence.

INTRODUCCIÓN

Continuando con el auge de los sistemas a muchas capas [1-5] (Pozos Cuánticos, Superredes, etc.), en los últimos años se ha consolidado el estudio de las llamadas Heteroestructuras cuasiregulares (QH: *Quasiregular Heterostructures*) [6-12]. Estas consisten en una sucesión de capas construida de acuerdo con cierta regla de sustitución, de manera que resulta un sistema no periódico pero con un algoritmo preciso de secuenciación de las distintas capas (QS: *Quasiregular Sequences*). Entre las QSs más estudiadas están las de Fibonacci, Thue Morse, Rudin-Shapiro, etc. La de Fibonacci, por ejemplo, se forma según la regla de sustitución

$$\begin{aligned}\xi_{\text{FIBO}}(A) &= AB \\ \xi_{\text{FIBO}}(B) &= A\end{aligned}\quad (1)$$

Una de las características más interesantes de estas estructuras es que muchas excitaciones elementales poseen espectro singular continuo, a diferencia, por ejemplo, de los pozos cuánticos en que típicamente es discreto, y de las superredes donde es a bandas. Bajo condiciones bastante generales esta propiedad se ha demostrado en varias de estas QSs aunque la demostración no es válida para la Rudin Shapiro [6, 11, 12].

Entre las magnitudes que han demostrado su utilidad para describir a las QSs están la Matriz Característica, la proporción de letras y la razón áurea.

La Matriz Característica (CM) es aquella matriz que en su fila i y columna j tiene un entero no

negativo igual al número de veces que está la letra j -ésima en la regla i -ésima. En casos sencillos este procedimiento puede generar sucesiones periódicas; un ejemplo de esto es

$$\begin{aligned}\xi_{\text{per}}(A) &= AB \\ \xi_{\text{per}}(B) &= A\end{aligned}\quad (2)$$

con CM

$$M_{\text{per}} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\quad (3)$$

Pero también puede generar sucesiones no periódicas. Por ejemplo, para la sucesión de Fibonacci

$$M_{\text{FIBO}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\quad (4)$$

Esta representación matricial sigue la descripción usada en [13]. Otras descripciones también son posibles, destacando entre ellas la proyección desde una estructura periódica de orden superior [14] y las recurrencias progresivas y regresivas [15]. Las generaciones posteriores a la cuarta son diferentes pero el número total de elementos sigue la ley de Fibonacci.

El objetivo del presente trabajo es demostrar que de la CM se puede extraer de manera sencilla la información acerca de la razón áurea y las proporciones en que una letra aparece en la QS.

Relaciones de Recurrencia de la Población de Letras

Llamemos $F_n^{(i)}$ a la cantidad de letras de tipo i que aparecen en la generación n -ésima. Entonces es evidente que entre estas magnitudes existe la siguiente relación de recurrencia

$$F_{n+1}^{(i)} = \sum_j M_{ji} F_n^{(j)} \quad (5)$$

Dividamos ahora ambos miembros por la población de una letra específica arbitrariamente escogida; la k -ésima, por ejemplo. Tendremos entonces

$$\frac{F_{n+1}^{(i)}}{F_n^{(k)}} = \sum_j M_{ji} \frac{F_{n+1}^{(j)}}{F_n^{(k)}} \quad (6)$$

Introduzcamos ahora las magnitudes

$$\tau_n^{(i)} = \frac{F_{n+1}^{(i)}}{F_n^{(i)}}; \quad f_n^{jk} = \frac{F_n^{(j)}}{F_n^{(k)}} \quad (7)$$

La primera de ellas es la relación entre las poblaciones de la letra i en la generación n y la siguiente. La segunda magnitud introducida es la proporción de cada letra respecto de una de ellas, la k -ésima en este caso, siempre en la generación n . Definamos, además, los límites de estas magnitudes cuando la generación en cuestión se hace muy grande, esto es

$$\tau^{(i)} = \lim_{n \rightarrow \infty} \tau_n^{(i)}; \quad f^{jk} = \lim_{n \rightarrow \infty} f_n^{jk} \quad (8)$$

La primera de estas dos es la llamada relación o razón áurea. La segunda es simplemente la relación o proporción entre las letras de una sucesión. Con estas notaciones, la relación de recurrencia entre las poblaciones de letras en dos generaciones sucesivas, cuando $n \rightarrow \infty$, se lee

$$\sum_j [M_{ji} - \tau^{(i)} \delta_{ij}] f^{jk} = 0 \quad (9)$$

Y aquí se tiene el primer resultado interesante de este análisis: la razón áurea es autovalor de la traspuesta de la CM (o autovalor por la izquierda de la CM). Además, el autovector correspondiente da las poblaciones relativas de las distintas letras.

Es evidente que

$$\tau_n^i = f_{n+1}^{ik} f_n^{ki} \tau_n^k \quad (10)$$

y por lo tanto todas las letras tienen la misma relación o razón áurea en la sucesión, o sea en la generación de índice infinito. Esta reflexión lleva a pensar que

no todos los autovalores de la CM tienen interpretación de razón áurea. Esto, por otra parte, es evidente del hecho de que la demostración hecha concluye que la razón áurea es autovalor de la CM, pero no lo contrario, o sea que habrá otros autovalores de la CM que no tengan esta interpretación.

Ejemplos

El caso de Fibonacci es elocuente. Los autovalores de la CM son

$$\lambda_1 = \frac{1-\sqrt{5}}{2}; \quad \lambda_2 = \frac{1+\sqrt{5}}{2} \quad (11)$$

mientras que el autovector que da las proporciones de letras es

$$v_2 = \begin{pmatrix} \tau \\ 1 \end{pmatrix} \quad (12)$$

El otro autovector tiene componentes 1 y $-\tau$.

Las QS de Thue-Morse, period doubling, circular, binary non-Pisot, ternary non-Pisot, e incluso la más complicada Rudin-Shapiro son ejemplos en que esta manera de calcular las razones áureas y proporciones es mucho más cómoda que la utilizada hasta el momento, consistente en escribir las relaciones de recurrencia y hacer maniobras algebraicas largas y costosas. En estos casos, al igual que en el de Fibonacci, hay un único autovector con todas sus componentes positivas y por lo tanto el único susceptible de ser interpretado como proporción de la presencia de las distintas letras.

Un ejemplo algo más raro es la QS

$$\begin{aligned} \xi(A) &= B \\ \xi(B) &= BBB \end{aligned} \quad (13)$$

con CM

$$M = \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix} \quad (14)$$

Lo raro está en que la sucesión no tendrá Aes. Los autovalores y autovectores son, respectivamente

$$\lambda_1 = 0; \quad \lambda_2 = \tau = 3 \quad (15)$$

$$v_1 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}; \quad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Y este resultado confirma que la razón áurea de la letra B es 3, a la A se le puede asignar esta misma razón áurea, y la relación entre las Bes y las Aes es de uno a cero.

