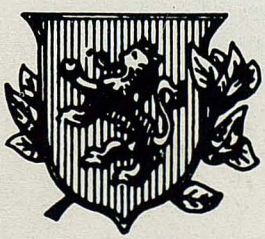
REVISTA
DE LA

ACADEMIA DE CIENCIAS

Exactas
Físicas
Químicas y
Naturales

DE
ZARAGOZA



Serie 2.ª
Volumen 47

1992

INDICE DE MATERIAS

Rev. Academia de Ciencias. Zaragoza 47 (1992)

CONTINUA WITHOUT SETS

Prof. F. Gonzalez Asenjo
Pittsburgh University
Department of Mathematics
Pittsburgh, PA 15260, USA

I.- MOTIVATION AND PURPOSE.

1. FROM THE PHENOMENOLOGY OF THE CONTINUUM.

Initially, we perceive an indefinite extension imprecisely, a spread $\underline{C}$; this perception can be visual, aural, or tactile. Next we perceive in $\underline{C}$ an uncertain number of aspects $\underline{C}^r, \underline{C}^s, \ldots$; these aspects provide $\underline{C}$ with the beginning of a structure. Then in each of these aspects we perceive regions $\underline{x}^r, \underline{y}^r, \underline{x}^s, \ldots$, which give the aspects their internal composition. Yet it would be an error to assume that because we label these regions with distinct letters, each is a fully separated entity. To think thus is to succumb to the insidious atomism induced by written language, a prejudice we must avoid at all costs if we are ever to understand the nature of a concrete continuum as we really perceive it, that is, free of the characteristics injected by our inherited preconceptions, preconceptions as prevalent as they are inadequate. We need a binary predicate of distinguishability to represent the fact that $\underline{x}^r$ is distinguishable from $\underline{y}^s$ but not necessarily vice versa: each distinction implies a direction. Subsequently, we perceive that some regions and aspects are united to one another; some not. Then we see that some regions and aspects are part of other regions and aspects; some not. And finally, we perceive that some regions and aspects lie between other regions and aspects, which represents a fundamental ternary relation of betweenness that is -- together with the two binary relations of unity and being a part of --

essential to the constitution of any concrete continuum and especially to its topological structure.

2. THE INADEQUACY OF CONTINUA BUILT ON SET THEORY.

The current mathematical theory of the continuum is Dedekind's, based on sets. In this theory, the continuum is built from the bottom up, with deceiving looseness. In fact, it is so weakly put together that using nonstandard analysis it is not difficult to totally unglue it and fill its uncountable holes with equally uncountable infinitesimals and infinites. These so easily created openings in the continuum then become interspersed with an unlimited number of variably sized monads, that is, monads, and monads within monads, and so on ad infinitum, all of which makes the original continuum impossible to recapture with further Dedekind cuts. In themselves, these additions and interpolations are a positive feature in that, phenomenologically speaking, many continua are indeed soft, penetrable, and amenable to unlimited interspersion. The key shortcomings of Dedekind's theory are unity's feeble role among the parts as well as the way the theory overlooks the fact that continua are genetically prior to sets, prior to any abstract gathering and separating. A surprising failure given the historical fact that there was geometry for centuries before there were sets.

The irony of Dedekind's approach lies in its characterization of the continuum through the least absolute and least important of its properties, divisibility, when actually what characterizes the continuum is its indivisibility, the encroachment of parts on one another, and the unbreakable linkages. This is why Dedekind's continuum is so unreal, and why it cannot tinguishable aspects, not cuts; (ii) that some of its regions form a unity while others do not, union being a true or false predicate and not a set-theoretic operation that has no exceptions; (iii) that it has no holes -- no empty regions -- and itself is never a universal depository of aspects and regions but is always open to endless downward analysis and upward synthesis;

(iv) that some of its regions interpenetrate mutually while others do not; and finally (v) that there are always regions and more regions between regions.

3. WHAT IS A CONTINUUM?

A continuum is an extension with distinguishable and often overlapping aspects, aspects in which, in turn, overlapping and nonoverlapping regions can be distinguished. In this extension each distinction discloses an in-between: every division engenders its own bridge. A continuum that is absolutely separable is not a continuum at all. The continuum mends its own tearings like flesh that heals its own wounds. In the continuum, holes are abstract illusions, somewhat like the black holes that current cosmology interprets as not really holes but as a kind of in-between, a link between two worlds that establishes the continuity of one into the other. When we consider an aspect as an intermediate--provisional--universal continuum all by itself, it becomes a "black hole" through which to travel from one "universal" level to another.

A theory for a concrete continuum -- a continuous continuum -- must reflect several fundamental characteristics: (i) its primitiveness -- the fact that the continuum is an ultimate aspect of reality to be described, not a building to be formally erected on a base rooted in set theory; (ii) the fact that the continuum does not gather regions in a set-theoretic fashion -- an axiom of comprehension would distort the unlimited openness, both upward and downward, with which a continuum presents itself in the middle of a vast expanse; (iii) the fact that each "universal" continuum is only an aspect temporarily taken as a universe, an "intermediate universal continuum," to use technically the paradoxical expression already employed (this expression well conveys the antinomic situation in which we find ourselves, perceiving as we must only a limited portion of reality that fades indefinitely toward the very big and the very small, despite our obsessive attempts to describe the world in its entirety as though it were a closed aggregate).

4. MINGLING VERSUS GATHERING.

Dedekind's continuum, being an upwardly directed construction, misses the point that in the real world wholes and parts exchange roles more freely than our habit of "collecting into a whole" allows us to perceive. This tendency to gather, to collect and enclose, is a major source of fallacious descriptions of reality. Wholes do not necessarily gather their parts; instead, the relations between parts and wholes are often like a mingling together. Our technical use of the word mingling will convey that a part often extends beyond many of its wholes. Mingling is of course compatible with gathering, but it generally transcends gathering and erases all stratifications of the part-whole relation. The predicate "being a part of" will be used technically in this sense of mingling, not in that of a set-theoretic gathering of elements or of an inclusion of one set into another.

5. IN THE MIDDLE OF THINGS.

Franz Kafka noted that things never present themselves "by their roots" -- nor in their totality, we must add -- "but by some point or other situated toward the middle." This is an accurate description of man's perceptive predicament. We can only approach reality at its middle, a reality which in effect may well have no roots and be utterly uncollectable in its entirety.

The intermediate universal continuum $\underline{C}$ is at the middle of all its aspects, and each aspect $\underline{C}^r$ is at the middle of all its regions $\underline{x}^r, \underline{y}^r, \underline{z}^r, \ldots$. Only set-theoretic preconceptions make us fail to see how entities are observed in direct experience. Yet $\underline{C}$ itself is only an aspect of a vaster expanse, an aspect that can itself be "horizontally" analyzed into further aspects and regions that are at a lower level than the vaster expanse. This level structure is reminiscent of Hao Wang's theory of positive and negative types -- which also excludes initial levels -- except for two considerations: (i) types are closed to one another, while levels share the same aspects,

taken alternatively as aspects or universes, and (ii) types are arranged in linear succession, whereas levels ramify nonlinearly upward and downward.

6. UNION AS A PREDICATE.

Some objects cannot unite even if one puts them in the same box in perpetuity. The moment that uniting entities becomes synonymous with collecting them into the same set, unification loses its concrete meaning and becomes an abstract, trivialized operation. In what follows, union is to be considered a binary predicate that holds if and only if two continua are truly unified, that is, if there is a definite coalescence between them, be they continguous or mediated by other continua. Thus the intermediate universal continuum $\underline{C}$ coalesces with each of its aspects $\underline{C}^r, \underline{C}^s, \ldots$, unifies with each of them -- which is why they are its aspects, not because of any membership relation. Similarly, each aspect $\underline{C}^r$ coalesces with each of its regions $\underline{x}^r, \underline{y}^r, \underline{z}^r, \ldots$, unifies with each of its regions -- which is why they are *its* regions. $\underline{C}$ and its aspects $\underline{C}^r, \underline{C}^s, \ldots$ are therefore essentially amalgamated.

7. BETWEENNESS.

The intermediate universal continuum $\underline{C}$ interjects itself between any two of its aspects; in symbols, $\underline{B}(\underline{C}^r, \underline{C}, \underline{C}^s)$. Similarly, each aspect interjects itself between any two of its regions, $\underline{B}(\underline{x}^r, \underline{C}^r, \underline{y}^r)$. Betweenness is not an external relation, for $\underline{B}(\underline{C}^r, \underline{C}, \underline{C}^r)$ and $\underline{B}(\underline{x}^r, \underline{C}^r, \underline{x}^r)$ are valid special cases of the two symbolic expressions just given; here, betweenness parallels the intimate coalescence of unification, as well as the intimate meaning of presence that "being a part of" provides. Further, because between two different regions of a given aspect a third can always be interjected, betweenness satisfies a density property similar to the one that rational numbers satisfy in their usual ordering. For this reason it is impossible to produce finite models of continua. Finally, betweenness -- like distinguishabity -- is a

one-directional relation; i.e., $\underline{y}^r$ may lie between $\underline{x}^r$ and $\underline{z}^r$ but not between $\underline{z}^r$ and $\underline{x}^r$.

8. PLACES VERSUS POINTS.

Our ordinary conception of the space-time continuum, even the relativistic one with its overlapping points of view, is a misleading model of the real continuum: it still relies on points as ultimates and converts separation into an absolute. Actually, there are no perceived points, only aspects and regions, and these can overlap just as the different perspectives provided by each system of reference overlap, which is impossible for points. To say that there are no points but only aspects and regions means that the continuum has no ultimate components, neither points nor elements of any kind -- in particular, no atomic regions. The terms for all predicate formulas, then, must be places, regions, or aspects but never points, places that are provisional units of discourse in the same way that any partially disengaged object of our perception is a unit of discourse. Given that there is not the slightest evidence to support the existence of a bottom or a top in our cosmos, we shall operate on the more adequate assumption that there is no truly indivisible atom, just as there is no complete universe, and that, further, it is in the nature of things to be irremediably caught in the middle; i.e., <u>there is only the middle</u> -- an ontological as well as a perceptual fact, in accordance with the principle that places are concrete while points are incurably unreal.

Aspects differ from regions in that they themselves can be temporarily taken as universes; regions cannot. Also, aspects are in the middle of each of their regions as well as between any two of them, that is, aspects are a kind of floating medium for their regions; yet a region is not necessarily in the middle of any two other regions. Aspects, like phases, are snapshots of a transcendent reality, the temporary visitation of a far-reaching entity, i.e., a higher-level continuum whose origin is perhaps remotely located -- just as light visits us daily from afar. This unfolding of aspects within an aspect

takes place at all levels in the never-ending upward and downward ramifications of continua, each aspect always relating to higher and lower aspects. This parallels the macroscopic amplification of microscopic structures and processes that Pascual Jordan considered the essence of life. It also parallels the unlimited self-similarity of fractals without regard for the up or down scale in which they are considered, that is, the fractal's entire geometric pattern reappears without change whatever the size and scope of observation; Benoit Mandelbrot found this characteristic dominant in most natural forms and processes. Regions, since they are patches of aspects, do not have this ability to move from one level to another.

9. RELATIONS OF THE PRIMITIVE PREDICATES.

Since unity, being a part of, and betweenness are independently introduced predicates, it is desirable to differentiate between the following cases. (i) Each aspect coalesces with every one of its regions and is at the middle of any of them. (ii) Many regions intersect because they have another region as a common part without the latter coalescing with any of the former. (iii) Regions are always mediated by other regions but none of these regions need necessarily coalesce. (iv) An aspect is part of many of its regions without their necessarily coalescing. Whenever a region is part of an aspect and vice versa, we say that the whole is part of the part. To see a family trait in a face is an example of an aspect being part of a region, as is the case when parsley gives flavor to a taste -- parsley giving something of itself but not all of itself, staying on the outside yet inside. The obstacle to perceiving these standard facts is that our thinking is overwhelmingly controlled by atomism, a prejudice that is exceedingly difficult to escape because the moment we put down symbols on paper we are already automatically committing ourselves to simple location; and yet, we constantly perceive regions that are "larger" than the aspects of which they are a region.

10. THE INCOMPLETENESS OF ANY ANALYSIS OF THE CONTINUUM.

We must differentiate between the following: (i) distinguishability, i.e., given two regions, a third one is unified with the first but not with the second, a nonsymmetric fact; (ii) nonintersection, i.e., two regions have no third region as a common part; (iii) disunity, i.e., two regions do not coalesce; plus (iv) combinations of all of these. It is possible for a region to be part of another without either one being distinguishable from the other. Further, unity is a form of connection even for nonintersecting regions, regions seemingly disconnected at first sight; indeed, it is common for regions to be united despite a wide breadth of in-between -- painters demonstrate this when they distribute color on a canvass. Finally, disunited regions can have parts in common.

Perceptively we may concentrate on an aspect and exclude its source, but this should not make us forget that any continuum can become a stage at a given level, a phase. Parts II and III will present a horizontal analysis of such stages, an analysis that can be repeated at all levels, keeping in mind that these levels have no linear ordering and no beginning and no end. Because of the infinite ramification of levels up and down at whatever aspect our mind starts, no description of the continuum can ever be complete nor be completed without violating its true nature.

The intermediate universal continuum $\underline{C}$, then, not only is not a universal class but also is not a product graph or a complete universal-algebraic object. Rather, it is a universal-algebraic fragment taken at the middle of things, as it must be; as such, it contains all the internal transformations relative to $\underline{C}$ -- the tearings and joinings between its aspects, for example. And, again, $\underline{C}$ can only be a fragment because it is itself an aspect of many higher-level continua, just as each of $\underline{C}$'s aspects can in turn be taken as a universe: this is the inevitable principle of the relativity of all continua. When attempting to comprehend this fragmentation, this purely relative articu-

lation of any intermediate universal continuum, the built-in atomism of set theory is a serious obstacle; in fact, it destroys our ability to conceive continuity correctly.

11. TOPOLOGY OF CONTINUA WITHOUT SETS.

Ordinary topology is so dependent on one-directional part-whole relationships, and it so weakens the role of the relations of unity and betweenness in the structure of extension, that it is not capable of dealing with the phenomenology of space. Coalescence, two-way part-whole relationships, and mediation of regions and aspects are the essential features of continua. We need a topology based on direct perception, not on abstraction, that is, not on our tendency to collect isolated items. We must always bear in mind that wholes do not necessarily gather parts, and that they often mingle with their parts.

In ordinary topology, closed simple boundaries enclose a region. Concrete boundaries, however, have direction, and may be impenetrable in one direction but open in another. Even Gestalt psychology commits the mistake of holding that a "good Gestalt" is necessarily a form that implies a closed boundary. But an "unfinished" drawing made with an open line should not be judged as demanding that we imagine the completion of a closed contour for our theoretical satisfaction. We can see faces in a cloud surrounded by unfinished, soft, penetrable contours, and these diffuse properties are essential to the face we perceive, inseparable from the perception's quality; in other words, the contours are not only open but must be so described theoretically.

The existence of soft, one-directional boundaries implies that not every simple closed boundary divides an inside from an outside. Actually, some boundaries are their own inside. Only our obsession with univocity makes this difficult to accept, even if it is a usual characteristic of perception. We believe that we draw a line made of points even if we know that we can neither

draw nor perceive a point. And yet, that boundaries can be their own interior should not be surprising when we know that Peano curves fill a plane, that Lebesgue curves have positive area, and that there are surfaces with positive volume.

In his "Topology without Points" Karl Menger outlines a system in which the ultimate objects are "lumps" rather than points. He ends the paper with the following recommendation: "By a lump, we mean something with a well defined boundary. But well defined boundaries are themselves results of limiting processes rather than objects of direct observation. Thus, instead of lumps, we might use at the start something still more vague -- something perhaps which has various degrees of density or at least admits a gradual transition to its complement. Such a theory might be of use for wave mechanics."[1] But what is a curve in a topology of lumps? In Part III we shall offer an answer to this question, "lumps" for us meaning either regions or aspects, regions $\underline{x}^r, \underline{y}^r, \ldots$ and aspects $\underline{C}^r, \underline{C}^s, \ldots$ being local continua, and the intermediate universal continua $\underline{C}', \underline{C}'', \ldots$ being global ones -- a purely relative distinction.

If we can manage not to succumb to prevalent formal-ontologic prejudices, it is easy to see that a boundary does not provide a razor-sharp division, that it often overlaps both its inside and outside -- if it has them -- and is frequently a kind of no-man's-land between, and encroaching upon, the regions immediately beside it.

12. IN SYNTHESIS.

To recapitulate briefly, a multiplicity of regions and aspects -- one constituting a boundary, say -- cannot be gathered in a set-theoretic manner. We distinguish continua partially before we can divide them, and even then new regions keep appearing between regions. Hence, to distinguish is to add; to divide is to multiply the number of viewpoints, to create further possible interpenetrations and coalescences: each act of distinction reveals an in-

between. To sever is too final to be realistically concrete, for there is often a gradual transition from a region to its complements. Distinguishability of entities, again, implies their giving something of themselves but not all.

There is a never ending descent and a never ending ascent in the analysis of the continuum. There is no empty continuum, and there is no absolutely universal continuum. Further, the distinction between local and global transformations is blurred. Every local change has global implications, and every global change affects local relations.

Unity, being a part of, and betweenness are primitive and irreducible relations, each contributing differently to the constitution of continua. It is a mistake to think that this proliferation of predicates makes the continuum a less exact structure. Although Whitehead said that "exactness is an ideal of thought and is only realized in experience by the selection of a route of approximation," it is misleading to identify exactness with clear-cut, razor-sharp definition. There is a stage in the description of reality beyond which sharpness becomes synonymous with inexactness. In a realm of diffuse entities, exactness demands diffuseness. This is a routine paradox of scientific thinking that we must simply accept. Exactness in the sense of full adequacy often requires moving away from our most cherished, most habit-soothing notions, notions which -- like those of set and membership -- only function as barriers after a point, preventing us from grasping the concrete structure of space as it really is.

II.- A FORMAL LANGUAGE.

13. THE PRIMITIVE SYMBOLS.

A three-sorted classical predicate calculus is assumed. The three sorts of individual variables are

(i) $\underline{C}, \underline{C}', \underline{C}'', \ldots$ to range over intermediate universal continua; i.e., aspects taken provisionally as universes (multiverses would be perhaps more appropriate, since $\underline{C}$ is not a closed whole of all its aspects).

(ii) $\underline{C}^r, \underline{C}^s, \underline{C}^t, \ldots$ to represent aspects of an intermediate universal continuum $\underline{C}$; i.e., partial continua of a universal continuum (these aspects are not to be taken as elements or "finished" products of any kind, for each of them may itself be taken as a universe). It is of course possible for the variable superindex $\underline{r}$ to be identical to $\underline{s}$, i.e., for $\underline{C}^r$ to be $\underline{C}^s$.

(iii) $\underline{x}^r, \underline{y}^r, \underline{z}^r, \ldots$, regions of an aspect $\underline{C}^r$, also occasionally denoted $\ldots, \underline{x}^r_{-1}, \underline{x}^r_0, \underline{x}^r_1, \underline{x}^r_2, \ldots$ for convenience.

We take an aspect -- any aspect at the middle of things -- and consider it provisionally as a universe $\underline{C}$. Then we find that $\underline{C}$, in turn, displays aspects $\underline{C}^r, \underline{C}^s, \ldots$, and that each of these aspects then displays regions $\underline{x}^r, \underline{y}^r, \ldots, \underline{x}^x, \underline{y}^s, \ldots$. This constitutes the horizontal analysis of $\underline{C}$'s structure. But we can take any universe $\underline{C}$ and observe that it is merely an aspect of a more sweeping universe $\underline{C}'$, actually, of many universes $\underline{C}', \underline{C}'', \underline{C}''', \ldots$; that is, $\underline{C}$ can be taken as a $\underline{C}'^r$, a $\underline{C}''^s$, a $\underline{C}'''^t$, etc.; this is analysis in the upward direction. Or still, we can take an aspect $\underline{C}^r$ of a given universe $\underline{C}$, and consider it, in turn, as a universe $\underline{C}'$ in itself, with its own aspects

$\underline{C}'^{r}, \underline{C}'^{s}, \ldots$; this is downward analysis. The upward and downward directions of analysis can be pursued limitlessly along all their ramified paths; i.e., we are unavoidably "in the middle of things." There is no way to reach a bottom or a ceiling because neither exists; there is no empty continuum, no absolutely universal continuum. We cannot collect all aspects but can only analyze their structure in a fragmentary way, that is, at a given horizontal level or, in a limited way, at some of their neighboring levels. There is, then, no essential difference between $\underline{C}$ and $\underline{C}^{r}$ insofar as their structure is concerned; $\underline{C}$ is merely the irreducible multiplicity of all its aspects. Whereas Cantor sets collect elements, universal continua spread aspects. The $\underline{C}^{r}, \underline{C}^{s}, \ldots$ are the internal composition of $\underline{C}$. The intermediate universal continua $\underline{C}', \underline{C}'', \ldots$, of which $\underline{C}$ is an aspect, establish the external relations of $\underline{C}$.

In addition we have:

(iv) Three undefined binary predicates, $\cup$, "union," π, "being a part of," and $\underline{N}(\underline{x}^{r}, \underline{y}^{r})$, "$\underline{y}^{r}$ is a neighborhood of $\underline{x}^{r}$." (Union is not an operation.)

(v) A ternary predicate, $\underline{B}$, "betweenness."

Intermediate universal continua, aspects, and regions are all continua; they differ in some of their formal properties as specified by the axioms that follow. Intuitively, these various continua emerge in the order of their successive discernibility. As we concentrate on an aspect $\underline{C}$, arbitrarily chosen, and take it as a world in itself, we should not have the illusion that it is anything more than a patch in the middle of vaster spreads, just as $\underline{C}$'s aspects and regions are patches in $\underline{C}$'s midst.

14. FOUR DEFINED PREDICATES.

(i) $=$, equality.

DEFINITION 14.1: $\underline{x}^r = \underline{y}^r$ if and only if (iff) $(\underline{z}^r)(\cup \underline{x}^r\underline{z}^r = \cup \underline{y}^r\underline{z}^r)$ & $(\underline{z}^r)(\underline{z}^r\pi\underline{x}^r = \underline{z}^r\pi\underline{y}^r)$ & $(\underline{z}^r)(\underline{x}^r\pi\underline{z}^r = \underline{y}^r\pi\underline{z}^r)$.

DEFINITION 14.2: $\underline{x}^r = \underline{y}^s$ iff $(\underline{u}^r)(\underline{v}^s)((\cup \underline{x}^r\underline{u}^r = \cup \underline{y}^s\underline{u}^r)$ & $(\cup \underline{x}^r\underline{v}^s = \cup \underline{y}^s\underline{v}^s))$ & $(\underline{u}^r)((\underline{u}^r\pi\underline{x}^r = \underline{u}^r\pi\underline{y}^s)$ & $(\underline{x}^r\pi\underline{u}^r = \underline{y}^s\pi\underline{u}^r))$ & $(\underline{v}^s)((\underline{v}^s\pi\underline{x}^r = \underline{v}^s\pi\underline{y}^s)$ & $(\underline{x}^r\pi\underline{v}^s = \underline{y}^s\pi\underline{v}^s))$.

Since $\underline{x}^r$, $\underline{C}^r$, and $\underline{C}$ are different sorts of variables, $\underline{x}^r = \underline{C}^r$, $\underline{x}^r = \underline{C}$, and $\underline{C}^r = \underline{C}$ are never the case. However, regions of one aspect can be part of other aspects. Note further that $\underline{x}^r \neq \underline{y}^r$ implies $(E\underline{z}^r)((\cup \underline{x}^r\underline{z}^r$ & $\neg \cup \underline{y}^r\underline{z}^r) \vee (\neg \cup \underline{x}^r\underline{z}^r$ & $\cup \underline{y}^r\underline{z}^r)) \vee (E\underline{z}^r)((\underline{z}^r\pi\underline{x}^r$ & $\neg(\underline{z}^r\pi\underline{y}^r)) \vee (\neg(\underline{z}^r\pi\underline{x}^r)$ & $\underline{z}^r\pi\underline{y}^r)) \vee (E\underline{z}^r)((\underline{x}^r\pi\underline{z}^r$ & $\neg(\underline{y}^r\pi\underline{z}^r)) \vee (\neg(\underline{x}^r\pi\underline{z}^r)$ & $\underline{y}^r\pi\underline{z}^r))$.

(ii) Two complements, $\cup$–Comp and π–Comp.

DEFINITION 14.3: $\cup$–Comp $(\underline{x}^r, \underline{y}^r)$ iff $(\underline{z}^r)(\cup \underline{z}^r\underline{y}^r = \neg \cup \underline{z}^r\underline{x}^r)$.

DEFINITION 14.4: π–Comp$(\underline{x}^r, \underline{y}^r)$ iff $(\underline{z}^r)(\underline{z}^r\pi\underline{y}^r = \neg(\underline{z}^r\pi\underline{x}^r))$ & $(\underline{z}^r)(\underline{y}^r\pi\underline{z}^r = \neg(\underline{x}^r\pi\underline{z}^r))$. Both complements are relative to a given aspect $\underline{C}^r$, but neither gathers or fills regions.

(iii) Δ, distinguishability.

DEFINITION 14.5: $\underline{x}^r\Delta\underline{y}^s$ iff $(E\underline{z}^t)(\cup \underline{z}^t\underline{x}^r$ & $\neg \cup \underline{z}^t\underline{y}^s)$. Δ is not symmetric. The definition extends to aspects as follows: $\underline{C}^r\Delta\underline{C}^s$ iff $(E\underline{z}^t)(\cup \underline{z}^t\underline{C}^r$ & $\neg \cup \underline{z}^t\underline{C}^s)$.

15. AXIOMS FOR UNION.

AXIOM 15.1: $(\underline{x}^r) \cup \underline{x}^r\underline{x}^r$.

AXIOM 15.2: $(\underline{x}^r)(\underline{y}^s)(\cup \underline{x}^r\underline{y}^s \Rightarrow \cup \underline{y}^s\underline{x}^r)$. True in particular if $\underline{r} = \underline{s}$. Union is not necessarily transitive.

AXIOM 15.3: $(\underline{x}^r) \cup \underline{x}^r\underline{C}^r$.

AXIOM 15.4: $(\underline{C}^r) \cup \underline{C}^r\underline{C}$. Every aspect is united with all its regions and the intermediate universal continuum is united with all its aspects. This is the continuum's first basic structural principle.

AXIOM 15.5: $(\underline{x}^r)(E\underline{y}^r) \neg \cup \underline{x}^r\underline{y}^r$. Unity of regions is not universal.

AXIOM SCHEME 15.6: For each $\underline{i} = 2,3,\ldots$, Axiom 15.6.i means $(\underline{x}_1^r)(\underline{x}_2^r)\ldots(\underline{x}_i^r)$
$(\underline{x}_1^r \neq \underline{x}_2^r \;\&\; \underline{x}_1^r \neq \underline{x}_3^r \;\&\; \ldots \;\&\; \underline{x}_{i-1}^r \neq \underline{x}_i^r \Rightarrow$
$(E\underline{x}_{i+1}^r)(\underline{x}_1^r \neq \underline{x}_{i+1}^r \;\&\; \ldots \;\&\; \underline{x}_i^r \neq \underline{x}_{i+1}^r \;\&\; \cup \underline{x}_1^r\underline{x}_{i+1}^r \;\&\; \ldots \;\&\; \cup \underline{x}_i^r\underline{x}_{i+1}^r))$. This axiom implies the existence of an infinity of regions provided that two distinct regions exist.

Note that two aspects $\underline{C}^r$ and $\underline{C}^s$ may not be united, and that the intermediate universal continuum $\underline{C}$ and a region $\underline{x}^r$ may not be united either. Further, unity is independent of simple location; regions which may not be considered contiguous or even near to one another in a superficial sense may form a union.

16. AXIOMS FOR BEING A PART OF.

$\underline{x}^r \pi \underline{y}^r$ reads "$\underline{x}^r$ is a part of $\underline{y}^r$" or, alternatively, "$\underline{y}^r$ is a whole of x^r."

AXIOM 16.1: $(\underline{x}^r)(\underline{x}^r \pi \underline{x}^r)$. Every region is a part of itself.

AXIOM 16.2: $(\underline{x}^r)(\underline{y}^r)(\underline{z}^r)(\underline{x}^r \pi \underline{y}^r \ \& \ \underline{y}^r \pi \underline{z}^r \Rightarrow \underline{x}^r \pi \underline{z}^r)$. Transitivity of π; π is neither necessarily symmetric, nor necessarily asymmetric or antisymmetric, i.e., it is possible to have $\underline{x}^r \neq \underline{y}^r$ and $\underline{x}^r \pi \underline{y}^r \ \& \ \underline{y}^r \pi \underline{x}^r$. Being a part of is a form of presence, not an inclusion.

AXIOM 16.3: $(\underline{x}^r)(E\underline{y}^r)(\underline{x}^r \neq \underline{y}^r \ \& \ \underline{y}^r \pi \underline{x}^r)$.

AXIOM 16.4: $(\underline{x}^r)(E\underline{y}^r)(\underline{x}^r \neq \underline{y}^r \ \& \ \underline{x}^r \pi \underline{y}^r)$. The infinite descent and infinite ascent implied by these two axioms means that each region is the whole of other parts and a part of other wholes. No region is an atom or a complete whole. This property of being unlimitedly analyzable and synthesizable is the continuum's second basic structural principle. The continuum has no points, only places that have places, and it cannot be boxed into a set.

AXIOM 16.5: $(\underline{x}^r)(\underline{x}^r \pi \underline{C}^r \vee \underline{C}^r \pi \underline{x}^r)$. This is a nonexclusive disjunction. Further, $\underline{C}^r \pi \underline{C}^s$ is possible, aspects can be part of other aspects; also, $\underline{C}^r \pi \underline{C}^r$, $\underline{C}^r \pi \underline{x}^s$, $\underline{x}^s \pi \underline{C}^r$, $\underline{x}^r \pi \underline{y}^s$, $\underline{C} \pi \underline{x}^r$, $\underline{x}^r \pi \underline{C}$, $\underline{C}^r \pi \underline{C}$, and $\underline{C} \pi \underline{C}^r$ are all possible (examples are given in Part IV).

AXIOM 16.6: $(\underline{x}^r)(\underline{y}^r)(E\underline{z}^r)(\underline{x}^r \neq \underline{z}^r \ \& \ \underline{y}^r \neq \underline{z}^r \ \& \ \underline{x}^r \pi \underline{z}^r \ \& \ \underline{y}^r \pi \underline{z}^r)$. This axiom subsumes two regions into a third one.

17. AXIOMS FOR BETWEENNESS.

AXIOM 17.1: $(\underline{x}^r)(\underline{y}^r) \neg \underline{B}(\underline{x}^r,\underline{y}^r,\underline{x}^r)$. This is a form of irreflexivity: no region stands between any region $\underline{x}^r$ and $\underline{x}^r$.

AXIOM 17.2: $(\underline{x}^r)(\underline{y}^r)(\neg \underline{B}(\underline{x}^r,\underline{x}^r,\underline{y}^r) \ \& \ \neg \underline{B}(\underline{x}^r,\underline{y}^r,\underline{y}^r))$. A third region is necessary for linkage between regions.

AXIOM 17.3: $\underline{B}(\underline{x}^r,\underline{y}^s,\underline{z}^t) \ \& \ \underline{B}(\underline{z}^t,\underline{u}^m,\underline{v}^n) \ \& \ \underline{y}^s \neq \underline{u}^t \Rightarrow \underline{B}(\underline{y}^s,\underline{z}^t,\underline{u}^m)$. The only transitivity of betweenness, which holds in particular if $\underline{r} - \underline{s} - \underline{t} - \underline{m} - \underline{n}$.

AXIOM 17.4: $\underline{x}^r - \underline{u}^s \ \& \ \underline{y}^r - \underline{v}^s \ \& \ \underline{z}^r - \underline{w}^s \ \& \ \underline{B}(\underline{x}^r,\underline{y}^r,\underline{z}^r) \Rightarrow \underline{B}(\underline{u}^s,\underline{v}^s,\underline{w}^s)$. Substitutivity of betweenness -- invariance with respect to aspects.

AXIOM 17.5: $(\underline{x}^r)(\underline{y}^r)(\underline{x}^r \neq \underline{y}^r \Rightarrow (E\underline{z}^r)\underline{B}(\underline{x}^r,\underline{z}^r,\underline{y}^r))$. Density of betweenness in the regions of a given aspect, even if $\underline{x}^r \pi \underline{y}^r$, or $\cup \underline{x}^r\underline{y}^r$, or both. This is the continuum's third basic structural principle.

AXIOM 17.6: $(\underline{x}^r)(\underline{y}^r)\underline{B}(\underline{x}^r,\underline{C}^r,\underline{y}^r)$. Each aspect is at the middle of all its regions, including $\underline{B}(\underline{x}^r,\underline{C}^r,\underline{x}^r)$.

AXIOM 17.7: $(\underline{C}^r)(\underline{C}^s)\underline{B}(\underline{C}^r,\underline{C},\underline{C}^s)$. $\underline{C}$ is not a container but rather an intermediate between all its aspects, as well as the center of each of them, that is, $\underline{B}(\underline{C}^r,\underline{C},\underline{C}^r)$.

AXIOM 17.8: $(\underline{x}^r)(\underline{y}^s)\underline{B}(\underline{x}^r,\underline{C},\underline{y}^s)$. $\underline{C}$ is the intermediate of all its regions; in particular, $\underline{B}(\underline{x}^r,\underline{C},\underline{x}^r)$.

That each aspect $\underline{C}^r$ is at the center of all its regions and $\underline{C}$ is at the center of all its aspects and all its regions is a form of multiple location

which betweenness provides. Note also that $\underline{B}(\underline{x}^r,\underline{y}^r,\underline{z}^r)$ is compatible with $\neg\ \underline{B}(\underline{z}^r,\underline{y}^r,\underline{x}^r)$; betweenness is not symmetric, and has to be thought of as having direction. Further, $\underline{B}(\underline{C}^r,\underline{C}^s,\underline{C}^t)$ is possible, and so are $\underline{B}(\underline{C}^r,\underline{x}^s,\underline{C}^t)$, $\underline{B}(\underline{C}^r,\underline{x}^s,\underline{x}^t)$, $\underline{B}(\underline{x}^r,\underline{x}^s,\underline{C}^t)$, $\underline{B}(\underline{x}^r,\underline{C}^s,\underline{x}^t)$, and $\underline{B}(\underline{x}^r,\underline{x}^s,\underline{x}^t)$. If $\underline{B}(\underline{x}^r,\underline{C}^s,\underline{x}^t)$, then $\underline{C}^s$ contributes to the external continuity of $\underline{x}^r$ and $\underline{x}^t$. Aspects, therefore, are not necessarily fully separated: they can link and be linked by regions and other aspects, and are always linked by the intermediate universal continuum.