Algo más *rara* aún es la QS

$$\begin{aligned}\xi(A) &= A \\ \xi(B) &= ABB\end{aligned}\quad (16)$$

con

$$M = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \quad (17)$$

y

$$\lambda_1 = \tau^{(1)} = 1; \quad \lambda_2 = \tau^{(2)} = 2 \quad (18)$$

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (19)$$

La interpretación es trivial. Si la semilla es A, o cualquier palabra que sólo contenga Aes, $\tau^{(1)}$ es la razón áurea de todas las letras y el autovector v_1 da la relación entre el número de Aes y el de Bes. Si la semilla tiene sólo Bes, o ambas letras, el segundo autovalor y su correspondiente autovector contienen estas informaciones.

La relación de recurrencia entre dos generaciones sucesivas es la misma cualquiera sea la semilla o palabra inicial; no tenemos que estar hablando necesariamente de una cadena canónica. Esto no implica que cualquiera sea la semilla obtendremos la misma razón áurea, como acabamos de mostrar.

CONCLUSIONES

Hemos demostrado que el proceso de diagonalización de la Matriz Característica permite obtener de una manera simple y expedita la razón áurea de cualquier Sucesión Cuasirregular. Adicionalmente, obtenemos las proporciones en que están presentes las diferentes letras en la Sucesión. Incluso en casos de QSs no primitivas este resultado es correcto; el segundo y tercer ejemplo desarrollados en el presente trabajo dan fe de ello. Esta interpretación de los autovalores y autovectores de la CM, aparentemente anodina, puede tener una importancia capital en aplicaciones pues nos dice la cantidad de cada material que contendrá el sistema y, por lo tanto su costo en materia prima. En los trabajos [16, 17] se pueden ver más casos concretos de lo que aquí hemos demostrado.

REFERENCIAS

1. ESAKI, L. (1984): **J. Physique Colloques**, 45, C5-3.
2. BASTARD, G. (1989): "Wave mechanics applied to semiconductor heterostructures", Éditions de Physique, Paris.
3. VINTER, B. and C. WEISBUCH (1991): "Quantum Semiconductor Structures", Academic Press, San Diego.
4. GARCÍA-MOLINER, F. y otros (1994): "Electrones y Fonones en Pozos Cuánticos", editado por la Universidad Nacional de Educación a Distancia (UNED), Madrid, España.
5. GELLER, M.R. (2001): "Quantum phenomena in low-dimensional systems", <http://www.arXiv:cond-mat/0106256>.
6. PÉREZ-ÁLVAREZ, R. and F. GARCÍA-MOLINER (2001): "Quasiregular Heterostructures", capítulo invitado en "Contemporary Problems of the Condensed Matter Physics", Nova Science Publishers, ed. por S. Vlaev y M. Gaggero-Sager.
7. MACIA, E. and F. DOMÍNGUEZ-ADAME (2000): "Electrons, phonons and excitons in low dimensional aperiodic systems", Madrid, Editorial Complutense.
8. International Workshop on Aperiodic Crystals, March 11-20, (1986): Les Houches, France. Edited by D Gratias and L Michel, **Journal de Physique**, 47, Colloque C3, supplément au No 7 (Juillet).
9. BOMBIERI, E. and J.E. TAYLOR (1986): "Which distributions of matter diffract? An initial investigation", **J. de Physique**, 47, Colloque C3, supplément au n° 7, C3-19-c3-28.
10. _____ (1987): "Quasicrystals, tilings, and algebraic number theory; some preliminary connections", **Contemporary Mathematics** 64, 241-264.
11. BOVIER, A. and J.M. GHEZ (1993): **Commun. Math. Phys.** 58(1), 45.

12. BOVIER, A. and J.M. GHEZ (1995): **J. Phys. A: Math. Gen.** 28, 2313.
13. LU, J.P. y otros (1986): **Phys. Rev. B** 33, 915.
14. MAC DONALD, A.H. (1987) en *Interface, Quantum Wells and Superlattices*, ed. por C.R. Leavens y R. Taylor, Plenum, New York.
15. WANG, Y.Y. and J.C. MAAN (1989), **Phys. Rev. B** 40, 1955.
16. KOLAR, M. **et al.** (1993): **J. Phys. A: Math. Gen.** 26, 7343.
17. ZHU, Y-y and N-b Ming (1990): **Phys. Rev. B** 42:6, 3676.

CALOR Y TRABAJO EN LA ENSEÑANZA DE LA TERMODINÁMICA

A. González Arias, Dpto. Física Aplicada, Universidad de La Habana

RESUMEN

Calor y Trabajo se encuentran en la base de la Termodinámica. Ambos conceptos son parte integrante de la Primera Ley, y el primero de ellos también está asociado a la Segunda Ley. Sería de esperar, por tanto, que ambos términos estuvieran muy bien definidos en la literatura docente, sin dar lugar a dudas o ambigüedades. En la práctica sucede todo lo contrario. De aquí que se examinen en detalle las definiciones de calor y trabajo proporcionadas por varios textos conocidos de Física General y Termodinámica. El procedimiento conduce a la introducción de dos definiciones razonablemente simples, modernas e independientes, y que no desmerecen en rigurosidad. Adicionalmente se destaca el hecho de que expresiones bastante extendidas, tales como "el sistema hace trabajo" y "se hace trabajo sobre el sistema" equivalen a utilizar indistintamente, de manera confusa y poco satisfactoria, dos referencias diferentes para describir diferentes etapas de un mismo proceso.

ABSTRACT

Heat and Work are in the basement of Thermodynamics. Both terms are included in the First Law, and the first term is also associated to the Second Law. Therefore, it would be reasonable to expect that both expressions would be very well defined in the common Thermodynamic textbooks, without giving rise to doubts or misconceptions. In practice the entire contrary happens. Due to this reason, the definitions of work and heat provided in several well-known textbooks are analyzed here. This procedure leads to the introduction of two reasonably simple, modern and independent definitions, without lack of rigorousness. Additionally, it is discussed the fact that widespread expressions such as "the system does work" and "work is done on the system" means that we are indistinctly using, in a confusing and no satisfactory way, two different reference frames to describe different stages of the same process.

INTRODUCCIÓN

La Termodinámica se imparte usualmente, de una u otra forma y con mayor o menor profundidad, en todas las especialidades asociadas a las Ciencias de la Vida y de la Tierra, en todas -o en casi todas- las Ingenierías, en gran parte de la Enseñanza Técnica y Pedagógica y desde luego, en Física, en Química y en toda la Enseñanza Media Superior (pre-universitaria).

Por otra parte, los conceptos *calor* y *trabajo* se encuentran en la base de la Termodinámica. Ambos son parte integrante de la primera ley, y el primero de ellos también está asociado a la segunda mediante la eficiencia de las máquinas térmicas. Tener una noción clara de lo mismos es esencial para que cualquier estudiante aprenda a razonar correctamente en términos termodinámicos y sea capaz de profundizar por sí mismo en el estudio de la disciplina. Sería de esperar, por tanto, que ambos términos estuvieran muy bien definidos en la literatura docente, sin dar lugar a dudas o ambigüedades. En la práctica sucede todo lo contrario, como se verá más adelante.