18. AXIOMS FOR COMPLEMENTATION.

AXIOM 18.1: $(\underline{x}^r)(E\underline{y}^r)$ $\cup$–Comp$(\underline{x}^r,\underline{y}^r)$.

AXIOM 18.2: $(\underline{x}^r)(E\underline{y}^r)$ π–Comp$(\underline{x}^r,\underline{y}^r)$.

π–Comp$(\underline{x}^r,\underline{y}^r)$ is compatible with $\cup\ \underline{x}^r\underline{y}^r$, and $\cup$–Comp$(\underline{x}^r,\underline{y}^r)$ is compatible with $\underline{z}^r\pi\underline{x}^r$ & $\underline{z}^r\pi\underline{y}^r$, $\underline{x}^r\pi\underline{z}^r$ & $\underline{y}^r\pi\underline{z}^r$, $\underline{x}^r\pi\underline{C}^r$ & $\underline{y}^r\pi\underline{C}^r$, and $\underline{C}^r\pi\underline{x}^r$ & $\underline{C}^r\pi\underline{y}^r$; i.e., there can be a gradual transition from a region to each of its complements, a transition that Karl Menger mentioned as desirable. Actually, more than one region may be part of $\underline{x}^r$ and its $\cup$–complement; to complement is to fill up even if it involves spilling over.

19. AXIOMS FOR DISTINGUISHABILITY.

Δ is clearly irreflexive, nonsymmetric, and nontransitive.

AXIOM 19.1: $(\underline{x}^r)\underline{C}^r\Delta\underline{x}^r$. An aspect is distinguishable from any of its regions even though regions are indistinguishable from their aspect. Regions may also be distinguishable from the intermediate universal continuum.

AXIOM 19.2: $(\underline{x}^r)(E\underline{y}^r)(\underline{x}^r \Delta \underline{y}^r)$. If $\underline{x}^r \Delta \underline{y}^r$ and $\underline{x}^r \pi \underline{y}^r$ then we say that $\underline{x}^r$ is distinguishable in $\underline{y}^r$. Distinguishability is not a clear-cut separation, for $\underline{x}^r \Delta \underline{y}^r$ is compatible with $\underline{y}^r \Delta \underline{z}^r$ and $\neg(\underline{x}^r \Delta \underline{z}^r)$.

AXIOM 19.3: $\underline{C}^r \Delta \underline{C}$. For each aspect $\underline{C}^r$ there exists at least a region $\underline{x}^s$ united with $\underline{C}^r$ but not with $\underline{C}$.

AXIOM 19.4: $\underline{C} \Delta \underline{C}^r$. No aspect $\underline{C}^r$ is united with all the regions of all the other aspects of $\underline{C}$.

AXIOM 19.5: $(\underline{C}^r)(\underline{C}^s)(\underline{C}^r \Delta \underline{C}^s)$.

Regions behave as all aspects do whenever $\underline{x}^r \Delta \underline{y}^s$. Further, $\underline{B}(\underline{x}^r, \underline{y}^s, \underline{z}^t)$ is consistent with $\neg \cup \underline{x}^r \underline{y}^s$, $\neg \cup \underline{y}^s \underline{z}^t$, and $\neg \cup \underline{x}^r \underline{z}^t$, and therefore consistent with $\underline{x}^r \Delta \underline{y}^s$, $\underline{y}^s \Delta \underline{z}^t$, $\underline{x}^r \Delta \underline{z}^t$ or their negations. In other words, betweenness is independent of unity.

20. RELATIONAL AXIOMS.

AXIOM 20.1: $(\underline{x}^r)(E\underline{y}^r)(\underline{z}^r)(\cup \underline{x}^r \underline{z}^r \Rightarrow \underline{z}^r \pi \underline{y}^r \vee \underline{y}^r \pi \underline{z}^r)$. Mingling of all regions united to a given region.

AXIOM 20.2: $(\underline{x}^r)(E\underline{y}^r)(\underline{z}^r)(\underline{z}^r \pi \underline{x}^r \vee \underline{x}^r \pi \underline{z}^r \Rightarrow \cup \underline{z}^r \underline{y}^r)$. $\cup$-gathering of all the parts and wholes of a given region.

AXIOM 20.3: $\underline{x}^r \pi \underline{y}^r \;\&\; \underline{x}^r \pi \underline{z}^r \;\&\; \underline{x}^r \neq \underline{y}^r \;\&\; \underline{x}^r \neq \underline{z}^r \;\&\; \underline{y}^r \neq \underline{z}^r \Rightarrow \underline{B}(\underline{y}^r, \underline{x}^r, \underline{z}^r)$. A part mediates between the wholes it is a part of.

AXIOM 20.4: $\underline{x}^r \pi \underline{y}^r \;\&\; \underline{z}^r \pi \underline{y}^r \;\&\; \underline{x}^r \neq \underline{y}^r \;\&\; \underline{z}^r \neq \underline{y}^r \;\&\; \underline{x}^r \neq \underline{z}^r \Rightarrow \underline{B}(\underline{x}^r, \underline{y}^r, \underline{z}^r)$. A whole mediates between any two of its parts.

Note that regions -- $\underline{x}^r$ and $\underline{y}^r$, and $\underline{y}^r$ and $\underline{z}^r$, say -- can be contiguous in the sense of $\underline{B}(\underline{x}^r, \underline{y}^r, \underline{z}^r)$ without $\underline{x}^r$ and $\underline{y}^r$ on the one hand, and $\underline{y}^r$ and $\underline{z}^r$ on the other, having any part in common.

AXIOM 20.5: $(\underline{x}^r)(\underline{y}^r)(\mathrm{E}\underline{z}^r)(\underline{w}^r)(\underline{w}^r \pi \underline{z}^r \vee \underline{z}^r \pi \underline{w}^r \equiv (\underline{w}^r \pi \underline{x}^r \vee \underline{x}^r \pi \underline{w}^r) \;\&\; \cup \underline{w}^r \underline{y}^r)$. Axiom of $\cup$-comprehension: $\underline{z}^r$ is the mingling of all parts and all wholes of $\underline{x}^r$ united to $\underline{y}^r$.

The following additional definitions are now in order to differentiate between the various kinds of separation obtainable.

DEFINITION 20.1: Two regions $\underline{x}^r$ and $\underline{y}^r$ are called cohesive iff $\cup \underline{x}^r \underline{y}^r \;\&\; (\underline{x}^r \pi \underline{y}^r \vee \underline{y}^r \pi \underline{x}^r)$.

DEFINITION 20.2: $\underline{x}^r$ and $\underline{y}^r$ are called disjoint iff $\cup \underline{x}^r \underline{y}^r \;\&\; \neg(\underline{x}^r \pi \underline{y}^r \vee \underline{y}^r \pi \underline{x}^r)$.

DEFINITION 20.3: $\underline{x}^r$ and $\underline{y}^r$ are called detached iff $\neg \cup \underline{x}^r \underline{y}^r \;\&\; (\underline{x}^r \pi \underline{y}^r \vee \underline{y}^r \pi \underline{x}^r)$.

DEFINITION 20.4: $\underline{x}^r$ and $\underline{y}^r$ are called severed iff $\neg \cup \underline{x}^r \underline{y}^r \;\&\; \neg(\underline{x}^r \pi \underline{y}^r \vee \underline{y}^r \pi \underline{x}^r)$. Note that even severance is not complete because it is compatible with the existence of a common part $\underline{z}^r$, i.e., $\underline{z}^r \pi \underline{x}^r \;\&\; \underline{z}^r \pi \underline{y}^r$.

DEFINITION 20.5: $\underline{x}^r$ and $\underline{y}^r$ intersect -- in symbols, $\cap \underline{x}^r\underline{y}^r$ -- iff $(E\ \underline{z}^r)(\underline{z}^r\pi\underline{x}^r\ \&\ \underline{z}^r\pi\underline{y}^r)$.

DEFINITION 20.6: $\underline{x}^r$ and $\underline{y}^r$ form a gap iff $\cup \underline{x}^r\underline{y}^r\ \&\ \neg \cap \underline{x}^r\underline{y}^r$.

DEFINITION 20.7: $\underline{y}^r$ is a wound between $\underline{x}^r$ and $\underline{y}^r$ iff $\underline{B}(\underline{x}^r,\underline{y}^r,\underline{z}^r)\ \&\ \neg(\cup \underline{x}^r\underline{y}^r \vee \cup \underline{y}^r\underline{z}^r)$.

DEFINITION 20.8: $\underline{y}^r$ is an incision between $\underline{x}^r$ and $\underline{y}^r$ iff $\underline{B}(\underline{x}^r,\underline{y}^r,\underline{z}^r)\ \&\ \neg(\underline{y}^r\pi\underline{x}^r \vee \underline{x}^r\pi\underline{y}^r \vee \underline{y}^r\pi\underline{z}^r \vee \underline{z}^r\pi\underline{y}^r)$.

DEFINITION 20.9: $\underline{y}^r$ is an internal immediate constituent of $\underline{x}^r$ iff $\underline{B}(\underline{x}^r,\underline{y}^r,\underline{z}^r)\ \&\ \cap \underline{x}^r\underline{y}^r$ (or alternatively of $\underline{z}^r$ iff $\underline{B}(\underline{x}^r,\underline{y}^r,\underline{z}^r)\ \&\ \cap \underline{y}^r\underline{z}^r$).

DEFINITION 20.10: $\underline{y}^s$ is an external immediate constituent of $\underline{x}^r$ iff $\underline{r} \neq \underline{s}\ \&\ \underline{B}(\underline{x}^r,\underline{y}^s,\underline{z}^t)\ \&\ \cap \underline{x}^r\underline{y}^s$ (or alternatively of $\underline{z}^t$ iff $\underline{s} \neq \underline{t}\ \&\ \underline{B}(\underline{x}^r,\underline{y}^s,\underline{z}^t)\ \&\ \cap \underline{y}^s\underline{z}^t$).

DEFINITION 20.11: $\underline{z}^r$ has multiple $\cup$-location in $\underline{x}^r$ and $\underline{y}^r$ iff $\neg \cap \underline{x}^r\underline{y}^r\ \&\ (\cup \underline{z}^r\underline{x}^r\ \&\ \cup \underline{z}^r\underline{y}^r)$.

DEFINITION 20.12: $\underline{z}^r$ has multiple π-location in $\underline{x}^r$ and $\underline{y}^r$ iff $\neg \cup \underline{x}^r\underline{y}^r\ \&\ (\underline{z}^r\pi\underline{x}^r\ \&\ \underline{z}^r\pi\underline{y}^r)$. Multiple location has two meanings: (i) that of a region being united with nonintersecting regions, and (ii) that of a region being part of disunited regions.

III.- TOPOLOGY WITHOUT SETS.

21. NEIGHBORHOODS.

Associated with each region $\underline{x}^r$ there are other regions $\underline{y}^r$ called $\underline{x}^r$-neighborhoods; the latter are not necessarily wholes or parts of $\underline{x}^r$. This relation is denoted $\underline{N}(\underline{x}^r,\underline{y}^r)$.

DEFINITION 21.1: An aspect $\underline{C}^r$ is called a topological continuum iff the following six axioms are satisfied .

AXIOM 21.1: $(\underline{x}^r)\underline{N}(\underline{x}^r,\underline{x}^r)$. Every region is a neighborhood of itself.

AXIOM 21.2: $(\underline{x}^r)(E\underline{y}^r)(\underline{x}^r \neq \underline{y}^r \ \& \ \underline{N}(\underline{x}^r,\underline{y}^r))$. Each region has at least one neighborhood other than itself.

AXIOM 21.3: $(\underline{x}^r)(\underline{y}^r)(\underline{N}(\underline{x}^r,\underline{y}^r) \Rightarrow \cup\, \underline{x}^r\underline{y}^r)$. Each region is united to all its neighborhoods.

AXIOM 21.4: $\underline{N}(\underline{x}^r,\underline{y}^r) \ \& \ \underline{N}(\underline{x}^r,\underline{z}^r) \Rightarrow (E\underline{w}^r)(\underline{w}^r \neq \underline{y}^r \ \& \ \underline{w}^r \neq \underline{z}^r \ \& \ \underline{N}(\underline{x}^r,\underline{w}^r) \ \& \ (\underline{w}^r\pi\underline{y}^r \vee \underline{y}^r\pi\underline{w}^r) \ \& \ (\underline{w}^r\pi\underline{z}^r \vee \underline{z}^r\pi\underline{w}^r))$. Given two neighborhoods of the same region, there is always a third that mingles with the other two.

AXIOM 21.5: $\underline{x}^r \neq \underline{y}^r \Rightarrow (E\underline{z}^r)(\underline{N}(\underline{x}^r,\underline{z}^r) \ \& \ \neg\, \underline{N}(\underline{y}^r,\underline{z}^r))$. The first separation property.

AXIOM 21.6: $\underline{N}(\underline{x}^r,\underline{y}^r) \Rightarrow (\underline{z}^r\pi\underline{y}^r \Rightarrow (E\underline{w}^r)(\underline{N}(\underline{z}^r,\underline{w}^r) \ \& \ \underline{N}(\underline{w}^r,\underline{y}^r))) \ \&$
$(\underline{y}^r\pi\underline{z}^r \Rightarrow (E\underline{w}^r)(\underline{N}(\underline{w}^r,\underline{z}^r) \ \& \ \underline{N}(\underline{y}^r,\underline{w}^r)))$.

DEFINITION 21.2: A region $\underline{x}^r$ is a topological continuum iff the six preceding axioms obtain relativized to all the parts and wholes of $\underline{x}^r$.

Note that, in contrast to ordinary set-theoretic topology, there is no empty neighborhood (Axiom 21.4 involves a mingling, not an intersection of neighborhoods). Note also that given that an aspect can be part of some of its neighborhoods (as can a topological region $\underline{x}^r$ to which we relativize Definition 21.1), we cannot talk about $\underline{C}^r$ (or respectively $\underline{x}^r$) as being the "entire" topological space or the "entire" topological continuum. Even topologically speaking, topological continua are "at the middle of things." Further, note once more that $\underline{N}(\underline{x}^r,\underline{y}^r)$ is compatible with $\neg(\underline{x}^r\pi\underline{y}^r)$ & $\neg(\underline{y}^r\pi\underline{x}^r)$ -- a neighborhood is not necessarily contiguous in the part-whole sense to the region for which it is a neighborhood.

22. SEPARATION PROPERTIES.

In addition to Axiom 21.5, some topological continua may have one or both of the following separation properties.

AXIOM 22.1: $\neg\cup\underline{x}^r\underline{y}^r \Rightarrow (E\underline{z}^r)(E\underline{w}^r)(\underline{z}^r \neq \underline{w}^r \;\&\; \underline{N}(\underline{x}^r,\underline{z}^r) \;\&\; \underline{N}(\underline{y}^r,\underline{w}^r) \;\&\; \neg\cap\underline{z}^r\underline{w}^r)$.

AXIOM 22.2: $\underline{N}(\underline{x}^r,\underline{y}^r) \;\&\; \underline{z}^r\pi\underline{y}^r \;\&\; \underline{z}^r \neq \underline{x}^r \Rightarrow (E\underline{w}^r)(\underline{N}(\underline{x}^r,\underline{w}^r) \;\&\; \neg(\underline{z}^r\pi\underline{w}^r \vee \underline{w}^r\pi\underline{z}^r))$.

Note that in keeping with the general motivation, the separation properties given by Axioms 21.5, 22.1, and 22.2 do not imply absolute division: two united regions may be separated by distinct neighborhoods just as two disunited regions may be separated by distinct neighborhoods, and yet in both cases one region may be part of the other. There is, then, no assurance of complete divisibility in the sense of a Dedekind cut--which actually cuts nothing that was not already cut to begin with.

23. CLUSTER REGIONS: INTERIOR AND EXTERIOR.

DEFINITION 23.1: A region y^r is a cluster region of x^r -- in symbols $C\ell(x^r,y^r)$ -- iff $(z^r)N(y^r,z^r) \Rightarrow ((Ew^r)(w^r\pi z^r \ \&\ (w^r\pi x^r \vee x^r\pi w^r) \ \&\ \neg(w^r\pi y^r)))$ $\vee\ ((Ew^r)(z^r\pi w^r \ \&\ (w^r\pi x^r \vee x^r\pi w^r) \ \&\ \neg(y^r\pi w^r)))$. That is, given a neighborhood of a cluster region y^r of x^r, either (i) this neighborhood has a part that is either a part or a whole of x^r but is not a part of the cluster region itself, or (ii) it is the part of a region w^r that is either a part or a whole of x^r but is not itself a whole of the cluster region y^r.

DEFINITION 23.2: A region x^r is open iff every cluster region of x^r is a part of x^r.

DEFINITION 23.3: A region x^r is closed iff every cluster region of x^r is a whole of x^r.

A region can be both open and closed because regions can be both part and whole of another region.

DEFINITION 23.4: A region x^r is interior to y^r iff $(z^r)(N(x^r,z^r) \Rightarrow z^r\pi y^r)$.

DEFINITION 23.5: A region x^r is exterior to y^r iff $(z^r)(N(z^r,x^r) \Rightarrow y^r\pi z^r)$.

A region can be both interior and exterior to another one.

24. BOUNDARIES.

Since we have no points, the topology that is being described here is in line with Karl Menger's geometry of lumps. An obvious question ensues: What is a line in a geometry of lumps? The following definition answers this question.

DEFINITION 24.1: A sequence of regions $\ldots,\underset{\sim}{x}^r_{-k},\ldots,\underset{\sim}{x}^r_{-1},\underset{\sim}{x}^r_0,\underset{\sim}{x}^r_1,\ldots,\underset{\sim}{x}^r_k,\ldots$ is called a one-sided line iff all the following expressions hold:
$\ldots,\underset{\sim}{B}(\underset{\sim}{x}^r_{-k},\underset{\sim}{x}^r_{-k+1},\underset{\sim}{x}^r_{-k+2}),\ldots,\underset{\sim}{B}(\underset{\sim}{x}^r_0,\underset{\sim}{x}^r_1,\underset{\sim}{x}^r_2),\ldots,\underset{\sim}{B}(\underset{\sim}{x}^r_{k-1},\underset{\sim}{x}^r_k,\underset{\sim}{x}^r_{k+1}),\ldots$. Lines have direction and can of course have two directions (two-sided lines).

DEFINITION 24.2: A sequence of regions $\underset{\sim}{x}^r_1,\underset{\sim}{x}^r_2,\ldots,\underset{\sim}{x}^r_n,\underset{\sim}{x}^r_1$ is called a one-sided closed line iff the following expression holds:
$\underset{\sim}{B}(\underset{\sim}{x}^r_1,\underset{\sim}{x}^r_2,\underset{\sim}{x}^r_3)$ & ... & $\underset{\sim}{B}(\underset{\sim}{x}^r_{n-1},\underset{\sim}{x}^r_n,\underset{\sim}{x}^r_1)$ & $\underset{\sim}{B}(\underset{\sim}{x}^r_n,\underset{\sim}{x}^r_1,\underset{\sim}{x}^r_2)$.

DEFINITION 24.3: The sequence $\underset{\sim}{x}^r_1,\underset{\sim}{x}^r_2,\ldots,\underset{\sim}{x}^r_n,\underset{\sim}{x}^r_1$ is called a two-sided closed line iff it is a one-sided closed line and $\underset{\sim}{x}^r_1,\underset{\sim}{x}^r_n,\underset{\sim}{x}^r_{n-1},\ldots,\underset{\sim}{x}^r_1$ is also a one-sided closed line.

Note that a finite number of regions suffices to determine a closed line, although, given the density of betweenness, between any two regions in a line an infinity of regions is always interspersed. One-sided closed lines which are not two-sided are similar to Möbius strips in that they can never encircle an inside; only two-sided lines can (see the following definitions). These one-sided closed lines are "soft" in the sense that they can be travelled in one direction but have holes in the other direction. In contrast, a two-sided closed line can even be its own inside.

DEFINITION 24.4: The region $\underset{\sim}{y}^r$ constitutes by itself a one-sided boundary of $\underset{\sim}{x}^r$ iff $(\underset{\sim}{z}^r)(\underset{\sim}{B}(\underset{\sim}{x}^r,\underset{\sim}{w}^r,\underset{\sim}{z}^r) \Rightarrow \underset{\sim}{w}^r - \underset{\sim}{y}^r)$. If, in addition, $(\underset{\sim}{z}^r)(\underset{\sim}{B}(\underset{\sim}{z}^r,\underset{\sim}{w}^r,\underset{\sim}{x}^r) \Rightarrow \underset{\sim}{w}^r - \underset{\sim}{y}^r)$, then $\underset{\sim}{y}^r$ is called a boundary of $\underset{\sim}{x}^r$.

DEFINITION 24.5: Let $\underset{\sim}{x}^r_1,\ldots,\underset{\sim}{x}^r_n,\underset{\sim}{x}^r_1$ be a two-sided closed line, this line is called a boundary of $\underset{\sim}{z}^r$ iff
$(E\underset{\sim}{y}^r_1)\underset{\sim}{B}(\underset{\sim}{z}^r,\underset{\sim}{x}^r_1,\underset{\sim}{y}^r_1)$ & $(E\underset{\sim}{y}^r_2)\underset{\sim}{B}(\underset{\sim}{z}^r,\underset{\sim}{x}^r_2,\underset{\sim}{y}^r_2)$ & ... & $(E\underset{\sim}{y}^r_n)\underset{\sim}{B}(\underset{\sim}{z}^r,\underset{\sim}{x}^r_n,\underset{\sim}{y}^r_n)$. The region $\underset{\sim}{z}^r$ is

called the inside of the boundary. If there is no region $\underline{z}^r$ which is the inside of the closed line $\underline{x}_1^r,\ldots,\underline{x}_n^r,\underline{x}_1^r$, and $(E\underline{y}^r)(\underline{B}(\underline{x}_1^r,\underline{x}_2^r,\underline{y}^r)$ & $\underline{B}(\underline{x}_2^r,\underline{x}_3^r,\underline{y}^r)$ & ... & $\underline{B}(\underline{x}_{n-1}^r,\underline{x}_n^r,\underline{y}^r)$ & $\underline{B}(\underline{x}_n^r,\underline{x}_1^r,\underline{y}^r))$, then we say that the closed line is its own inside, even if some or all the regions of the closed line are the inside of other boundaries.

DEFINITION 24.6: A region $\underline{x}^r$ is called a medium of a one-sided line iff $\ldots,\underline{B}(\underline{y}_{-k}^r,\underline{x}^r,\underline{y}_{-k+1}^r),\ldots,\underline{B}(\underline{y}_0^r,\underline{x}^r,\underline{y}_1^r),\ldots,\underline{B}(\underline{y}_k^r,\underline{x}^r,\underline{y}_{k+1}^r),\ldots$ all hold. In this case, the regions of the one-sided line are "immersed" in $\underline{x}^r$.

DEFINITION 24.7: A boundary $\underline{x}_1^r,\ldots,\underline{x}_n^r,\underline{x}_1^r$ is a universal boundary iff it is a boundary of every region $\underline{z}^r$ of $\underline{C}^r$.

Note that a kind of compactness, a Heine-Borel type of property, obtains in that despite the density of betweenness a boundary is determined by a finite number of regions.

Jordan's Theorem revisited. Every boundary of a region $\underline{z}^r$ articulates the topological continuum $\underline{C}^r$ into four domains: (i) the inside $\underline{z}^r$ of the boundary, (ii) the boundary itself, (iii) the $\cup$-complement of $\underline{z}^r$, and (iv) the π-complement of $\underline{z}^r$, the latter two being the two corresponding outsides of the boundary. If the boundary is its own inside, then it has no outside. The four domains may overlap; further, a region $\underline{z}^r$ may have more than one boundary, and a boundary more than one inside.

25. DEGREES OF TRANSITION.

Let $\cup$-Comp $(\underline{x}^r,\underline{y}^r)$ or π-Comp $(\underline{x}^r,\underline{y}^r)$. According to the density property of betweenness, $(E\underline{z}^r)\underline{B}(\underline{x}^r,\underline{z}^r,\underline{y}^r)$ in either case; this region $\underline{z}^r$ provides what Menger calls "a gradual transition" of a region to its complement. Because of the density of betweenness, such transition is always infinitely gradual.

In general, we can now define degrees of transition in terms of neighborhoods as follows.

DEFINITION 25.1: The degree of transition from a region $\underline{x}^r$ to a region $\underline{y}^r$ is finite and equal to $\underline{k}$ iff a finite number of $\underline{x}^r$-neighborhoods $\underline{z}_1^r,\ldots,\underline{z}_k^r$ satisfy $\underline{B}(\underline{x}^r,\underline{z}_1^r,\underline{z}_2^r)$ & $\underline{B}(\underline{z}_1^r,\underline{z}_2^r,\underline{z}_3^r)$ & ... & $\underline{B}(\underline{z}_{k-1}^r,\underline{z}_k^r,\underline{y}^r)$, and $\underline{k}$ is the least number of neighborhoods with such property. The degree of transition from $\underline{y}^r$ to $\underline{x}^r$ may be different or nonexistent.

26. TEARINGS AND JOININGS.

DEFINITION 26.1: $\underline{T}(\underline{x}^r,\underline{y}^s,\underline{z}^s)$ -- read "$\underline{y}^s,\underline{z}^s$ are a tearing of $\underline{x}^r$" -- iff $\cup\,\underline{x}^r\underline{y}^s$ & $\cup\,\underline{x}^r\underline{z}^s$ & $\neg\cup\,\underline{y}^s\underline{z}^s$ & $\neg\cap\,\underline{y}^s\underline{z}^s$.

DEFINITION 26.2: $\underline{J}(\underline{x}^r,\underline{y}^r,\underline{z}^s)$ -- read "$\underline{z}^s$ is a joining of $\underline{x}^r$ and $\underline{y}^r$" -- iff $\neg\cup\,\underline{x}^r\underline{y}^r$ & $\neg\cap\,\underline{x}^r\underline{y}^r$ & $\cup\,\underline{x}^r\underline{z}^s$ & $\cup\,\underline{y}^r\underline{z}^s$. Tearings and joinings take place inter and intra aspects -- the latter when $\underline{r} = \underline{s}$ -- and for each region and pair of regions there are, respectively, at least as many tearings and joinings as aspects, as the following axioms state.

AXIOM 26.1: $(\underline{x}^r)(\underline{C}^s)(E\underline{y}^s)(E\underline{z}^s)\underline{T}(\underline{x}^r,\underline{y}^s,\underline{z}^s)$.

AXIOM 26.2: $(\underline{x}^r)(\underline{y}^r)(\underline{C}^s)(E\underline{z}^s)\underline{J}(\underline{x}^r,\underline{y}^r,\underline{z}^s)$.

27. HOMEOMORPHISMS.

DEFINITION 27.1: A function between two aspects or on an aspect into itself is a relation $\underline{F}(\underline{x}^r,\underline{y}^s)$ that satisfies $\underline{F}(\underline{x}^r,\underline{y}^s)$ & $\underline{F}(\underline{x}^r,\underline{z}^s) \Rightarrow \underline{y}^s = \underline{z}^s$. Note that $\underline{F}(\underline{x}^r,\underline{y}^s)$ is compatible with $\underline{F}(\underline{x}^r,\underline{y}^t)$ provided that $s \neq t$ and $\underline{y}^s \neq \underline{y}^t$. A

function can have many values but no more than one distinct one for each aspect.

DEFINITION 27.2: A biunivocal function between $\underline{C}^r$ and $\underline{C}^s$ is a function $\underline{F}(\underline{x}^r,\underline{y}^s)$ that satisfies $\underline{F}(\underline{x}^r,\underline{y}^s)$ & $F(\underline{z}^r,\underline{y}^s) \Rightarrow \underline{x}^r = \underline{z}^r$.

DEFINITION 27.3: Let us indicate by $\underline{x}^s,\underline{y}^s,\ldots$ the respective images of $\underline{x}^r,\underline{y}^r,\ldots$ under a function $\underline{F}(\underline{x}^r,\underline{x}^s)$. Given a biunivocal function $\underline{F}(\underline{x}^r,\underline{x}^s)$, this function is a homeomorphism from $\underline{x}^r$ to $\underline{x}^s$ iff $(\underline{z}^r)(\underline{N}(\underline{x}^r,\underline{z}^r) = \underline{N}(\underline{x}^s,\underline{z}^s))$ & $(\underline{z}^r)(\underline{C\ell}(\underline{x}^r,\underline{z}^r) = \underline{C\ell}(\underline{x}^s,\underline{z}^s))$ & $(\underline{z}^r)(\underline{y}^r)(\underline{B}(\underline{x}^r,\underline{y}^r,\underline{z}^r) = \underline{B}(\underline{x}^s,\underline{y}^s,\underline{z}^s))$ & $(\underline{z}^r)(\underline{y}^r)(\underline{B}(\underline{y}^r,\underline{x}^r,\underline{z}^r) = \underline{B}(\underline{y}^s,\underline{x}^s,\underline{z}^s))$ & $(\underline{z}^r)(\underline{y}^r)(\underline{B}(\underline{y}^r,\underline{z}^r,\underline{x}^r) = \underline{B}(\underline{y}^s,\underline{z}^s,\underline{x}^s))$. Homeomorphisms preserve the neighborhoods and cluster regions of $\underline{x}^r$ as well as its relations of betweenness; $\underline{x}^r$ and $\underline{x}^s$ are then called homeomorphic regions.

Note that homeomorphic regions are compatible with their tearing and joining; not surprising, since tearings are not absolute separations and joinings are not complete fusions. Tearings and joinings may not alter the topological structure of continua, this structure being determined by the neighborhood systems, the cluster regions, and the relations of betweenness. The topological properties of a topological continuum -- the ones preserved by homeomorphisms -- are those of boundary and inside, for example, not the complements. Rather than studying those qualitative, intrinsic properties of space invariant under "stretchings and bendings without tearings or joinings," topology in the concrete sense that we are giving it deals with properties of general location that are compatible with some tearings and joinings, properties such as being the inside of a boundary, being a boundary, being a neighborhood, being a cluster region, and, most fundamental of all, being in between. Topology in this sense does not necessarily imply the preservation of unions or part-whole relationships.

28. DIMENSION.

DEFINITION 28.1: The dimension of a boundary is the number of its regions.

DEFINITION 28.2: The inner dimension of an aspect $\underline{C}^r$ at a region $\underline{x}^r$ is the least dimension (greater than one) of all boundaries of which $\underline{x}^r$ is an inside. Inner dimension is a local property of aspects, and varies from region to region.

DEFINITION 28.3: The outer dimension of $\underline{C}^r$ is the supremum of all its inner dimensions at each of its regions, if such supremum (a natural number) exists.

DEFINITION 28.4: If all aspects $\underline{C}^r, \underline{C}^s, \ldots$ of $\underline{C}$ have an outer dimension, the supremum of these dimensions, if it exists, is the dimension of C.

DEFINITION 28.5: The dimension of a region x^r is the degree of transition of $\underline{x}^r$ to its $\cup$-complement.

Except for this last definition, all other dimensions are defined in terms of the $\mathcal{B}$-predicate, and are, therefore, homeomorphic invariants. Note also that inner and outer do not bear the usual connotations of simple location. This is in line with the fact that the $\cup$-complement of a region $\underline{x}^r$, for example, may have as parts, parts of $\underline{x}^r$. The difficulty in comprehending all this originates in the visual fallacy that a boundary absolutely severs a two-dimensional continuum, a fallacy induced, say, by simply drawing a line on paper. This leads us to believe unconsciously that any line, closed or not, stands out as if detached, figure against a clearly and absolutely separated background. For this reason, the graphic structure of writing -- which is the

way formal logic is usually presented -- immediately prejudices our thinking and deforms our perception to make it fit the nature of our visual symbols and the well-ordered theories derived from them. Thus, we are always surprised by the stubborn resistance that reality has to being described in neat linear order. Like it or not, a boundary $\underset{\sim}{x}^r_1,\ldots,\underset{\sim}{x}^r_n,\underset{\sim}{x}^r_1$ is compatible with $\underset{\sim}{B}(\underset{\sim}{x}^r_2,\underset{\sim}{x}^r_1,\underset{\sim}{x}^r_3)$ and $\underset{\sim}{B}(\underset{\sim}{x}^r_4,\underset{\sim}{x}^r_1,\underset{\sim}{x}^r_3)$ -- a fact of nature.

IV.- TOWARD A CONTINUOS GRAMMAR.

29. INTERPRETATIONS AND MODELS.

Although interpretations are syntactic when based on functions that map the objects of one formal theory into those of another, they still have an essentially semantic character because the values of the function -- the formal objects of the second theory -- give meaning to the formal objects of the first theory. This is the case also when a concrete object of any nature -- a road sign, a souvenir trinket -- refers to another concrete object, the latter providing the meaning of the former. These and other examples should be sufficient reminders that logic's current set-theoretic approach to interpretations and models is not only relatively recent but of a special kind as well. The set-theoretic approach imposes grave restrictions on the nature and structure of models, and makes semantics too much a servant of set theory, hence far removed from the infinite variety of concrete meanings, meanings always imbued with overlappings, inconclusiveness, gradual fadings, multiple locations, and continuous connotations. The following interpretations and models are continua as specified by the definitions of this section. (Note that the standard logical definition of satisfiability does not necessarily require that the values of the interpreted formal variables be members of a set.)

Let $\underset{\sim}{C}$ be a concrete intermediate universal continuum having aspects

$\underset{\sim}{C}^r, \underset{\sim}{C}^s, \ldots$, which in turn have regions $\underset{\sim}{x}^r, \underset{\sim}{y}^s, \ldots$, all satisfying the axioms of Part II, i.e., $\cup$, π, and $\underline{B}$ each refers to a specific relation between the entities of $\underset{\sim}{C}$. Given the correct number of appropriate entities of $\underset{\sim}{C}$ for each predicate, the concrete relations $\cup$, π, and $\underline{B}$ always either hold or do not hold for such entities -- not both -- and this holding property is given as composing the structure of $\underset{\sim}{C}$, and denoted by $\models$. (Again, there is no reason to attach to "holding" the meaning of an $\underline{n}$-tuple of individuals that belongs to an $\underline{n}$-ary relation. Holding is a primitive correspondence that obtains for some interpreted formal entities in some structures $\underset{\sim}{C}$ and not in others.)