Ha sido precisamente la falta de coincidencia entre las definiciones propuestas por diferentes

autores lo que nos condujo a analizar en detalle las definiciones de calor y trabajo que aparecen en varios textos conocidos de Física General y Termodinámica. Casi todos han sido utilizados alguna vez como textos básicos o de consulta en diferentes facultades de la Universidad de La Habana y de otras universidades nacionales. Varios aún se utilizan, en nuevas ediciones o incluso en la misma edición. Como resultado del análisis, fue posible encontrar dos definiciones razonablemente simples y compatibles con los conceptos modernos de calor y trabajo. A pesar de la relativa simplicidad de sus enunciados, las definiciones propuestas resultan ser lo suficientemente generales como para no desmerecer en rigurosidad científica.

Todas las definiciones analizadas a continuación pertenecen a textos que siguen la formulación usual de Poincaré (1908), basada en definiciones de variables mensurables. En el apéndice 1 se exponen brevemente algunos criterios sobre los textos que siguen la formulación matemática alternativa de Carathéodory (1909), identificada por su carácter lógico, abstracto y postulacional¹. La práctica ha demostrado ampliamente que, desde el punto de vista didáctico, ésta última formulación resulta totalmente inadecuada para una primera aproximación al estudio de la Termodinámica.

Calor

Tabla 1.

En la Tabla 1 se reproducen las definiciones de calor y trabajo proporcionadas por los textos mencionados anteriormente. Unas se destacan por su falta de precisión y transparencia (definición de trabajo en /3/ y /4/), otras por su falta de actualización /1/, y otras son francamente desafortunadas y pueden conducir a falsas interpretaciones de la realidad /2/.

En /2/ el calor y el trabajo se definen en función de los “parámetros externos e internos”. Según esa misma referencia, “...los parámetros externos son cantidades que están determinadas por la posición de cuerpos externos que no entran en nuestro sistema. Ejemplo: el volumen del sistema ...la intensidad de un campo de fuerzas ... etc.”. Por el contrario, la presión del sistema, su densidad y magnetización son parámetros internos. Y aunque no aparece claramente expresado en el texto, aparentemente la temperatura T es también un parámetro interno. En la definición de calor y trabajo se menciona la energía “transmitida por un sistema”. No queda claro si la definición también sería válida para la energía “recibida por un sistema”. Sin embargo, no hay necesidad de esclarecer esta aparente omisión para llegar fácilmente a la conclusión de que las definiciones en /2/ resultan bastante confusas y llegan incluso a contradecir los criterios comúnmente establecidos sobre el calor y el trabajo.

Por ejemplo, considere la solidificación de cualquier sustancia a una presión determinada y el correspondiente calor de solidificación, - $Q_{\text{FUSIÓN}}$. Si a este proceso se le aplican las definiciones en /2/, como el cambio de estado tiene lugar a una temperatura T esencialmente constante, no habrá variación de los parámetros internos. El volumen (parámetro externo) variará durante la solidificación, en mayor o menor grado según la sustancia considerada. Habría que considerar, por tanto, que la transmisión de energía al medio ambiente durante la solidificación es en forma de trabajo, por estar asociada a una variación de volumen a T constante, y no de calor, como ocurre en realidad. Pudiera alegarse que la variación del volumen trae aparejada una variación de la densidad (parámetro interno) y, por tanto, además del trabajo habría también flujo de calor. Pero es que las definiciones propuestas en /2/

Ref.	Calor (Q)	Trabajo (W)
2	Se hace un recuento histórico hasta llegar al equivalente mecánico del calor, sin que se proporcione una definición independiente de la de trabajo.	$\int F \cdot dr$
3	Cantidad de energía transmitida por un sistema sin variación de sus parámetros externos.	Cantidad de energía transmitida por un sistema con variación de sus parámetros externos.
4	Flujo de energía a través de la superficie que limita al sistema, como consecuencia de un gradiente de temperatura.	Calor y trabajo se refieren a dos mecanismos diferentes de transferencia de energía entre sistemas y su medio exterior.
5	Energía transferida de uno a otro sistema en escala atómica, sin que se realice trabajo, mediante interacción térmica	a) Hecho sobre un sistema: incremento (positivo o negativo) de la energía media de un sistema adiabáticamente aislado. b) Si hay decremento (positivo o negativo) el trabajo es hecho por el sistema.
6	Energía media transferida de un sistema a otro como consecuencia de una interacción puramente térmica	Intercambio de energía por “interacción mecánica pura” entre dos sistemas térmicamente aislados

no especifican lo que sucede cuando varían los parámetros internos. Solo se refieren a la variación o constancia de los parámetros externos (ver Tabla 1). No es difícil llegar a la conclusión de que, basándose solamente en estas definiciones, resulta muy difícil -por no decir imposible- diferenciar claramente el calor del trabajo en determinados procesos.

El análisis anterior también es válido para la fusión, sublimación, licuefacción y, en general, para cualquier cambio de estado de agregación. Es válido igualmente para cualquier cambio de fase de primer orden que tenga lugar a p y T constantes con $\Delta H = Q_p \neq 0$ (cambios de fase cristalina, recristalizaciones). En la expresión anterior $H = E + pV$ es la entalpía y Q_p el calor a presión constante.

Los ejemplos anteriores, donde hay presente un flujo de calor a T constante, también tienden a invalidar la definición propuesta en /3/, donde se exige la existencia de un gradiente de temperatura para que exista el mencionado flujo.

Las definiciones de calor en las referencias /4/ y /5/ son aceptables, siempre y cuando se defina

previamente el significado de “interacción térmica”. Como la interacción térmica ocurre a nivel microscópico, donde átomos y moléculas se mueven desordenadamente, es posible adoptar para el calor una definición independiente del concepto de interacción térmica de la siguiente forma:

“Calor: transferencia de energía en forma microscópica y desordenada”. (1)

La definición anterior no sólo permite caracterizar el flujo de calor entre dos sistemas; también permite considerar el calor intercambiado de una región a otra de un mismo sistema.

Si un sistema incrementa su energía interna E al recibir una cierta cantidad de calor Q proveniente del medio ambiente, el proceso suele simbolizarse por la expresión analítica

$$Q = \Delta E. \tag{2}$$

El convenio de signos es aquí muy sencillo. Un valor positivo de Q indica un incremento ΔE de la energía interna, donde el sistema A gana o recibe energía en forma de calor, proveniente del medio ambiente A' (ver Figura 1). Un valor negativo indicará, por tanto, una disminución de la energía interna, con el calor pasando de A hacia A' .

El calor así definido es una magnitud perfectamente *mensurable* en forma independiente del trabajo. Históricamente, como unidad de medida se tomó la *caloría*, definida como la cantidad de calor necesaria para elevar la temperatura de 1 g de agua en 1°C^a.

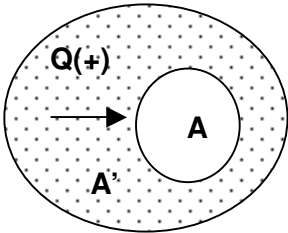


Figura 1

Trabajo

La relativa simplicidad con que se logra definir el significado del calor desaparece cuando la transferencia de energía se refiere al trabajo. La razón es que existen diferentes tipos de trabajo (mecánico, eléctrico, magnético) y no resulta evidente encontrar lo que es común a todos ellos y que, a la vez, los diferencia del calor. Así, por ejemplo, las definiciones de trabajo en /4/ y /5/ dejan bastante que desear. En /5/ no se considera la posibilidad de trabajo eléctrico o magnético, y la definición en /4/ resulta bastante confusa.

Con el ánimo de llegar a una definición de trabajo que fuese lo más general posible y que, además, fuera compatible con la definición en (1), se analizaron las relaciones para el trabajo y la energía que usualmente se derivan en los cursos de Mecánica. Estas relaciones aparecen resumidas en la Tabla 2, donde se utilizan los símbolos E_c , E_p y E_M para designar la energía cinética, potencial y mecánica. El trabajo de la fuerza resultante y los trabajos conservativo y no conservativo se designan por W_R , W_C y W_{NC} , respectivamente. Las expresiones de la Tabla, en el caso (I), se refieren a algún cuerpo o sistema A sobre el cual actúa una fuerza aplicada por algún otro sistema A' . En Mecánica el sistema A' usualmente no se toma en cuenta; ni siquiera se menciona. Así, un cuerpo cualquiera al que se le aplica una fuerza de fricción variará su energía mecánica, sin que usualmente nos interese lo que le sucede al otro cuerpo cuya superficie ha interaccionado con el cuerpo objeto de estudio. La representación correspondiente al flujo de energía en forma de trabajo sería similar a la Figura 1, sustituyendo el calor por el trabajo W .

Tabla 2.

(I) Mecánica de la partícula	
Teorema del trabajo y la energía	$W_R = \Delta E_c$
Trabajo de una fuerza conservativa	$W_C = - \Delta E_p$
Trabajo no conservativo	$W_{NC} = \Delta E_M$
(II) Mecánica de los Sistemas de Partículas	
Teorema del trabajo y la energía	$W_{R(externo)} + W_{R(interno)} = \Delta E_c^b$

En el caso (II) las variaciones de energía en el sistema pueden ser originadas tanto por fuerzas externas como por fuerzas internas.³ Sin embargo, siempre será posible *redefinir* nuestro sistema para considerar el trabajo de interés como externo al mismo, de forma que no haya diferencias esenciales con el caso (I).

Un error bastante extendido, y que usualmente induce a confusión en los estudiantes, es referirse al trabajo realizado *por el sistema*. El trabajo se define para las fuerzas, no para los sistemas. En un determinado sistema habrá tantos trabajos como fuerzas haya desplazándose.

^aActualmente la caloría se encuentra redefinida sobre la base de las unidades del Sistema Internacional; 1 cal = 4.1868 J.
^bLa energía cinética del sistema de partículas se define como la suma de las energías cinéticas de todas las partículas en el sistema.

En Mecánica las fuerzas involucradas en un proceso y el correspondiente trabajo realizado siempre son *macroscópicos*, y en principio se podría considerar al trabajo como una *transferencia de energía en forma macroscópica*, para diferenciarlo del calor microscópico en (1). Sin embargo, en los casos eléctrico y magnético el trabajo también puede ser *microscópico* (i.e., trabajo de polarización y de magnetización). No obstante, en estos casos siempre aparece (o desaparece) un cierto *orden* asociado a la transmisión de energía, a causa de la redistribución ordenada de las cargas o a la alineación de los dipolos eléctricos o magnéticos. Esta particularidad de los trabajos eléctrico y magnético permite redefinir el trabajo desde el punto de vista termodinámico, de forma razonablemente simple y general, sin incompatibilidades con los conocimientos previos de Mecánica, y bien diferenciado de la definición de calor en (1), como

"Trabajo: transferencia de energía en forma ordenada y/o macroscópica" (3)

Las definiciones anteriores (1) y (3) han sido adoptadas por el autor durante la impartición de varios cursos introductorios de Termodinámica, con resultados razonablemente satisfactorios. Una versión escrita de estos cursos puede ser consultada en la WEB^a.

Para establecer un convenio de signos similar al del calor en la Figura 1, bastaría considerar positivo el trabajo W' cuando hay un incremento de energía interna ΔE . Sin embargo, este proceder introduciría la dificultad de que, al aplicar la primera ley, obtendríamos expresiones donde se encontrarían mezclados parámetros de estado del sistema y del medio ambiente. Esto haría prácticamente imposible la definición posterior de cualquier función termodinámica que dependa exclusivamente de los parámetros del sistema.

Así, por ejemplo, si consideramos el trabajo reversible de expansión o compresión de un gas ideal (Figura 2), es correcto considerar el trabajo W' como positivo cuando el gas se comprime y se incrementa su energía interna. La correspondiente expresión analítica sería

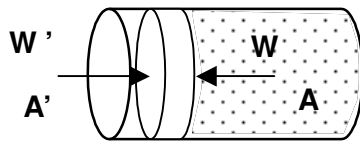


Figura 2

$$W' = \Delta E \quad (4)$$

Si se utiliza la notación diferencial con el fin de simplificar las expresiones^b, es posible expresar el trabajo como $\delta W' = p'dV$, donde p' es la presión *externa* al sistema. Al sumar las relaciones (2) y (4) para llegar a la primera Ley, se obtiene:

$$Q + W' = \Delta E$$

$$\delta Q = dE - p'dV \quad (5)$$

Como p' es la *presión externa*, y no la ejercida por el gas, no sería entonces posible definir posteriormente las restantes funciones de estado o potenciales termodinámicos del sistema^c pues la energía interna en (5) es función de estado^d, pero p' no es un parámetro de estado del sistema, sino del medio ambiente.

Afortunadamente, es posible sustituir en (4) la presión externa p' por la presión p del sistema con sólo cambiar el signo del trabajo. Esta última afirmación se demuestra fácilmente al recordar que *el trabajo reversible es el único que puede ser representado analíticamente en Termodinámica*. Durante el trabajo reversible el proceso ocurre esencialmente en equilibrio (o tan cercano a él como se quiera). Examinando el caso simplificado de una fuerza constante en la Figura 2, tendremos que $\mathbf{F}' = -\mathbf{F}$, y como el área de la pared móvil es común al sistema y al medio ambiente, $p' = p$. Entonces, si d es el desplazamiento de la pared durante el proceso y V el correspondiente volumen barrido;

$$W' = \mathbf{F}' \cdot \mathbf{d} = F'd\cos\theta = p'Ad\cos\theta = p'V\cos\theta,$$

con una expresión análoga para W , donde aparece p en vez de p' y el resto de los parámetros no varía. Como los valores de F' , A , d , V y p' son modulares, tanto el signo de W' como el de W dependerán exclusivamente del valor de $\cos\theta$. En la compresión W' es positivo ($\theta = 0$) y W es negativo ($\theta = 180^\circ$). Durante la expansión se intercambian los signos. En general, para ambos procesos se cumple que $W' = -W$. La generalización a fuerzas variables en tres dimensiones es inmediata⁷.