Further, let us agree to distinguish the formal italicized symbols of the language in Part II from the corresponding interpretive entities of $\underset{\sim}{C}$ by writing the latter in bold face. Thus, $(\underline{x}^r, \underset{\sim}{x}^r)$, $(\underline{C}^r, \underset{\sim}{C}^r)$, $(\underline{C}, \underset{\sim}{C})$ indicate a specific valuation of the formal symbols $\underline{x}^r$, $\underline{C}^r$, and $\underline{C}$, a valuation which assigns to the latter the concrete continua $\underset{\sim}{x}^r$, $\underset{\sim}{C}^r$, and $\underset{\sim}{C}$ respectively.

DEFINITION 29.1: Given a concrete intermediate universal continuum $\underset{\sim}{C}$ and a valuation $(\underline{x}_i^r, \underset{\sim}{x}_i^r)$, $(\underline{C}^r, \underset{\sim}{C}^r)$, then we say that this valuation satisfies the formula $\underline{x}_1^r \pi \underline{x}_2^s$ iff $\underset{\sim}{x}_1^r \pi \underset{\sim}{x}_2^s$ holds in $\underset{\sim}{C}$. In symbols $(\underline{x}_1^r, \underset{\sim}{x}_1^r), (\underline{x}_2^s, \underset{\sim}{x}_2^s) \models \underline{x}_1^r \pi \underline{x}_2^s$. Similarly, $(\underline{C}^r, \underset{\sim}{C}^r), (\underline{x}_1^s, \underset{\sim}{x}_1^s) \models \underline{C}^r \pi \underline{x}_1^s$ iff $\underset{\sim}{C}^r \pi \underset{\sim}{x}_1^s$ holds in $\underset{\sim}{C}$.

$(\underline{x}_1^r, \underset{\sim}{x}_1^r), (\underline{x}_2^s, \underset{\sim}{x}_2^s) \models \cup \underline{x}_1^r \underline{x}_2^s$ iff $\cup \underset{\sim}{x}_1^r \underset{\sim}{x}_2^s$ holds in $\underset{\sim}{C}$.

$(\underline{x}_1^r, \underset{\sim}{x}_1^r), (\underline{x}_2^s, \underset{\sim}{x}_2^s), (\underline{x}_3^t, \underset{\sim}{x}_3^t) \models \underline{B}(\underline{x}_1^r, \underline{x}_2^s, \underline{x}_3^t)$ iff $\underline{B}(\underset{\sim}{x}_1^r, \underset{\sim}{x}_2^s, \underset{\sim}{x}_3^t)$ holds in $\underset{\sim}{C}$. For compound expressions, satisfiability is defined as usual in terms of a valuation $(\underline{x}_i^r, \underset{\sim}{x}_i^r)$, $(\underline{C}^r, \underset{\sim}{C}^r)$ where for a given $\underline{x}_i^r, \underset{\sim}{x}_i^r$ is a fixed region of $\underset{\sim}{C}^r$, and for a given $\underline{C}^r, \underset{\sim}{C}^r$ is a fixed aspect of $\underset{\sim}{C}$.

DEFINITION 29.2: We call a concrete intermediate universal continuum $\underset{\sim}{C}$ a model of a formula ϕ in the formal language of Part II iff for all valuations of the variables that occur in ϕ, ϕ is satisfied by each of these valuations.

All models of formulas with continua as variables must, of course, be infinite. The density property of betweenness requires this.

30. CONTINUITY OF CONCRETE LANGUAGES VS. DISCRETENESS OF ABSTRACT ONES.

There are aspects of a drawing which cannot be fully conveyed by words -- the gaze in a portrait, say, or the figure's attitude. There are also aspects of reality that cannot be conveyed by either a drawing or by words -- the atmosphere of a situation, a scent, and the many other nonvisual qualities that lie close to the boundary of our senses. Even making allowances for all this, our usual conception of language as a tool continues to give us an exceedingly distorted and limited view of how language really functions. We must eradicate our firmly established misconception that words are merely finite strings of separated symbols ready-made to be fed to a digital computer; these rigid strings are only the words' skeleton. Karl Bühler and Jost Trier, examining the semantic aspects of language, introduced the expression "field of a word," a continuous field of meaning which overlaps other words' fields and which varies substantially according to the company that a given word keeps. Parallel, then, to the conventionally conceived syntactic object -- the word, the sentence, the paragraph -- we have, concretely speaking, a realm of continuous realities of various kinds. Structural semanticists present the following essential characteristics of a semantic field: totality, orderliness (ramified, not linear), reciprocal determination of its parts, absence of gaps, incomplete distinguishability. In the words of H. Schwarz: "The relations of concepts within a field can be of different types: one must consider subordinations, supraordinations, and coordinations, as well as, of

course, the interferences of conceptual spheres (even up to multiple superposition). Frequently, it is much less important to determine the exact external limits of a field...than to establish the centers of gravity and their reciprocal disposition."[4] A word as a syntactic object is interpreted by a continuous region of meaning; a sentence by an aspect; and a paragraph by either an aspect or an intermediate universal continuum, depending on the paragraph's context -- or lack of it. Semantic reality, then, consists of continua of various kinds, continua subject to a variety of "subordinations, supraordinations, and coordinations." It is this reality we shall now examine, keeping in mind that a discrete syntax is only an approximation of the kind of continuous syntax that concrete continuous semantics demands.

To understand that semantically words are regions in the continuum of meaning is the first step in building a grammar close to the actual use of language. Discrete written language is a crystallization of continuous regions of thought, not thought the product of discontinuous language. But the moment we become aware that the concrete sentence is not the printed one, all our semantic models necessarily become infinite; words can no longer be interpreted by isolated objects because words are an inextricable part of the sentences in which they occur, sentences which constantly modify the words' field of meaning: the sentence is routinely part of the word. This coordination of continuous semantic realities opens our mind to simple facts that otherwise would pass unrecorded; in particular, that a concept which emerges in the middle of a text incorporates the sequence of sentences that converge on it, as we shall now make clear.

31. RICKERT'S THEORY OF DEFINITION.

In his Zur Lehre von der Definition, Heinrich Rickert provides a cogent look into the concrete nature of the concept as a semantic entity. Some quotations are appropriate here to do justice to this disregarded but important work. Rickert points out that a concept depends on the thought processes that

precede it. "Ordinarily the concept is considered as a preliminary stage to thought, and a judgment as a relation between two concepts." Yet "the content of a concept... is a series of judgments. We do not realize this very clearly because we never have occasion to complete verbally such act of concept formation, expressing it in a sentence... . We can then compare the content of our knowledge with a spread of threads in which nodal fixed points are the concepts, while the threads that go from one node to another would represent the relations between concepts, that is, the judgments. If we conceive the threads in their direction toward the nodes, we have an analogy of the synthetic definition, for here the judgments meet in the concept." "The concept divides into its judgments." "In a strict sense, thought only moves...in the level of judgments, and this fact throws light on the theory of the concept."[5]

This approach has important logical and phenomenological consequences. To begin with it is concrete, for it conveys the true facts of the mind, and as Rickert says, "the concept of gravitation is identical with the law of gravitation; and laws are always judgments." To consider concepts as composed of judgments -- sentences -- is a more realistic and promising phenomenological point of departure than the usual one of seeing in the concept the incarnation of a single Platonic Idea forever identical to itself. We should not "look in a word for the 'essence' of a thing which the concept must express."[5] Words are devices to express complexes of judgments taken as aspects present in a semantic region. A concept is more than the limit of a convergent sequence of sentences in the manner of a Kantian idea; in effect, the concept has as parts the sentences that converge on it, in accordance with the principle that in a any field the whole is part of the part. Concepts are independent only to the extent to which they are composed of different streams of sentences.

Rickert's approach to the way concepts are constituted in the mind also

forces us to realize that essentialism is a crippling phenomenological error, an antiphenomenalistic error that obscures the way meaning evolves in our consciousness. Just as regions can be part of aspects and vice versa, concepts are the mingling of sentences. A concept emerges from thoughts that it preserves as parts, and then becomes the constituent of other thoughts. Terms absorb a continuous stream of sentences to become part of new sentences.

This is to some extent Tarski's approach in his paper "Methodological investigations on the definability of concepts." Here concepts are defined in relation to two collections of sentences, the first being an immediate part of the explicit definition of the concept and the second a broader collection that provides a general frame of reference within which the definition is to function. No concept can be defined logically without sentences being given prior to the definition. In this, logic reflects the concrete fact that concepts are sentences, sentences that routinely converge on a nodal point where they then change semantic direction. Hence, concepts can even be inconsistent when they embrace contradictory sentences, just as sentences can be inconsistent when they embrace contradictory paragraphs. The truth and falsity of opposing sentences is simply projected into the antinomic concept that absorbs and coordinates them, with the result that the field of meaning acquires the richest possible polarization.

32. CONTINUOUS SYNTAX.

Not only do concepts incorporate sentences and sentences paragraphs -- or even entire volumes -- but sentences can also be part of one another. If these concrete relationships are to be explicitly described by grammar, we must find a way to blend the formal symbols with the area in which they are placed, with their specific neighborhoods. This area would become the immediate syntactic context of the symbol, part of the symbol, and a place where alien meanings could occur and which could accommodate the overspill of neigh-

boring as well as distant symbols.

Symbols, then, are inevitably variable; they are a function of context and, as such, continuous linguistic realities despite their detached appearance. Every word is many words, for a symbol is not complete until the syntactic whole in which it is inserted has been incorporated by it. Specifically, the blank that appears between the letters of a word, or between the words of a sentence, is the channel in which currents of meaning -- Rickert's "threads" -- move constantly along. There is a syntactic continuum in grammar that must be emphasized if the semantic continuum is to be captured and systematically registered. In this syntactic continuum, to read a word means: I discern a complex of sentences converging toward the particular area in which the word is placed, an area in which neighboring statements dominate but where carry-overs from previous reading and thinking are freely present. Hence, discernment is a process, usually subconscious and normally submerged or even superseded by further reading and thinking, for often "the resolution of concepts into judgments cannot be continued indefinitely and..., therefore, not all judgments have subjects and predicates consisting of defined concepts, that is, of judgments."[5] Therefore, the analysis of presences in the context area of a word should distinguish between the undefined concept -- a region without any aspects -- and the defined one. In the latter case, the analysis of the convergence of aspects into the concept being discerned is like the computation of the sum of an infinite series. However, the key point with either the defined or undefined concept is that, syntactically speaking, the blank surrounding a word is not a blank but a failure to perceive the in-between. To see the blank as a blank only reveals a blank in our thoughts.

A major obstacle in developing the rules of a continuous syntax lies, once more, in our linguistic habit -- also the habit of essentialist phenomenology -- of not being able to articulate theoretically that the whole is part of the part, a factor we must consistently learn to recognize. We keep looking for essences where there is only process, expecting to identify

the Idea where we should be perceiving the sum of a sequence of minglings.

33. FINAL REMARKS.

One advantage of continua without sets is that they allow us to avoid the post-Cantor obsession with the principle of comprehension, that is, the set-theoretic compulsion to collect and seal, to gather, always gather (an "anal obsession," Freud would claim). From the viewpoint of continua without sets it is impossible to collect all regions which satisfy a given property, for something is always left out in any act of collecting. From the same viewpoint, it is easy to understand that no continuous sentence can be interpreted forever by a fixed aspect -- have a fixed meaning. Even in scientific writing concretely understood, the most one can say is that the interpretations of a sentence converge toward a clearly discernable cluster of aspects. Thus, the true model of all expression is the unfinished work that remains forever uncoagulated and flowing in open-ended continuity. This means that even the position of a word is never a matter of simple location, since each word spreads over the entire page, without regard for distance. Regrettably, an expert linguist like A. J. Greimas gives distance a dominant role in his topological studies of grammar, ignoring nonmetrical properties. It is not distance that matters when one is trying to express something, but rather unity, mingling, and betweenness. Given the ceaseless cross-references to which thinking -- articulate and inarticulate -- is constantly subject, meaning is always meaning *in statu nascendi*. And yet there *is* truth, as there is truth in saying that a portrait expresses a person despite our changing perceptions. Atomism, which hides continuity and cuts the threads that are the flesh and blood of language, is the greatest obstacle to our understanding of this fluid truth.

NOTES

1. Menger, K, "Topology without points," Rice Institute Pamphlet 27, 1, 1940, p. 107.

2. It is interesting that Frege said, in connection with the empty set, "There cannot be an empty class if we take a class to be a collection or totality of individuals." Frege, G., "A critical elucidation of some points in E. Schroeder's Vorlesungen ueber die Algebra der Logik," Translations from the Philosophical Writings of Gottlob Frege, Blackwell, 1960, p. 102. Zermelo, in turn, "refused to take as sets collections that are too 'big,' that of all 'things' or that of all ordinals, for example." Van Heijenoort, J., From Frege to Gödel, Harvard University Press, 1967, p. 200.

3. Whitehead, A. N., The Concept of Nature, Cambridge University Press, 1920, p. 59.

4. See Geckeler, H., Strukturelle Semantik und Wortfeldtheorie, W. Fink Verlag, 1971, III-B.

5. Rickert, H., Zur Lehre von der Definition, J. Mohr Verlag, 1929. All quotations from Chapter III.

Rev. Academia de Ciencias. Zaragoza 47 (1992)

EL ESPEJO DE NUESTRA HISTORIA: LOS GRANDES MATEMATICOS ARAGONESES

Concepción Romo Santos
Departamento de Algebra
Universidad Complutense
MADRID.

INTRODUCCION :

Del 5 de Octubre de 1991 al 6 de enero de 1992, los viajeros que se han detenido en Zaragoza, han podido disfrutar de una excelente exposición titulada "El espejo de nuestra historia".

La exposición antes citada es fruto de la colaboración entre el Arzobispado y el Ayuntamiento de Zaragoza y ha sido posible gracias al trabajo desarrollado por más de cien especialistas pertenecientes a diversos ámbitos científicos. Este esfuerzo colectivo confluye en un sólo objetivo: ofrecer al visitante una ambiciosa muestra de lo que ha sido la historia de la diócesis zaragozana desde sus orígenes hasta nuestros días.

La historia de la diócesis cesaraugustana ha estado unida de manera indisoluble a la historia de la ciudad. Según la tradicción, sus orígenes se remontan a tiempos apostólicos y nacen de un acontecimiento excepcional sucedido hacia el año 40 : la aparición de la Virgen al apóstol Santiago a orillas del río Ebro.

La cristiandad zaragozana, citada por primera vez en un manuscrito de San Cipriano del S. III, atravesó una profunda crisis como consecuencia de las persecuciones llevadas a cabo por los emperadores romanos. Santa Engracia y sus dieciocho compañeros, San Lorenzo y San Vicente sufrieron martirio y Valero (San Valero, obispo de Zaragoza y patrón de la ciudad) murió, según la versión más aceptada, en su exilio de Roda de

Isábena. De forma paralela se fue organizando la sede episcopal, que tras la paz de la Iglesia comenzó una lenta recuperación, alcanzando su máximo esplendor en época visigótica -y concretamente en el siglo VII- gracias a la contribución de los obispos Braulio y Tajón. Esta etapa expansionista, caracterizada por el establecimiento de los primeros contactos con la Santa Sede y por una importante actividad política e intelectual, quedó interrumpida con la ocupación musulmana. Pese a ello, la permisividad de los invasores hizo posible el mantenimiento de una mozarabía activa, únicamente anulada durante el siglo X y parte del XI.

La toma de la ciudad por las tropas de Alfonso I el Batallador en 1118 tuvo como consecuencia la restauración de la sede episcopal. Al mismo tiempo, la situación estratégica de la ciudad fue convirtiendo a ésta en cabecera diocesana y en capital del reino aragonés. Los contactos con la Santa Sede se intensificaron, y se establecieron los límites diocesanos, con los siguientes conflictos jurisdiccionales, a éstos se sumaron otros derivados de la elevación de la sede a arzobispado por Juan XXII en 1318, hecho que significó la emancipación de la metropolitana de Tarragona, los causados por las vinculaciones políticas de la iglesia zaragozana y por el enfrentamiento entre las iglesias de La Seo (Catedral de El Salvador) y Santa María la Mayor (Basílica del Pilar), que se solventó con su unión en 1675. Todos estos acontecimientos, esenciales para explicar la historia de un pueblo, y muchos otros ligados a tradiciones y devociones locales, aparecen "reflejados" en "El espejo de nuestra historia".

Como objetivo de nuestro trabajo nos proponemos el conocer un poco mejor a los grandes científicos aragoneses. Para ello haremos un estudio de los libros expuestos en "El espejo de nuestra historia".

1. LOS GRANDES MATEMATICOS ARAGONESES HASTA EL SIGLO XVIII

En el siglo XI Zaragoza será refugio de escritores y científicos musulmanes y judíos: hubo un ambiente de tolerancia que acoge a

sabios cordobeses emigrados; los mismos monarcas son entusiastas cultivadores de las ciencias. Al-Muqtadir y su hijo Almutamín fueron muy buenos astrónomos, matemáticos y filósofos. Se hizo famosa una obra del segundo, "Libro de la perfección", elogiada por un discípulo de Maimónides. El zaragozano Avenpace (1070-1138) será el primer comentarista entre los musulmanes españoles de las obras de Aristóteles, lo que hace compatible con escribir un tratado de botánica.

Destacaremos la buena situación en este siglo de la taifa aragonesa de los Banu Hud, pues en la ciudad de Huesca quizá nacieron y se formaron Moisés Sefardí y Abraham bar Hiyá, aunque su concreción cultural se realizara en territorio cristiano.

Moisés Sefardí fué médico y astrónomo. Fué médico de Alfonso I de Aragón. Escribió un opúsculo en latín para determinar los eclipses y probablemente, una traducción latina de las tablas de Aljuarismi, que algunos manuscritos le atribuyen; esas tablas fueron utilizadas por Adelardo de Bath.