Sustituyendo $W' = -W$ en (4) y p' por p , se obtiene finalmente:

$$Q = \Delta E + W$$

$$\delta Q = dE + pdV \quad (6)$$

^awww.geocities.com/fisgeo2000, ó www.fisica.uh.cu/bibvirtual/fis1y2.

^bEste razonamiento puede extenderse a estudiantes sin conocimientos previos de cálculo considerando una variación ΔV muy pequeña, de forma que la presión p' se puede considerar aproximadamente constante durante la variación de volumen.

^cPor ej., la energía libre de Gibbs, $G = E - TS + pV$.

^d $E = E(p, V, T, \dots)$

Ahora tanto la presión como el volumen son parámetros de estado de nuestro sistema. Si, por ejemplo, la presión se mantiene constante,

$$Q_p = \Delta E + p\Delta V = \Delta H,$$

y el calor a presión constante pasa a ser función del estado del sistema. La importancia de las funciones de estado o potenciales termodinámicos es que, como su valor depende sólo de los estados inicial y final del proceso, constituyen la única vía para calcular calores y trabajos en procesos reales *irreversibles*, en los cuales las características del sistema en los estados intermedios se desconocen totalmente.

Note que el convenio de signos inicial no ha variado: en ausencia de calor, la energía interna aumenta cuando el gas se comprime (y δW , referido ahora a las variables del sistema, es negativo). Cuando el gas se expande el trabajo es positivo, y si además $Q = 0$, la energía interna disminuye. Este último es el convenio empleado usualmente en los textos de Termodinámica, pero sin que se analicen en detalle los criterios que le dieron origen. Cuando hay presente trabajo eléctrico o magnético el convenio de signos no se altera. Analíticamente, considerando un proceso donde no hay intercambio de calor y $\delta Q = 0$, tendremos que $\Delta E = -W$. Durante la expansión $dV > 0$; $\delta W = pdV > 0$, y el sistema cede energía a los alrededores. Durante la compresión $dV < 0$, $\delta W < 0$ y el sistema gana energía.

A la luz del análisis anterior queda perfectamente esclarecido que expresiones tales como “el sistema hace trabajo” y “se hace trabajo sobre el sistema”, utilizadas frecuentemente para caracterizar las variaciones de energía de un gas, o el trabajo neto realizado por una máquina térmica durante un ciclo, no son satisfactorias. En primer lugar, como se mencionó anteriormente, el trabajo es realizado por las fuerzas y no por el sistema.

En segundo lugar, de la Figura 2 queda perfectamente claro que tanto en la expansión como en la compresión reversible siempre estarán presentes dos fuerzas trabajando en equilibrio y los dos correspondientes trabajos: uno positivo y el otro negativo. Considerar que en un caso “el sistema hace trabajo” y en el otro “se hace trabajo sobre el sistema” equivale a cambiar arbitrariamente el origen de referencia de las fuerzas que actúan: en un caso sólo se considera la fuerza ejercida por el sistema y en el otro la fuerza ejercida por el medio ambiente, ignorando en cada caso la presencia de la otra fuerza.

Finalmente, en la expresión analítica del trabajo de expansión, la presión que siempre aparece es la del sistema y no la del medio ambiente. Si estamos analizando un sistema determinado, y todos los parámetros de estado se refieren a ese sistema, deberíamos referirnos siempre al trabajo realizado por nuestro sistema y no al realizado por el medio ambiente, cuyas características usualmente se desconocen. Tanto durante la expansión como en la compresión el gas trabaja.^a En el primer caso el trabajo es positivo. En el segundo, negativo. Y nada más.

CONCLUSIONES

- A partir del análisis de diversas definiciones poco satisfactorias de calor y trabajo que aparecen en textos conocidos de Mecánica y Termodinámica se proponen las siguientes expresiones generales para el calor y el trabajo. *Calor: transferencia de energía en forma microscópica y desordenada. Trabajo: transferencia de energía en forma ordenada y/o macroscópica.*
- Se hace un llamado de atención sobre las definiciones de calor y trabajo que aparecen en algunos libros de texto, para las cuales es posible encontrar fácilmente casos particulares que dan lugar a contradicciones, dificultando la correcta comprensión de estos términos por parte del estudiante.
- Expresiones como “el sistema hace trabajo” y “se hace trabajo sobre el sistema” equivalen a utilizar indistintamente dos referencias; una asociada al sistema de interés y la otra al medio ambiente. Se debería utilizar siempre una misma referencia, ligada al sistema, y no alternarla con otra asociada al medio ambiente, cuyos parámetros de estado usualmente se desconocen y no son objeto de interés.

APÉNDICE 1

Los textos que siguen la formulación de Carathéodory también suelen presentar ambigüedades en las definiciones de calor y trabajo. Por ej., H.B. Callen⁸ proporciona dos definiciones... ¿independientes?... para el calor, una cuantitativa y otra cualitativa. Sin pretender profundizar en el aspecto gnoseológico o epistemológico de la validez de tal procedimiento^b, pues el tema se escapa totalmente de los objetivos de éste artículo, resulta indispensable señalar lo siguiente.

La definición “cualitativa” proporcionada en [8] introduce, directa o indirectamente, conceptos tales como “modos atómicos ocultos” o “promedio

^aLas fuerzas por él ejercidas realizan trabajo

^b... por ejemplo; ¿por qué dos definiciones donde debería bastar con una?

estadístico". Estos conceptos requieren de conocimientos previos en otras disciplinas y, además, el "promedio estadístico" no es en realidad un término cualitativo, sino cuantitativo. Por otra parte, la definición "cuantitativa" propuesta consiste simple-

mente en definir el calor a partir de la primera ley, pero sin haber discutido ésta previamente. Resulta bastante difícil encontrar una justificación didáctica para este procedimiento en un curso básico de Termodinámica.

REFERENCIAS

¹©1995-2004 by Isidoro Martínez, Historia de la Termodinámica, (imartínez@etsin.upm.es)
<http://imartinez.etsin.upm.es/bk3/History.htm>

²FRISH, S. and A. TIMOREVA (1967): **Curso de Física General 1**, Ed. Mir.

³BAZAROV, I.O. (1969): **Thermodynamics**, Ed. R.

⁴SEARS, FRANCIS W. (1958): **Mecánica, Mov. Ondulatorio y Calor**, Ed. R. La Habana.

⁵REIF, F. (1965): **Berkeley Physics Course**, 5, Ed. R.

⁶_____ (1968): **Fundamentos de Física Estadística y Térmica**, Ed. del Castillo.

⁷SEARS, FRANCIS W. (1959): **Termodinámica**, Ed. Reverté.

⁸CALLEN, HERBERT B. (1981): **Termodinámica**, Editorial AC, España.

ACLARANDO LA LÓGICA BORROSA (FUZZY LOGIC)

Manuel Hernández Calviño*, Facultad de Física, Universidad de La Habana

RESUMEN

La Lógica Borrosa es un formalismo matemático que pretende emular la habilidad que tienen algunas personas para tomar decisiones correctas a partir de datos vagos o imprecisos y que están expresados lingüísticamente. En la última década la Lógica Borrosa se ha utilizado fundamentalmente para realizar sistemas de control de procesos o de ayuda a toma de decisiones, porque permite aprovechar la experiencia de un experto humano e implementar el sistema rápida y eficientemente. El presente artículo divulgativo pretende llevar al lector sus ideas básicas e ilustrar el método de desarrollar un controlador basado en ella. Finalmente se listan algunas de sus aplicaciones más importantes en Japón.

ABSTRACT

Fuzzy Logic is a mathematical formalism aimed to emulate the ability of certain people to take right decisions departing from data which is expressed linguistically in terms vague or imprecise. In the past decade Fuzzy Logic has been mainly used to implement process control systems and decisions making aid, because it allows to profit from human expertise reducing development time. This introductory paper presents the reader Fuzzy Logic basic ideas and illustrates the method to build a controller based on it. Finally, a list of some important applications in Japan is presented.

INTRODUCCIÓN

La Lógica Borrosa (LB) surgió como consecuencia natural de la observación de que ciertas personas tienen suficiente habilidad para tomar decisiones correctas a partir de un conjunto de datos que están expresados lingüísticamente de forma vaga o imprecisa (borrosos), casi siempre utilizando adjetivos o adverbios como **mucho**, **alto**, **normal**, **muy**, etc. Tales personas pueden controlar eficientemente un proceso tecnológico¹ (un buen ejemplo cubano es el tradicional puntista en un central azucarero que controla el proceso de cristalización del azúcar), diagnosticar una enfermedad a partir de síndromes y síntomas (el médico clínico) o tomar una decisión acertada en el mercado de valores.

La LB es un cuerpo teórico que pretende emular tales capacidades mediante su formalización y forma parte junto a las redes neuronales, los algoritmos genéticos, los sistemas expertos, etc. de los muchos esfuerzos que se hacen para crear la llamada "inteligencia artificial". Mediante la utilización de la LB se pueden desarrollar sistemas de control de procesos o de ayuda a toma de decisiones, con las siguientes ventajas:

- Eficiencia y rapidez en la implementación, porque brinda una metodología para aprovechar la experiencia de un experto.
- Tiene éxito en aquellos casos en que los datos de entrada por su propia naturaleza son escasos, imprecisos o ruidosos.

- Hace innecesario modelar detalladamente el sistema que se pretende controlar o predecir, porque utiliza un método heurístico.
- Permite controlar sistemas con una no-linealidad muy marcada, donde las estrategias tradicionales, casi siempre controladores del tipo Proporcional-Integral-Derivativo (PID) ofrecen pobres resultados.

Trescientos años a.n.e., Aristóteles estableció su llamada Ley de Bivalencia que afirma que cualquier sentencia es verdadera o falsa (1, 0), pero no ambas cosas a la vez. La lógica aristotélica nos ha sido útil por más de 2000 años y está en los cimientos de la Matemática y en el principio de funcionamiento de nuestras computadoras. Pensadores posteriores sugirieron que el mundo está lleno de contradicciones, de cosas que son y no-son a un tiempo y que por tanto una tercera región debía ser considerada. En el siglo pasado, un matemático llamado Lukasiewicz² propuso inicialmente una lógica trivalente. Posteriormente experimentó con lógicas de cuatro y cinco valores y finalmente llegó a la conclusión que una lógica con infinitos valores era tan plausible como una lógica con un conjunto finito de ellos. La LB es precisamente eso, una lógica con infinitos valores que puede verse como una generalización de la lógica bivalente tradicional.

Muchas veces tenemos la falsa idea que describir algo con un término lingüístico y aparentemente vago, tiene menos información que si brindamos un valor exacto. Para comprenderlo mejor piénsese en el siguiente ejemplo: Es más rico en información decir **María es alta** que decir **María mide 1.90 m.**, porque

en el primer caso clasificar a María de **alta** realmente quiere decir: **María tiene una estatura que supera en dos desviaciones estándar a la media de los individuos de su sexo, raza y cultura**. La razón es simplemente que el concepto de persona **alta** es contextual y lleva implícito mucha información adicional. Esto es precisamente lo que se propone la LB, operar con conceptos aparentemente vagos o subjetivos pero que en realidad contienen mucha información, de los que se pueden obtener conclusiones útiles. En última instancia, la filosofía de la LB es que el razonamiento exacto es un caso particular y límite del razonamiento aproximado.

Surgida en los años 60, la LB pasó por una necesaria etapa de desarrollo teórico durante las siguientes dos décadas y por supuesto ha tenido y tiene muchos detractores³, quienes han tratado de demostrar con cierto éxito que es innecesaria. Sus defensores asumen una posición pragmática y argumentan que el éxito de su utilización justifica por sí solo el seguir impulsando su desarrollo. Lo cierto es que la década de los 90 ha sido testigo de una explosión de su utilización en los más diversos campos de aplicación, que van desde el control de una simple lavadora doméstica de ropa hasta el de una central electro-nuclear. En la actualidad son numerosas las publicaciones en el mundo y los grupos que se dedican a estudiar los llamados "sistemas borrosos".

La teoría de los Conjuntos Borrosos

Se considera que el padre de la LB es el azerbaijano Lofti Zadeh quien trabajando en la Universidad de Berkeley a principios de la década de los 60, publicó un par de trabajos^{4,5} ahora considerados medulares sobre lo que él denominó los Conjuntos Borrosos y cuya característica esencial es que, a diferencia de los conjuntos Booleanos clásicos, la propiedad de pertenencia de un elemento a un determinado conjunto, se describe por una función que puede variar continuamente entre 0 y 1, llamada Función de Pertenencia (FP). En el álgebra de Boole clásica, la propiedad de un ente de pertenecer a un conjunto específico sólo puede tomar dos valores (falso, verdadero) a los que se les asigna por convenio los valores extremos 0 y 1, pero no hay valores intermedios.

La FP puede interpretarse como el grado en que el elemento particular que estamos considerando cumple con las especificaciones que definen a los elementos del conjunto en cuestión y no debe interpretarse como la probabilidad de pertenencia. Si la probabilidad de que el elemento X pertenezca al conjunto A es de 0.8 y afirmamos que X pertenece al conjunto A, tenemos 80 % de probabilidad de acertar, pero el elemento intrínsecamente pertenece o no-pertenece a A. Cuando decimos que la FP de X es 0.8 queremos decir que cumple en nuestro

criterio con el 80 % de las características que definen a los elementos del conjunto A. En resumen, la probabilidad indica incertidumbre estadística mientras que la FP indica vaguedad y subjetividad. A los creadores de la LB les costó trabajo apartarse del concepto de probabilidad y convencerse de que la FP describe una cualidad distinta.

Para comprender mejor lo que estamos explicando veamos un ejemplo sencillo. Supongamos que la temperatura corporal de un paciente es clasificada en tres términos lingüísticos: **normal**, **moderada** y **elevada**. Supongamos también que definimos a los pacientes con temperatura **elevada** como aquellos con $T \geq 39^\circ\text{C}$. Está claro que en el contexto clínico, un paciente con $T = 38.99^\circ\text{C}$ representa la misma situación. Imaginemos que un sistema de monitoreo vigila a los pacientes y alerta cuando la temperatura es **elevada**. Se producirá una transición brusca al pasar de 38.9°C a 39°C , aunque el cuadro clínico quizás no haya cambiado sustancialmente. Es más, aumentar la precisión de la medición no resolvería el problema, porque un paciente con $T = 38.999^\circ\text{C}$ sigue clasificando como con temperatura **moderada** y el ruido de la medición puede activar la alarma frecuentemente.