Abraham ibn Ezrá fué un científico de esta época, nació en Tudela en 1809 y murió en Calahorra en 1167. Cultivó sobre todo el campo matemático, en especial el astronómico. Su obra principal son las tablas astronómicas conocidas como Tabulae pisanae, redactadas en 1145 para el merídiano de Pisa y que no se han conservado; pero él mismo las adaptó para los meridianos de Angers (1154) y de Winchester (1164), aunque solo han subsistido los cánones, de los que hay una recensión muy amplia escrita en 1154 en Dreux en latín y conocida como De rationibus tabularum. En hebreo escribió (1146) un tratado sobre el astrolabio.

En la Baja Edad Media (1250-1492) la situación difiere bastante de la Alta Edad Media. Terminadas ya las épocas de asimilación y de transmisión, precisamente entonces se produce la gran labor creadora que se manifiesta primordialmente en el campo de la astronomía. Estudiaremos los avances conseguidos en este campo en la Corona de Aragón, en la corte de Pedro IV el Ceremonioso.

Es preciso señalar la labor de los judíos fabricantes de instrumentos de cálculo (astrolabios), así como de relojes auxiliares.

Citaremos también la crítica de Hasday Cresques (Barcelona 1340-Zaragoza 1411) a la física aristotélica, que abrió nuevos horizontes científicos: es conocida su influencia en Pico della Mirandola y en Spinoza.

Y llegamos ya al siglo XVI. El zaragozano mosen Juan Andrés, escribió "Aritmética práctica", escrita para hacerse buen contador sin necesidad de maestro.

Son años abundantes en impresiones de tratados de astrología, como el "Tractatus astronomiae" de Guido Bonatti de Forlivio, exhibido.

En 1521, Gaspar Lax estando en Zaragoza escribió "Tractatus summularum magistri", exhibido. Lax fué un destacado filósofo, matemático y literato nacido en Sariñena (Huesca)en 1487. Doctor en Teología y catedrático en la Universidad de París, siendo uno de sus discípulos el famoso Luis Vives, quien así lo reconoce en su Tratado contra los dialécticos, y enseñando filosofía a San Francisco de Borja.

En Zaragoza fué regente de la Ración de la Mensa de Maestro Mayor de la catedral del Salvador (La Seo), luego tuvo dicha dignidad en propiedad y presidió su escuela hasta 1559, falleciendo el 23 de febrero de 1560. Fué enterrado en la iglesia parroquial de San Nicolás de Bari.

También en 1521 Pedro Sanchez Ciruelo escribió su Apotelesmata Astrologiae Christianae. En ella, siguiendo las teorías de Ptolomeo y los conocimientos físicos de Aristóteles, Ciruelo matiza el libre albedrío del individuo frente a una posible influencia de los astros sobre los seres.

Pedro Sanchez Ciruelo es uno de los humanistas españoles más sobresalientes. Nacido en Daroca (Zaragoza) en 1470 se formará en la Universidad de Salamanca donde aprendará astrología y matemáticas. Fué catedrático de la Universidad de Alcalá de Henares y canónigo magistral de la catedral de Salamanca. Muere en 1548 Salamanca.

En este estudio de los grandes científicos aragoneses no nos

podemos olvidar de Ramón Pignatelli (1734-1793). Pignatelli fué canónigo del cabildo catedralicio, Rector de la Universidad de Zaragoza en varias ocasiones, Protector del Canal Imperial de Aragón (1772-1793) que él llevó a cabo, fundador y director de la Real Sociedad Económica de Amigos del País (1782), además de ser el autor de la construcción de la Plaza de Toros en 1764, bautizada con el nombre de Coso de la Misericordia, pues su finalidad fué recaudar dinero para mantener la Real Casa de la Misericordia, el Hospicio de los niños cuyas mejoras apoyó el arzobispo Juan Sáenz de Buruaga (1768-1777) permitiendo trabajar a todos los que quisieran en dicha obra los días de fiesta de guardar "con tal que no lleven estipendio alguno por su trabajo",

Ramón de Pignatellí fué el gran aragonés del siglo XVIII.

En "El espejo de nuestra historia" se exponen dos libros impresos en 1796. "Elogio de Ramón de Pignatelli por la Real Sociedad Económica Arogonesa de Amigos del Pais", escrito por el conde de Sástago y el "Elogio fúnebre del señor D. Ramón Pignatelli" escrito por Juan Agustín García, los ejemplares expuestos son propiedad del palacio arzobispal de Zaragoza aunque proceden de la iglesia parroquial de San Miguel.

2.- LIBROS DE MATEMATICAS EN LA CATEDRAL DE LA SEO DE ZARAGOZA

Haremos un breve comentario de los principales libros de Matemáticas que se exhiben en la exposición : "El espejo de nuestra historia".

DECEM CONTIENENS TRACTATUS ASTRONOMIAE

GUIDO BONATTI DE FORLIVIO

1506

Libro impreso

Zaragoza. Catedral de La Seo, Biblioteca Capitular. Perteneció a D.B. Laurentis, canónigo de Santa María del Pilar.

Presenta capitulares grabados. Impreso en Venecia por Jacobum Pentiuz. Portada orlada, grabada en madera.

==========

Bonatti fué un gran astrónomo que vivió en el siglo XIII, estableció distinciones claras entre astronomía y astrología, separando esta última de otras formas de adivinación. Como astrónomo se mostró partidario de la teoría de las excéntricas para explicar los movimientos solares y lunares.

Astronomiae Tractatus decem de Guido Bonatti de Forlivio ha sido considerada tradicionalmente como el texto más importante de astrología escrito en el siglo XIII y presenta excepcional interés no sólo para la historia de la astrología sino también para la historia de la astronomía.

Según reflejó en el prólogo, escribió su tratado que definió como largo y prolijo, para uso de un sobrino y con intención de recopilar las teorías formuladas por autores anteriores y exponerlas de forma que pudieran ser comprendidas por quienes carecieran de amplios conocimientos científicos. Utilizó en consecuencia frecuentes alusiones y referencias a numerosos astrónomos y astrólogos de la antiguedad, evitando en lo posible la inclusión de "disputas sin pruebas" que hubieran alargado considerablemente la extensión del texto.

Dante hizo aparecer a Bonatti como inquilino del infierno pero esta circunstancia no impidió el éxito alcanzado por el Tractatus astronomiae. Abundantes copias manuscritas circularon por Europa durante la etapa anterior a la aparición de la imprenta, una de las cuales, extraordinariamente lujosa, perteneció a Enrique VII de Inglaterra. Los reyes de España también dispusieron de un códice, copiado en el siglo XIV y conservado en la Biblioteca del Escorial.

La presente edición veneciana, impresa a expensas de Melchor Sesse bajo el reinado del principe Leonardo Lauretano, destaca por su abundante y bella iconografía.

TRACTATUS SUMMULARUM MAGISTRI

GASPARIS LAX ARAGONENSIS

Gaspar Lax

1521

31 x 20 cm

Libro impreso

Zaragoza. Seo. Biblíoteca Capitular. Impreso en Zaragoza por Jorge Cocci

Portada con armas de Arzobispo Don Juan de Aragón Grabado de la Crucifixión.

==========

La obra está dedicada por su autor, Gaspar Lax, al arzobispo de Zaragoza Don Juan de Aragón, y parece ser un libro usado por escolares que escuchaban al propio autor en 1527.

Lax fué un destacado filósofo, matemático y literato nacido en Sariñena (Huesca) en 1487. Sus obras escritas e impresas hacen un número de diecinueve, y entre ellas están el Tractatus Summularum y el Tractatus Parvorum Logicalium, obra esta segunda dedicada asimismo al arzobispo Don Juan de Aragón e impresa por Jorge Cocci también en 1521.

DE RE AEDIFICATORIA, LIBRI DECEM LEON BAUTISTA ALBERTI

1541

20,5 x 16 cm

Tinta sobre papel

Zaragoza, Catedral de La Seo, Bibliotecaa del Cabildo

Impreso en Basilea (Suiza)

En el estilo de los Diez Libros de Arquitectura de Marco Vitrubio Polion, el arquitecto del emperador Augusto, descubierto por aquellos años, y de los que circulan numerosas copias manuscritas por la Roma humanista de Nicolás V, a la espera de su impresión, que no tiene lugar hasta 1483 por Giovanni Sulpicio da Verole; y con la pretensión de actulizarlos y ampliarlos, escribe León Battista Alberti su "De Re Aedificatoria Decem Libris".

Alberti nace en Génova en 1403, y fallece en Roma en 1472, apareciendo la edición principe del Re Aedificatoria en 1485, es pues, la obra de una vida, donde Alberti vierte todas sus experiencias, no solo del campo de la arquitectura sino de la filosofía y poesía.

La obra comprende una sistematización de los edificios clásicos , extraida de la observación directa de las ruinas, dibujadas y medidas por el propio autor, tal como también hacían en Roma Brunelleschi, Donatello y Massacio.

Por su formación humanista, Alberti es polifacetico; conoce la pintura, sobre la que escribe también un tratado contemporáneo al de Leonardo da Vinci, la escultura y el grabado, aunque considera a la arquitectura "la joya en que se insertan las demás piedras preciosas, que acrecientan su valor y belleza".

Alberti es fundamentalmente teórico, considera que la obra de arquitectura perfecta es la que se desarrolla solamente en el proyecto, ya que el proceso constructivo, con sus limitaciones, la empobrece. Así, la mayor parte de su obra arquitectónica se realiza sin su dirección, sobre los planos que redacta. Excepciones son el Templo Malatestiano, reforma de un antiguo convento de franciscanos para su conversión en panteón de Segismondo Pandolfo Malatesta, Señor de Rímini y su amada Isotta, y las iglesias de San Sebastián y San Andrés de Mantua, en las que materializa el modelo de iglesia descrito en su Re Aedificatoria, rematado en fachada por un arco triunfal a lo romano.

Rev. Academia de Ciencias. Zaragoza 47 (1992)

REPRÈSENTACIONES INTEGRALES DE LAS FUNCIONES DE BESSEL-CLIFFORD DE TERCER ORDEN

N. Hayek Calil
Departamento de Análisis Matemático
Facultad de Matemáticas
Universidad de La Laguna
38271 - La Laguna (Tenerife)

V.Hernández Suárez
Departamento de Matemáticas
Universidad de Las Palmas de Gran Canaria

e) Abstract:

In this paper, we establish some integral representations for the so-called Bessel-Clifford functions of the third order [6]. These representations involve the third order sinus f_1, f_2 and f_3, the Appell's functions P, Q, and R, and certain functions cer x and cei x (of similar structure to the well-known ber x and bei x, due to Kelvin), as well as some integral formulas of the Sonine and Weber types.

1. INTRODUCCION

Las funciones $C_{m,n}(x)$ de Bessel-Clifford de tercer orden fueron introducidas por N. Hayek en [6]. Dichas funciones representan una clase de naturaleza análoga a las de Bessel de igual orden, inicialmente estudiadas por P. Humbert [8], quién las definió mediante la fórmula:

$$J_{m,n}(x) = \frac{x^{m+n}}{3^{m+n}\,\Gamma(m+1)\,\Gamma(n+1)}\;{}_0F_2\left(m+1,\ n+1;\ \frac{-x^3}{27}\right) \qquad (1.1)$$

donde ${}_0F_2$ designa la función hipergeométrica triconfluente de tercer orden, fórmula que constituye una obvia generalización de la conocida representación [16]:

$$J_m(x) = \frac{x^m}{2^m\,\Gamma(m+1)}\;{}_0F_1\left(m+1;\ \frac{-x^2}{4}\right)$$

Las $J_{m,n}(x)$ serían posteriormente investigadas en varios trabajos por el propio Humbert ([9], [10], [11], [12]), así como por otros autores, entre ellos R.S. Varma [15], N.W. MacLachlan [14], P. Agarwall [1], P. Delerue

[3], y más recientemente han sido objeto de atención por parte de otros varios, como H. Dimovski [4], V. Kiryakova [13], y algunos más.

Entre las propiedades establecidas para las funciones $C_{m,n}(x)$ (véase [6]), destaca principalmente la relativa a su conexión con las $J_{m,n}(x)$:

$$C_{m,n}(x) = x^{-\frac{m+n}{3}} J_{m,n}\left(3\sqrt[3]{x}\right), \tag{1.2}$$

la expresión de su función generatriz:

$$e^{u+v-\frac{x}{uv}} = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} u^m v^n C_{m,n}(x), \tag{1.3}$$

el desarrollo (*):

$$C_{m,n}(x) = \sum_{r=0}^{\infty} \frac{(-x)^r}{\Gamma(m+r+1)\,\Gamma(n+r+1)\,r!} = \frac{1}{\Gamma(m+1)\,\Gamma(n+1)}\, {}_0F_2(m+1,\ n+1;\ -x) \tag{1.4}$$

determinadas fórmulas de recurrencia y, sobre todo, el hecho de ser natural generalización de un tipo de funciones $C_m(x)$, denominadas de Bessel-Clifford de primera especie, estudiadas en profundidad en varios trabajos por Hayek (entre los que cabe destacar [5]), y cuyo uso en múltiples campos teóricos y de aplicación, sustituye con manifiesta ventaja a las $J_m(x)$ de Bessel. En un trabajo nuestro anterior [7] fueron obtenidas otras interesantes propiedades de las funciones de Bessel-Clifford de tercer orden; entre ellas, relaciones de las mismas con los senos de tercer orden f_1, f_2 y f_3, investigados por Appell [2], desarrollos de éstos en términos de las $C_{m,n}(x)$, la ecuación diferencial de tercer orden que las mismas satisfacen, así como algunas otras ecuaciones relevantes de la Física Matemática cuyas soluciones son expresables en términos de ellas.

En el presente trabajo se investigan nuevas propiedades de estas funciones, ofreciéndose especialmente diversas representaciones integrales de las $C_{m,n}(x)$.

(*) La adopción de una nueva forma para la función generatriz (1.3), permite obtener los valores de las $C_{m,n}(x)$ en el caso de índices m y n enteros negativos, para el cual no tiene evidentemente sentido el desarrollo (1.4) que las define, (véase [6]).

Son deducidas expresiones en las que intervienen las funciones P, Q y R de Appell [2], representaciones integrales en función de los senos de tercer orden f_1, f_2 y f_3 y otras que contienen las funciones cer x y cei x (de estructura similar a las ber x y bei x, de conocida aplicación en varios contextos físicos). Se incluyen, asimismo, representaciones para las $C_{m,n}(x)$ de los tipos de integrales de Sonine y de Weber para las funciones de Bessel.

2. REPRESENTACIONES INTEGRALES DE LAS $C_{m,n}(x)$

2.1 Expresiones que incluyen las funciones P, Q y R de Appell

Haciendo uso de (1.2) con m = 0, n = 0, o bien siguiendo un proceso directo similar al desarrollado por Humbert [8] para expresar las $J_{m,n}(x)$ mediante una integral doble, se infiere la fórmula:

$$C_{0,0}\left(\frac{x^3}{27}\right) = \frac{i\sqrt{3}}{4\pi^2}\iint_D f_1\left[x\,P(\theta,\varphi)\right]\,d\theta\,d\varphi \qquad (2.1)$$

donde los puntos representativos de θ y φ se encuentran situados en un rectángulo de lados $\frac{2\pi}{\sqrt{3}}$ y $2\pi i$ (que son precisamente los períodos de la función $P(\theta,\varphi)$ de Appell).

Nota. La (2.1) generaliza la llamada integral tipo Parseval para la función de Bessel-Clifford ordinaria [5]:

$$C_0\left(\frac{x^2}{4}\right) = \frac{1}{\pi}\int_0^{\pi}\cos\,(x\,\text{sen}\,\varphi)\,d\varphi$$

El análisis precedente puede extenderse asimismo a la función $C_{m,n}(x)$ de índices no nulos, desprendiéndose la representación de la misma por una de las integrales siguientes (y con igual dominio de integración que para la $C_{0,0}(x)$):

$$C_{m,n}\left(\frac{x^3}{27}\right) = (-1)^{m+n}\,\frac{i\sqrt{3}}{4\pi^2}\left(\frac{x}{3}\right)^{-(m+n)}\iint e^{-xP(\theta,\varphi)-m(j\theta+j^2\varphi)-n(j^2\theta+j\varphi)}\,d\theta\,d\varphi \qquad (2.2)$$

$$C_{m,n}\left(\frac{x^3}{27}\right) = (-1)^{m+n}\,\frac{i\sqrt{3}}{4\pi^2}\left(\frac{x}{3}\right)^{-(m+n)}\,j^{n-m}\iint e^{-xQ(\theta,\varphi)-m(j\theta+j^2\varphi)-n(j^2\theta+j\varphi)}\,d\theta\,d\varphi \qquad (2.3)$$

$$C_{m,n}\left(\frac{x^3}{27}\right) = (-1)^{m+n}\ \frac{i\sqrt{3}}{4\pi^2}\left(\frac{x}{3}\right)^{-(m+n)} j^{m-n}\iint e^{-xR(\theta,\varphi)-m(j\theta+j^2\varphi)-n(j^2\theta+j\varphi)}\,d\theta\,d\varphi \qquad (2.4)$$

siendo P, Q y R las funciones de Appell y $j^3 = 1$.