Lo que propone Zadeh es definir para cada conjunto (pacientes con temperatura normal, moderada y elevada) una función de pertenencia tal como muestra la Figura 1. Así un paciente con $T = 36.5^\circ\text{C}$ definitivamente pertenece al conjunto de las personas con temperatura corporal **normal** ($FP = 1$), mientras que un paciente con $T = 38.2^\circ\text{C}$ pertenece al mismo tiempo a los conjuntos **moderada** ($FP = 0.6$) y **elevada** ($FP = 0.2$) y definitivamente no pertenece al conjunto de los pacientes con temperatura corporal **normal** ($FP = 0$).

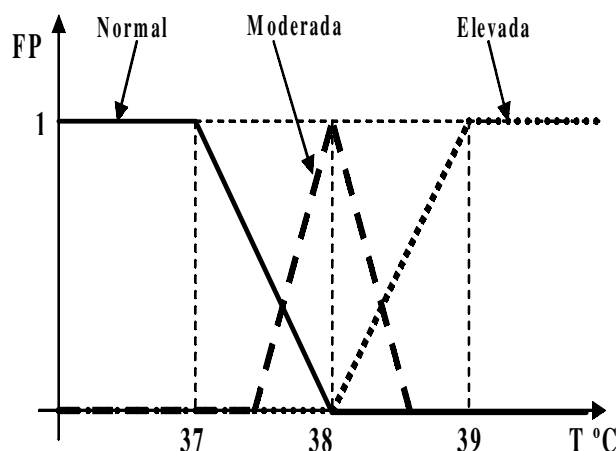


Figura 1. Funciones de pertenencia de la variable T.

Para poder operar con los Conjuntos Borrosos es necesario definir las operaciones elementales entre ellos, léase la UNIÓN, la INTERSECCIÓN y el COMPLEMENTO. Esto implica definir el modo de

calcular las FP a estos tres conjuntos. Sean $FP(X)$ y $FP(Y)$ las funciones de pertenencia correspondientes a los conjuntos borrosos X y Y . Zadeh propone:

Intersección o AND	$FP(X \text{ AND } Y) = \text{mínimo de } (FP(X), FP(Y))$
Unión u OR	$FP(X \text{ OR } Y) = \text{máximo de } (FP(X), FP(Y))$
Complemento o NOT	$FP(\text{Complemento } X) = 1 - FP(X)$

Estas definiciones tienen dos rasgos interesantes. En primer lugar el lector fácilmente se dará cuenta que las operaciones lógicas AND, OR y NOT clásicas con conjuntos booleanos cumplen con estas definiciones. Si $A = 0$ y $B = 1$, $A * B = 0$ porque toma el mínimo de ambos valores y $A + B = 1$ porque toma el máximo. Las operaciones con conjuntos borrosos son asociativas, conmutativas y distributivas y también cumplen con las leyes de DeMorgan. De esta forma la lógica de los conjuntos borrosos puede considerarse como una generalización de la lógica booleana⁶.

En segundo lugar, puede observarse que si la FP se interpretara como la probabilidad de pertenencia a un conjunto, entonces la FP del conjunto Intersección sería el producto de las FP individuales. Puesto que estamos hablando de funciones comprendidas en el intervalo cerrado 0-1, es evidente que $FP(X) * FP(Y) < \text{mínimo de } (FP(X), FP(Y))$, es decir, si hubiéramos definido la FP del conjunto intersección como el producto de las FP individuales, se obtendría un valor menor y por tanto representativo de una categoría lingüística inferior. Supongamos que la estatura de una persona se clasifica con los términos lingüísticos **alto**, **mediano** y **bajo** y su peso con **gordo**, **normal** y **flaco**. Supongamos también que alguien en particular se clasifica como perteneciente al conjunto **alto** con $FP = 0.6$ y al **gordo** con $FP = 0.5$. La FP correspondiente al conjunto intersección **alto AND gordo** es 0.5, mientras que si interpretáramos las FP individuales como probabilidades, la probabilidad de ser a un tiempo alto y gordo es de sólo 0.3.

Como el lector habrá podido observar las reglas de la LB no tienen nada de "borrosas" y están perfectamente definidas por lo que quizás sería más adecuado llamarla "lógica de las variables borrosas".

Los controladores basados en Lógica Borrosa

La ventaja fundamental que tiene la LB es que permite desarrollar sistemas de toma de decisiones o controladores en tiempo real, aprovechando la experiencia humana. Esto agiliza la etapa de implementación, porque evita tener que desarrollar modelos matemáticos exactos del proceso que se desea controlar, lo que es imprescindible con las estrategias de control convencionales. Hay muchos problemas de control donde tales modelos no existen o son difíciles de obtener. Otra ventaja de la LB es que tiene éxito en aquellas situaciones donde son pocos los datos de entrada, tienen poca resolución o la relación señal/ruido es pequeña.

Para comprender mejor cómo se desarrolla un controlador basado en LB, hagámoslo con un ejemplo hipotético. Supongamos que tenemos un proceso tecnológico donde es necesario para la buena marcha del mismo, controlar la temperatura **T** (variable de salida). Supongamos también que las variables físicas que sirven de indicadores en cada momento de que la temperatura debe ser modificada son la presión **P** y el flujo **F**. (Variables de entrada)

Con la ayuda de un experto en controlar este proceso, comenzaríamos por definir los términos lingüísticos de las dos variables de entrada y la de salida. Supongamos que para la temperatura el experto utiliza los términos **frío**, **tibio** y **caliente**; para la presión utiliza **alta**, **media** y **baja** y para el flujo **mucho**, **normal** y **poco**. El siguiente paso es definir las FP de las tres variables. En esta etapa del proceso el experto nos dirá, por ejemplo, a partir de cual temperatura T_1 él considera al proceso ciertamente **frío** ($FP = 1$) y la temperatura T_2 donde definitivamente deja de serlo ($FP = 0$). En el intervalo T_1, T_2 la FP toma valores intermedios. Así se procede con las dos variables de entrada y la variable de salida. Para calcular los valores de FP en la zona de transición, se prefiere utilizar funciones lineales, porque usar funciones más complejas no mejoran sustancialmente la prestación del controlador y sí incrementan apreciablemente el tiempo de cálculo. No hay que olvidar la naturaleza borrosa de las regiones que estamos definiendo.

La siguiente etapa consiste en definir las llamadas reglas de inferencia, las que representan en cierta medida "la inteligencia del controlador". Estas reglas emulan la forma en que el experto procedería para controlar la temperatura del proceso. Un ejemplo de una de ellas podría ser:

IF P es alta AND F es mucho THEN T es frío.