2.2. Expresiones en que intervienen los senos de tercer orden f_1, f_2 y f_3.

Al igual que para las $J_{m,n}(x)$, el cálculo simbólico u operacional de Heaviside [14] puede ser aplicado a las funciones de Bessel-Clifford de tercer orden, para derivar nuevas propiedades de éstas a través de sus imágenes.

Así, si se parte del desarrollo (1.4), se deduce sin dificultad:

$$x^m y^n C_{m,n}(xy) \supset \frac{1}{p^m q^n}\ e^{\frac{-1}{pq}}$$

Descomponiendo adecuadamente el producto del segundo miembro y aplicando el teorema de composición, sigue:

$$x^m y^n C_{m,n}(x\,y) = \int_0^x\int_0^y \frac{(x-u)^{m-m'}(y-v)^{n-n'}}{\Gamma(m-m'+1)\Gamma(n-n'+1)}\,u^{m'-1}v^{n'-1}\,C_{m'-1,n'-1}(uv)\,du\,dv \qquad (2.5)$$

Si se sustituye ahora m' por $\frac{2}{3}$ y n' por $\frac{1}{3}$, y se tiene en cuenta que [7]:

$$C_{\frac{-1}{3},\frac{2}{3}}(x) = \frac{\sqrt{3}}{2\pi}\,f_1(3\sqrt[3]{x}),$$

se infiere, tras oportunos cambios de variable:

$$C_{m,n}(z) = \frac{9\sqrt{3}}{2\pi\Gamma(m+\frac{1}{3})\Gamma(n+\frac{2}{3})}\int_0^1\int_0^1 (1-\xi^3)^{m-\frac{2}{3}}\,(1-\eta^3)^{n-\frac{1}{3}}\,\xi\,f_1(3\xi\eta\sqrt[3]{z})\,d\xi\,d\eta \qquad (2.6)$$

válida para $m > -\frac{1}{3}$, $n > -\frac{2}{3}$.

Análogamente, partiendo de relaciones similarmente deducidas de la (2.5) resulta, al poner $m' = \frac{4}{3}$, $n' = \frac{5}{3}$ y $m' = \frac{2}{3}$, $n' = \frac{4}{3}$, respectivamente, y usar las [7]:

$$C_{\frac{1}{3},\frac{2}{3}}(x) = \frac{\sqrt{3}}{2\,\pi\,\sqrt[3]{x^2}}\, f_2\left(3\sqrt[3]{x}\right)$$

$$C_{\frac{-1}{3},\frac{1}{3}}(x) = \frac{\sqrt{3}}{2\,\pi\,\sqrt[3]{x}}\, f_3\left(3\sqrt[3]{x}\right),$$

las siguientes representaciones integrales:

$$C_{m,n}(z) = \frac{9\sqrt{3}\; z^{\frac{-2}{3}}}{2\pi\Gamma(m-\frac{1}{3})\Gamma(n-\frac{2}{3})} \int_0^1\int_0^1 (1-\xi^3)^{m-\frac{4}{3}} (1-\eta^3)^{n-\frac{5}{3}}\, \xi\,\eta^2\, f_2(3\xi\eta\sqrt[3]{z})\; d\xi\; d\eta \qquad (m > \tfrac{1}{3},\ n > \tfrac{2}{3}) \qquad (2.7)$$

$$C_{m,n}(z) = \frac{9\sqrt{3}\; z^{\frac{-1}{3}}}{2\pi\Gamma(m+\frac{1}{3})\Gamma(n-\frac{1}{3})} \int_0^1\int_0^1 (1-\xi^3)^{m-\frac{2}{3}} (1-\eta^3)^{n-\frac{4}{3}}\, \eta^2\, f_3(3\xi\eta\sqrt[3]{z})\; d\xi\; d\eta \qquad (m > \tfrac{-1}{3},\ n > \tfrac{1}{3}) \qquad (2.8)$$

Nota. Las (2.6), (2.7) y (2.8) constituyen sendas generalizaciones de la representación integral de la función $C_n(z)$ de Bessel-Clifford mediante una integral del tipo de Poisson, establecida por Hayek [5]:

$$C_n(z) = \frac{2}{\Gamma(n+\frac{1}{2})\;\Gamma(\frac{1}{2})} \int_0^1 \left(1-\xi^2\right)^{n-\frac{1}{2}} \cos(2\sqrt{z}\;\xi)\; d\xi,$$

con intervención de los senos de tercer orden f_1, f_2 y f_3, respectivamente, en lugar del coseno ordinario.

Las (2.6), (2.7) y (2.8), pueden considerarse, por otra parte, como extensiones a las funciones de Bessel-Clifford de tercer orden de las fórmulas integrales para las $J_{m,n}(x)$ obtenidas por Delerue [3] y de contextura análoga a éstas.

De forma similar, cabe establecer expresiones de las $C_{m,n}(z)$ en función de los denominados senos hiperbólicos h_1, h_2 y h_3 de orden superior.

Por último, y en particular, dando oportunos valores a los subíndices m y n, se deducen diversas representaciones integrales; por ejemplo:

$$C_{\frac{2}{3},\frac{1}{3}}(z) = \frac{9\sqrt{3}}{2\,\pi} \int_0^1\int_0^1 \xi\, f_1\left(3\,\xi\,\eta\,\sqrt[3]{z}\right)\; d\xi\; d\eta$$

$$C_{\frac{2}{3},\frac{4}{3}}(z) = \frac{9\sqrt{3}}{2\pi}\int_0^1\int_0^1 \xi\,(1-\eta^3)\,f_1\left(3\,\xi\,\eta\,\sqrt[3]{z}\right)\,d\xi\,d\eta$$

$$C_{\frac{4}{3},\frac{5}{3}}(z) = \frac{9\sqrt{3}}{2\pi}\,z^{\frac{-2}{3}}\int_0^1\int_0^1 \xi\,\eta^2\,f_2\left(3\,\xi\,\eta\,\sqrt[3]{z}\right)\,d\xi\,d\eta$$

$$C_{\frac{5}{3},\frac{7}{3}}(z) = \frac{9\sqrt{3}}{2\pi}\,z^{\frac{-1}{3}}\int_0^1\int_0^1 (1-\xi^3)\,(1-\eta^3)\,\eta^2\,f_3\left(3\,\xi\,\eta\,\sqrt[3]{z}\right)\,d\xi\,d\eta,$$

entre otras. Y de éstas, con el uso de fórmulas de recurrencia, resultan relaciones que ligan entre sí a los senos de tercer orden, por ejemplo:

$$f_3(z) - z\,f_1(z) = \frac{z^4}{3}\int_0^1\int_0^1 \xi\,(1-\eta^3)\,f_1(\xi\,\eta\,z)\,d\xi\,d\eta$$

$$2f_2(z)-zf_3(z) = \frac{z^5}{9}\int_0^1\int_0^1 (1-\xi^3)\,(1-\eta^3)\,\xi\,f_1(3\,\xi\,\eta\,z)\,d\xi\,d\eta$$

2.3. Expresiones que contienen las funciones cer x y cei x

Las funciones cer x y cei x fueron definidas por Hayek [5], como parte real e imaginaria, respectivamente, de la función $C_0(-ix)$ (x real) (*):

$$C_0(-ix) = \text{cer } x + i\ \text{cei } x$$

Sus desarrollos en serie:

$$\text{cer } x = 1 + \sum_{k=1}^{\infty}\frac{(-1)^k\,x^{2k}}{[(2k)!]^2}$$

$$\text{cei } x = -\sum_{k=1}^{\infty}\frac{(-1)^k\,x^{2k-1}}{[(2k-1)!]^2}$$

son absoluta y uniformemente convergentes en todo compacto de $\mathbb{R}$, pudiendo por ello ser diferenciadas e integradas término a término.

(*) Las cer x y cei x están conexionadas con las conocidas funciones ber x y bei x, introducidas por Kelvin a partir de la expresión: $J_0(x\ i\ \sqrt{i}) = \text{ber } x + i\ \text{bei } x$, así $\text{ber}(2\sqrt{x}) = \text{cer } x$, $\text{bei}(2\sqrt{x}) = \text{cei } x$.

Al igual que en el apartado 2.2 anterior, si se recurre a las reglas del cálculo operacional, cabe inferir, en primer lugar, que:

$$\cos \frac{1}{p} \subset \operatorname{cer} x, \qquad \operatorname{sen} \frac{1}{p} \subset \operatorname{cei} x \tag{2.9}$$

Entonces, con el uso de la primera de (2.9) y de la conocida expresión simbólica:

$$f(\sqrt{p}) \subset \frac{1}{\sqrt{\pi x}} \int_0^{\infty} e^{-\frac{s^2}{4x}} h(s)\, ds \qquad (\text{con } f(p) \subset h(x))$$

resulta una nueva expresión, en la que al cambiar p por $\frac{p}{4 \operatorname{sen}^2\theta}$, se deduce:

$$\cos\left(\frac{2 \operatorname{sen}\theta}{\sqrt{p}}\right) \subset \frac{1}{2 \operatorname{sen}\theta \sqrt{\pi x}} \int_0^{\infty} e^{-\frac{s^2}{16 x \operatorname{sen}^2\theta}} \operatorname{cer}(s)\, ds$$

Ahora bien, si se tiene en cuenta la siguiente integral tipo Parseval [5]:

$$C_0\left(\frac{1}{p}\right) = \frac{1}{\pi} \int_0^{\pi} \cos\left(\frac{2 \operatorname{sen}\theta}{\sqrt{p}}\right) d\theta,$$

la integración respecto de θ de la última expresión simbólica conduce a:

$$C_{0,0}(x) = \frac{1}{2\pi\sqrt{\pi x}} \iint_C \frac{1}{\operatorname{sen}\theta} e^{-\frac{s^2}{16 x \operatorname{sen}^2\theta}} \operatorname{cer}(s)\, ds\, d\theta \tag{2.10}$$

viniendo dado el campo C de integración por $0 < \theta < \pi$, $0 < s < \infty$.

Partiendo de la segunda de (2.9) y siguiendo un proceso similar al precedente, se infiere esta otra representación integral:

$$C_{\frac{1}{2},\,1}(x) = \frac{1}{2\pi\sqrt{\pi x}} \iint_C e^{-\frac{s^2}{16 x \operatorname{sen}^2\theta}} \operatorname{cei}(s)\, ds\, d\theta \tag{2.11}$$

con igual campo C de integración.

2.4. Otras representaciones

La siguiente representación para las de Bessel-Clifford de tercer orden:

$$C_{\lambda_1+\mu_1+1,\lambda_2+\mu_2+1}(x) = \frac{2^2}{\Gamma(\mu_1+1)\,\Gamma(\mu_2+1)} \int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} \prod_{i=1}^{2} \operatorname{sen}^{2\lambda_1+1}\theta_1 \cos^{2\mu_1+1}\theta_1 \cdot$$

$$\cdot\, C_{\lambda_1,\lambda_2}(x \operatorname{sen}^2\theta_1 \operatorname{sen}^2\theta_2)\, d\theta_1\, \theta_2 \qquad (2.12)$$

en donde la integral doble contiene una función del mismo orden, pero de índices λ_1 y λ_2 inferiores que los de la función representada, generaliza la siguiente fórmula integral tipo Sonine para las funciones de Bessel-Clifford de 1ª especie [5]:

$$C_{\mu+\nu+1}(x) = \frac{2}{\Gamma(\nu+1)} \int_0^{\frac{\pi}{2}} C_\mu(x\operatorname{sen}^2\theta)\, \operatorname{sen}^{2\mu+1}\theta\, \cos^{2\nu+1}\theta\, d\theta$$

La (2.12) se establece por proceso directo, sustituyendo en la integral:

$$C_{\lambda_1,\lambda_2}(x \operatorname{sen}^2\theta_1 \operatorname{sen}^2\theta_2) \quad \text{por} \quad \sum_{k=0}^{\infty} \frac{(-1)^k x^k \operatorname{sen}^{2k}\theta_1 \operatorname{sen}^{2k}\theta_2}{\Gamma(\lambda_1+k+1)\,\Gamma(\lambda_2+k+1)},$$

con lo cual, tras el cálculo de algunas integrales conocidas, el segundo miembro de (2.12) se transforma en el sumatorio:

$$\sum_{k=0}^{\infty} \frac{(-1)^k x^k}{\Gamma(k+\lambda_1+\mu_1+2)\,\Gamma(k+\lambda_2+\mu_2+2)}$$

que, según (1.4), no es otra cosa que el desarrollo del primer miembro.

Análogamente, y también por comprobación directa, resulta:

$$\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} C_{m,n}\left(x\operatorname{sen}^2\theta_1\operatorname{sen}^2\theta_2\right) C_{m',n'}\left(z\cos^2\theta_1\cos^2\theta_2\right) \cdot$$

$$\cdot \operatorname{sen}^{2m+1}\theta_1 \operatorname{sen}^{2n+1}\theta_2 \cos^{2m'+1}\theta_1 \cos^{2n'+1}\theta_2 d\theta_1 d\theta_2 =$$

$$= \frac{1}{2^2} C_{m+m'+1,n+n'+1}(x+z) \qquad (2.13)$$

fórmula que generaliza la segunda integral tipo Sonine para las de Bessel-Clifford de primera especie [5].

Por último, si se aplica la propiedad simbólica:

$$x^{\lambda_1} C_{\lambda_1,\lambda_2}(x) \supset \frac{1}{p^{\lambda_1}} C_{\lambda_2}\left(\frac{1}{p}\right),$$

se infiere, en virtud de la definición de la transformada generalizada de Laplace-Carson:

$$\int_0^\infty e^{-\frac{x}{a}} x^{\lambda_1} C_{\lambda_1,\lambda_2}(x)\, dx = a^{\lambda_1+1} C_{\lambda_2}(a), \qquad (2.14)$$

integral exponencial del tipo de Weber, que relaciona las funciones de tercer y primer orden.

3. BIBLIOGRAFIA

[1] AGARWAL, R.P. (1950):"Sur une généralisation de la transformation de Hankel". *Ann. Soc. Sci. Bruxelles. Ser. 1, 64, 164-168.*

[2] APPELL, P. (1897): "Sur certaines fonctions analogues aux fonctions circulaires", *C.R. Acad. Sc.*, t. 84, 540.

[3] DELERUE, P. (1953): "Sur le Calcul Symbolique à n variables et les fonctions hyperbesséliennes, II". *Ann. Soc.Sci. Bruxelles, 67, 3, 229-274.*

[4] DIMOVSKI, I.H.; KIRYAKOVA, V.S. (1987): "Generalized Poisson representations of hypergeometric functions ${}_pF_q, p<q$ using fractional integrals". *Proceedings of the Sixteenth Spring Conference of the Union of Bulgarian Mathematicians,* 205-212, Sofía.

[5] HAYEK, N.(1967): "Estudio de la ecuación diferencial $xy''+(\nu+1)\,y'+y=0$ y de sus aplicaciones". *Collectanea Mathematica,* Vol. XVIII, fasc. 1 y 2, 57-174, Barcelona.

[6] HAYEK, N. (1987): "Funciones de Bessel-Clifford de tercer orden". *Actas XII Jornadas Luso-Españolas de Matemática,* 346-351, Braga.

[7] HAYEK, N.; HERNANDEZ, V.(1991):"Sobre las funciones de Bessel-Clifford de tercer orden". *Actas XII C.E.D.Y.A. (II Congreso de Matemática Aplicada),* Oviedo.

[8] HUMBERT, P. (1930): "Les fonctions de Bessel du troisième ordre". *Atti. Pont. Accad. Sc. Nuovi Lincei, Anno 83, fasc. III*, 128-146.

[9] HUMBERT, P. (1934): "Nouvelles remarques sur les fonctions de Bessel du troisième ordre". *Atti. Pont. Accad. Sc. Nuovi Lincei,* 87, 323-331.

[10] HUMBERT,P. (1935): "Sur une équation aux dérivées partielles". *Atti.*

Pont. Accad. Sc. Nuovi Lincei, Tome 88.

[11] HUMBERT, P.; COLOMBO, S. (1965): "Le Calcul Symbolique et ses applications a la Physique Mathématique". *Mémorial des Sc. Math. Gauthier-Villars,* fasc. 158, Paris.

[12] HUMBERT,P.; DELERUE,P. (1950): "Sur l'equation différentielle de la fonction de Bessel du troisième ordre, d'indices nuls". *Ann. Soc. Sci. Bruxelles,* I, 64, 159-163.

[13] KIRYAKOVA, V.; SPIROVA, S. (1989): "Representations of the solutions of Hyper-Bessel differential equations via Meijer's G-function". *Complex Analysis and Applications*, 87, 284-297, Sofía.

[14] MCLACHLAN, N.W.; HUMBERT, P. (1941): "Formulaire pour le Calcul Symbolique", *Mémorial des Sc. Math. Gauthier-Villars,* Paris.

[15] VARMA, R.S. (1939): "Sur les fonctions de Bessel du troisième ordre", *J. Ecole Polytechn,* III, 145, 33-35.

[16] WATSON, G.N. (1958): "A treatise on the theory of Bessel Functions". *Cambridge University Press,* second edition.

Rev. Academia de Ciencias. Zaragoza 47 *(1992)*

BOUNDS FOR THE ZEROS OF POLYNOMIALS

Ioannis K. Argyros
Department of Mathematics
Cameron University
Lawton, OK - 73505

Abstract. A series of new bounds for the norm of the solution of the polynomial equation is obtained.

A.M.S. Classification codes: 12B22, 46B15.

Key Words: Upper-lower bound, polynomial, Cauchy.