Puesto que en este ejemplo las variables **P** y **F** cada una tienen tres términos lingüísticos, existen nueve reglas como la anterior donde se hacen todas las combinaciones posibles. En realidad el experto sólo define, de acuerdo con su experiencia, cual es la consecuencia de cada una de las reglas, es decir, lo que va después del THEN. La Tabla 1 muestra un resumen de todos los casos posibles, para nuestro ejemplo concreto.

Tabla 1. Resultados de las nueve reglas de inferencia.

		Presión P		
		Baja	Media	Alta
Flujo F	Mucho	T es tibio	T es frío	T es frío
	Normal	T es caliente	T es tibio	T es frío
	Poco	T es caliente	T es caliente	T es tibio

Toda vez que las reglas de inferencia han sido definidas, el controlador está listo para ser probado y puesto a punto. El controlador funciona en un bucle cerrado donde ejecuta repetidamente la siguiente secuencia de acciones:

- Se leen los valores físicos de las variables **P** y **F**.
- Se calculan las FP de los tres términos lingüísticos de cada una de ellas. A este paso en la literatura especializada se le ha llamado **fuzzyficación**.
- Se evalúan los miembros izquierdos de las nueve reglas de inferencia y se le asigna el valor mínimo (operación AND) de ambos operandos al miembro derecho.
- Se procede a la llamada **defuzzyficación**, es decir a calcular el valor físico que debe tener la variable **T**.
- Se realizan las acciones necesarias para establecer la temperatura calculada.

De las operaciones anteriormente listadas sólo la penúltima necesita de una explicación adicional. En un determinado momento habrá valores de **P** y **F** cuyas FP son iguales a cero para ciertos términos lingüísticos. Esas reglas dan una consecuencia también nula (al tomar el valor mínimo) y se llaman reglas “inactivas” porque no contribuyen al valor de **T**. Entre las reglas “activas” puede haber varias que tengan una consecuencia descrita por el mismo término lingüístico, por ejemplo, **T es caliente**. Supongamos que dos reglas distintas brindan el resultado **T es caliente** con FP = 0.8 y **T es caliente** con FP = 0.6. Entre ellas se hace la operación OR y se toma el máximo de ambas, es decir al final nos quedamos con el resultado **T es caliente** con FP = 0.8.

Finalmente queda por resolver el problema de cómo combinar las reglas cuyas consecuencias están descritas por términos lingüísticos diferentes entre si, por ejemplo, **T es caliente** con FP = 0.8 y **T es tibio** con FP = 0.4. Se trata ahora de encontrar el mejor compromiso para el valor de **T** y para ello han sido propuestos varios métodos. Aquí explicaremos el llamado método del Centro de Gravedad

(CoG), que es uno de los más utilizados y que se ilustra en la Figura 2. Las FP **T es tibio** y **T es caliente** se truncan a sus valores máximos, a continuación se superponen las áreas resultantes y se calcula la coordenada **T** del Centro de Gravedad del área total. De esta forma obtenemos el valor de la temperatura que debe establecer el controlador en esta fase del proceso.

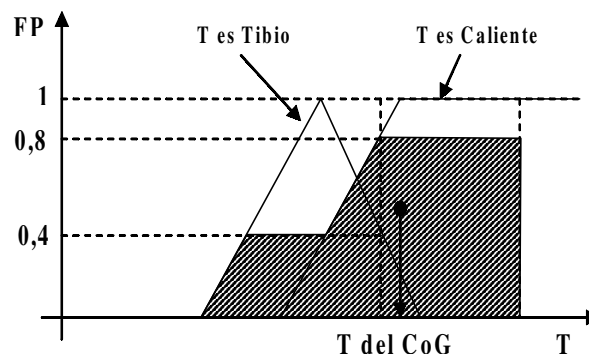


Figura 2. Método del Centro de Gravedad para calcular el mejor valor para **T**.

Para poner a punto y afinar la prestación de un controlador basado en LB se utilizan generalmente dos métodos. Uno de ellos es modificar las consecuencias de una o varias reglas de inferencia, sobre todo de aquellas que permanecen más tiempo activas y tienen mayor influencia en la variable de salida del controlador. El otro método es multiplicar el resultado de una regla de inferencia por un factor de ponderación comprendido ente 0 y 1. Por omisión se supone que todas las reglas tienen al inicio del proceso de puesta a punto, un factor de ponderación igual a 1. De esta forma podemos reforzar o debilitar la influencia de una regla en el resultado final.

Cuando el sistema a controlar tiene muchas variables de entrada y a la vez cada variable tiene varios términos lingüísticos, la matriz de reglas de inferencia puede ser muy complicada. Una ventaja de la LB es que puede comenzarse con una matriz incompleta donde sólo estén presentes aquellas reglas que resultan activas en el intervalo más común de las variables de entrada. Los expertos recomiendan utilizar para cada variable un número de términos lingüísticos impar y comprendido entre tres y siete. Esto se debe a que las personas tienden naturalmente a clasificar una variable en dos valores extremos y un valor intermedio y a que la memoria de corto plazo le es difícil recordar más de siete situaciones.

En una entrevista realizada en cierta ocasión a Zadeh, se le preguntó dónde inicialmente él esperó que la LB encontrara una mayor aplicación y contestó que en campos como la economía, la sociología y las artes y no precisamente en la ingeniería que es lo que ha sucedido en realidad.

Japón es el país que más uso ha hecho de la LB para el desarrollo de sistemas de control y toma de decisiones en los más diversos campos. Algunos ejemplos muy ilustrativos de aplicaciones en este país⁷ son:

- Mejoras de la seguridad en reactores nucleares. (*Hitachi Bernard Nuclear Fuel Division*)
- Control de trenes subterráneos para mejorar la precisión al detenerse y el ahorro de energía (*Hitachi*)
- Simulación de procesos legales (*Nagoya Univ., Meiki Gakum Univ.*)
- Control de un solo botón de una lavadora doméstica de ropa (*Matsushita, Hitachi*)
- Predicción temprana de terremotos (*Inst. of Seismology Bureau of Metrology, Japan*)
- Sustituto de un experto en actividades de la Bolsa de Valores (*Yamaichi, Hitachi*)
- Diagnóstico del cáncer (*Kawasaki Medical School*)

REFERENCIAS

¹UMBERS, I.G. and P.J. KING (1980): "An analysis of human decision-making in cement kiln control and the implications for automation", **International Journal of Man-Machine Studies**, 12, 11 - 23.

²LEJEWSKI, C. (1967): "Jan Lukasiewicz", **Encyclopedia of Philosophy**, 5, MacMillan, NY 104– 107.

³HAACK, S. (1979): "Do we need fuzzy logic", **International Journal of Man-Machine Studies**, 11, 437 - 445.

⁴ZADEH, L.A. (1965): "Fuzzy sets", **Information and Control**, 8, 338-353.

⁵_____ (1968): "Fuzzy algorithms", **Information and Control**, 12, 94-102.

⁶FOX, J. (1981): "Towards a reconciliation of fuzzy logic and standard logic", **International Journal of Man-Machine Studies**, 15, 213-220.

⁷SCHWARTZ, D.G. and G.J. KLIR (1992): "Fuzzy Logic Flowers in Japan", **IEEE Spectrum**, July.