Introduction. Consider the polynomial

$$P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 \tag{1}$$

where $a_0, a_1, \cdots, a_{n-1}$ are complex numbers. For every zero z of $P(z)$ we have

$$|z| \leq \max\{|a_0|, 1+|a_1|, \cdots, 1+|a_{n-1}|\}, \tag{2}$$

$$|z| \leq \max\{1, |a_0| + |a_1| + \cdots + |a_{n-1}|\}, \tag{3}$$

$$|z| \leq r, \tag{4}$$

where r is the unique positive zero of

$$G(z) = z^n - |a_{n-1}|z^{n-1} - \cdots - |a_1|z - |a_0|. \tag{5}$$

These are classical results [2],[3]. It is also known [1] that

$$|z| \leq L_m \tag{6}$$

where

$$L_m = \max \{S_m, 1+|a_{m+1}|, 1+|a_{m+2}|, \cdots, 1+|a_{n-1}|\}, \quad m = 0,1,2,\ldots,n-1$$

and S_m is the unique positive zero of

$$G_m(z) = z^{m+1} - |a_m|z^m - |a_{m-1}|z^{m-1} - \cdots - |a_1|z - |a_0|. \tag{7}$$

Finally we have

$$|z| \leq \max \{1, |a_0| + \cdots + |a_m|, 1+|a_{m+1}|, \cdots, 1+|a_{n-1}|\} \tag{8}$$

where $m \in \{0,1,2,\cdots,n-1\}$.

Note that for $m = 0$ and $m = n-1$, (6) yields (2) and (4), whereas for $m = 0$ and $m = n-1$ (8) yields (2) and (3).

In this paper we improve further the above upper bounds. We also provide a lower bound for the absolute value of the solution z.

Finally we show how to use the above bounds to bound the solutions of polynomial equations in a Normed space X.

Proposition 1. Let $k \in \{1,2,\cdots,n-1\}$ be fixed. Assume:

$$0 \neq |a_k| < \bar{N}_k \tag{9}$$

where

$$\bar{N}_k = \begin{cases} 1 + N_k, & k = 1,2,\ldots,n-2 \\ r_{n-1}, & k = n-1 \end{cases}$$

$$N_i = \max \{|a_{i+1}|, |a_{i+2}|, \cdots, |a_{n-1}|\}, \quad i = 1,2,\cdots,n-2,$$

r_k is the unique positive zero of

$$P_k(z) = (q_k-1)|a_k|z^k - |a_{k-1}|z^{k-1} - \cdots - |a_1|z - |a_0|$$

and q_k is such that

$$1 < q_k \leq \frac{\bar{N}_k}{|a_k|}, \quad k = 1,2,\cdots,n-1. \tag{10}$$

Then every zero of the complex polynomial

$$P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

satisfies

$$|z| \leq M_k \tag{11}$$

where

$$M_k = \begin{cases} \max\{r_k, 1+|a_{k+1}|, 1+|a_{k+2}|, \cdots, 1+|a_{n-1}|\}, & k = 1,2,\cdots,n-2 \\ r_{n-1}, & k = n-1 \end{cases}$$

Proof. We shall show that a complex number z satisfying $|z| > \bar{N}_k$, $|z| > r_k$, $k = 1,2,\cdots,n-1$ cannot be a zero of P.

We have

$$\begin{aligned} |P(z)| &= |z^n + \sum_{j=0}^{n-1} a_j z^j| \geq |z|^n - \sum_{j=0}^{n-1} |a_j||z|^j \\ &= |z|^n - q_k|a_k|\cdot|z|^k + q_k|a_k||z|^k \\ &\qquad - \sum_{j=0}^{k} |a_j|\cdot|z|^j - \sum_{j=k+1}^{n-1} |a_j||z|^j \\ &\geq (|z|^n - q_k|a_k||z|^k) + P_k(|z|) - N_k \sum_{j=k+1}^{n-1} |z|^j \\ &\geq (|z|^n - |z|^{k+1}) + P_k(|z|) - N_k \sum_{j=k+1}^{n-1} |z|^j \end{aligned}$$

(by the choice of q_k)

$$= [|z|^n - |z|^{k+1}][|z| - \bar{N}_k]/(|z|-1) + P_k(|z|).$$

Since $|z| > r_k$ and r_k is the unique positive zero of P_k, we have $P_k(|z|) > 0$ and thus $|P(z)| > 0$.

Remark 1. Proposition 1 can be sometimes applied even if (9) is violated. Consider the real equation

$$z^3 + z^2 - 4z + 1 = 0. \tag{12}$$

Here, $a_0 = a_2 = 1$, $a_1 = -4$ and for $k = 1$, $N_1 = |a_2|$ so (9) is violated. Set $z = x + 1$ then (12) is equivalent to

$$\begin{aligned} z &= x + 1 \\ x^3 + 4x^2 + x - 1 &= 0 \end{aligned} \tag{13}$$

and (9) is now satisfied in (13).

We now compare L_k and M_{k+1}.

Proposition 2. Assume:

$$\begin{aligned} &\text{(a)} \quad i \in \{1,2,\cdots,n-2\} \text{ is fixed;} \\ &\text{(b)} \quad |a_{i+1}| < \bar{N}_{i+1} \end{aligned} \tag{14}$$

and

$$\text{(c)} \quad \left.\begin{aligned} \frac{1+|a_{i+1}|}{|a_{i+1}|} < q_{i+1} &\le \frac{\bar{N}_{i+1}}{|a_{i+1}|} \\ 1 < q_{n-1} &= \frac{r_{n-1}}{|a_{n-1}|} \end{aligned}\right\}. \tag{15}$$

Then

$$M_{m+1} \le L_m \ , \ m = 0,1,2,\cdots,n-2$$

and

$$r_{n-1} = r.$$

Proof. We have

$$G_k(z) = z^{k+1} - |a_k|z^k - |a_{k-1}|z^{k-1} - \cdots - |a_0|$$
$$P_{k+1}(z) = (q_{k+1}-1)|a_{k+1}|z^{k+1} - |a_k|z^k - |a_{k-1}|z^{k-1} - \cdots - |a_0|.$$

Set $z = S_k$ then $G_k(S_k) = 0$ and

$$P_{k+1}(S_k) = [(q_{k+1}-1)|a_{k+1}| - 1]S_k^{k+1} > 0$$

by the choice of q_{k+1} $\qquad r_{m+1} < S_m$

$$\Rightarrow M_{m+1} \le L_m, \quad m = 0,1,2,\cdots,n-2.$$

Also $G(r_{n-1}) = 0$ by (14) and (15), so $r_{n-1} = r$.

Note that a comparison between the bounds given in (6) and (8) makes sense only if the degree of the polynomial G_m is the same with the degree of P_k.

Remark 2.

(a) Consider the polynomial

$$P(z) = z^3 + 5z^2 + z + 7.$$

Let $m = 0$ then the bound given by (2), (6) and (8) is 7. The bound given by (3) is 13 whereas the bound given by (11) for $k = 1$, $q_1 = 6$ is

$$M_1 = \max\{1.4,6\} = 6.$$

(b) Consider the polynomial

$$P(z) = z^2 + z + 1.$$

The bound on $|Z|$ given by (4) is

$r = 1.6180339$

whereas for $k = 1$, $q_1 = 1.6180339$

$$M_1 = r_1 = r.$$

The estimates obtained above can also be used to find upper and lower bounds for the norm of the solutions of polynomial equations in a normed space X.

Let

$$B(x) = B_n x^n + B_{n-1} x^{n-1} + \cdots + B_1 x + B_0 \qquad (28)$$

where the B_ℓ, $\ell = 1,2,\cdots,n$ is a bounded ℓ-linear operator on X [4] and $B_0 \in X$ is fixed.

Set

$$C(x) = \|B_n\| \cdot \|x\|^n + \|B_{n-1}\| \|x\|^{n-1} + \cdots + \|B_1\| \cdot \|x\| + \|B_0\|$$

and

$$D(x) = \frac{1}{\|B_n\|} C(x) \quad \text{if} \quad \|B_n\| \neq 0.$$

We can now apply Propositions 1 and 5 on $D(x)$ to obtain bounds on the norm of the solution x of (28).

References

[1] Deutsch, E. Bounds for the zeros of polynomials. Amer. Math. Month. 88,3(1981), pp. 205-206.

[2] Marden, M. The geometry of the zeros of a polynomial in a complex variable. American Mathematical Society, New York, 1949.

[3] Parodi, M. La localisation des Valeurs Caracteristiques des Matrices et ses Applications, Gauthier-Villars Paris, 1959.

[4] Rall, L.B. Computational solutions of nonlinear operator equations. John Wiley and Sons Publ. 1969.

Rev. Academia de Ciencias. Zaragoza 47 (1992)

ON THE APPROXIMATION OF QUADRATIC EQUATIONS IN BANACH SPACE USING FINITE RANK OPERATORS

Ioannis K. Argyros
Department of Mathematics
Cameron University
Lawton, OK - 73505

Abstract. In this paper a simple relation between the solution of the quadratic equation and the corresponding solution of the discretized equation is derived using finite rank quadratic operators.

Key words and phrases: Banach space, quadratic equation, quadratic operator.

A.M.S. 1980 classification codes: 46(B15), 65J15.

Introduction. Consider the quadratic equations

$$x = y + Q(x) \tag{1}$$

$$x = y + B(x,x) \tag{2}$$

in a Banach space X, where $y \in X$ is fixed, Q is a bounded quadratic operator on X, B is a symmetric bounded bilinear operator on X such that

$$Q(x) = B(x,x) \quad \text{for all} \quad x \in X, \ [12], \ [13]. \tag{3}$$

Obviously if (3) holds the solutions of equations (1) and (2) coincide.

A large number of very interesting applied problems that are particular cases of (1) (or (2)) arise in neutron transport [4], [9], elasticity theory [3], [8] and in some bending of beams problems [1], [5].

The iterations

$$x_{n+1} = y + Q(x_n) \tag{4}$$

or

$$x_{n+1} = y + B(x_n, x_n) \tag{5}$$

for some $x_0 \in X$ have been used to approximate a solution z^* of (1) or (2), [4], [9], [11].

Since the iterations (4) or (5) can rarely be executed in infinite dimensional spaces (see e.g. the equations in [4],[11]) if we set

$$x = y + Q_n(x) \tag{6}$$

where Q_n is a bounded quadratic operator on X with $\dim(\text{span}(\text{Rang}(Q_n))) = n$. In [6] we showed that (6) has a solution $x \in X$ if and only if

$$x = y + \sum_{i=1}^{n} v_i^* b_i \tag{7}$$

where the vector $\underline{v}^* = [v_1^*, \cdots, v_n^*]^{tr} \in \mathbb{R}^n$ is a solution of the quadratic system in $\mathbb{R}^n$

$$\underline{v} = \underline{y} + \overline{Q}_n(\underline{v}) \tag{8}$$

or

$$\underline{v} = \underline{y} + \overline{B}_n(\underline{v}, \underline{v}) \ , \tag{9}$$

and $\{b_i\}$, $i = 1, 2, \ldots, n$ is a unit norm basis for the span of the range of Q_n.

Here, $\underline{y}$ is fixed in $\mathbb{R}^n$, $\overline{Q}_n$ is a bounded quadratic operator on $\mathbb{R}^n$ and $\overline{B}_n$ is the unique symmetric bilinear operator such that

$$\overline{Q}_n(\underline{v}) = \overline{B}_n(\underline{v}, \underline{v}) \quad \text{for all } \underline{v} \in \mathbb{R}^n \ . \tag{10}$$

The explicit values of $\underline{y}$, $\overline{Q}_n$, $\overline{B}_n$ can be found in [4], [5] or [6].

Since each quadratic operator Q_n is of finite rank n, for each $x \in X$ there exist real numbers $v_i = v_i(x)$, $i = 1, 2, \cdots, n$ such that

$$Q_n(x) = \sum_{i=1}^{n} v_i b_i .$$

Moreover $X_n = \text{Rang}(Q_n) \approx \mathbb{R}^n$. Therefore for each $x \in X$ we can identify the element $Q_n(x) \in X$ with the vector $\underline{v} = [v_1, v_2, \cdots, v_n]^{tr}$. That is $Q_n(x) \approx \underline{v}$ and $\|Q_n(x)\|_X = \|\underline{v}\|_{\mathbb{R}^n} = \sum_{i=1}^{n} |v_i|$.

In this paper under certain assumptions we associate the solutions of (1) (or (2)), (6), and (8) (or (9)). In particular we show that if z^* is a solution of (1), $\underline{v}^*$ is a solution of (8) and $Q_n(z^*) \approx \underline{q} = [q_1, \cdots, q_n]^{tr} \in \mathbb{R}^n$ then

$$\underline{v}^* = Q_n(z^*) + O\left[\frac{1}{n}\right] . \tag{11}$$

From now on the norms in all spaces will be denoted by the same symbol $\| \ \|$. We will also identify equation (1) with (2), iteration (4) with (5) and equation (8) with (9).

MAIN RESULTS.

The following theorem for (1) is given in [4].

Theorem 1. Let B be a bounded bilinear operator on X and suppose y and z belong to X. Define $T : X \longrightarrow X$ by

$$T(x) = y + B(x,x) . \tag{12}$$

Set

$$a = \frac{1}{2\|B\|} - \|z\| , \tag{13}$$

$$b = a - \left[a^2 - \frac{\|T(z)-z\|}{\|B\|}\right]^{1/2} , \tag{14}$$

and assume b is nonnegative and $a \neq 0$. Then

(i) T has a unique fixed point in $U(z,a) = [x \in X \ / \ \|x - z\| < a]$;

(ii) this fixed point actually lies in $\overline{U}(z,b)$.

Corollary. Let B, y be as in Theorem 1. Assume $z = z^* \in X$ is a solution of (1) and $a > 0$.

Then z^* is the unique solution of (1) in $U(z^*,a)$. Moreover, the sequence (5) converges to z^* for any $x_0 \in U(z^*,a)$.

Proof. Take $z = z^*$ in Theorem 1.

Let z^* be a solution of (1) and $\underline{v}^* = v^*(n)$ be a solution of (8).

Assume that the sequence $\{Q_n(z)\}$ converges in norm to $\{Q(z)\}$ as $n \to \infty$ with $z = z^*$ or $x_0 \in X$. Then there exists $c > 0$, $N_1 \in N$ such that:

$$\|Q_n(z) - Q(z)\| \leq c \cdot \frac{1}{n} \tag{15}$$

for all $n \geq N_1$.

Let B, $\overline{B}_n$ be the unique symmetric bilinear operators associated with Q and $\overline{Q}_n$ respectively. Then it is well known [11], [12] that the following equations hold:

$$Q(x) = B(x,x) \quad \text{for all} \quad x \in X \tag{16}$$

$$\overline{Q}_n(\underline{v}) = \overline{B}_n(\underline{v},\underline{v}) \quad \text{for all} \quad \underline{v} \in \mathbb{R}^n \tag{17}$$

and

$$Q(w_1 + w_2) = Q(w_2) + 2B(w_1,w_2) + Q(w_2) \tag{18}$$

$$\overline{Q}_n(\underline{v}_1 + \underline{v}_2) = \overline{Q}_n(\underline{v}_1) + 2\overline{B}_n(\underline{v}_1,\underline{v}_2) + \overline{Q}_n(\underline{v}_2) \ . \tag{19}$$

Set:

$$a(n) = \frac{1}{2\|\overline{B}_n\|} - \|Q_n(z^*)\| \ , \tag{20}$$

$$b(n) = a(n) - \left[a(n)^2 - \frac{\|\overline{T}(Q_n(z^*)) - Q_n(z^*)\|}{\|\overline{B}_n\|} \right]^{1/2} \tag{21}$$

$$\overline{a}(n) = \frac{1}{2\|\overline{B}_n\|} - \|\underline{v}^*\| \ . \tag{22}$$

It is known [12], [13] that B is a bounded bilinear operator if and only if Q is and the following estimate holds:

$$\|Q\| \leq \|B\| \leq 2\|Q\| \ . \tag{23}$$

If $\{Q_n\}$, $n = 1, 2, \ldots$ is a bounded sequence of finite rank, n ($n = \dim(\text{spanRang}(Q_n))$) then the definition of the $\overline{Q}_n$ and $\overline{B}_n$ in [6] obviously implies that they are bounded operators on $\mathbb{R}^n$. Then there exists

$$D_1, D_2, E > 0$$

such that

$$D_1 \leq \|\bar{B}_n\| \leq D_2 \tag{24}$$

and

$$\|\bar{Q}_n\| \leq E \quad (\leq 2D_2 \text{ by (23)}) . \tag{25}$$

Define the real functions

$$f_1(n) = A_1\left[\frac{1}{n}\right]^2 + A_2\left[\frac{1}{n}\right] + A_3 , \qquad f_2(n) = A_1'\left[\frac{1}{n}\right]^2 + A_2'\left[\frac{1}{n}\right] + A_3' ,$$

$$f_3(d) = A_2''d^2 + A_2'd + A_3'' , \qquad f_4(r) = \bar{A}_1 r^2 + A_2 r + \bar{A}_3$$

where,

$$A_1 = 4D_2^2Ec^2 , \qquad A_1'' = D_1 ,$$

$$A_2 = 8D_2^2c(1 + 2E\|Q(z^*)\|) , \qquad A_2'' = -2D_1c ,$$

$$A_3 = 4D_2^2(\|Q(z^*)\| + \|\bar{T}(Q(z^*)) - Q(z^*)\|) - D_1 , \qquad A_3'' = Ec^2 ,$$

$$A_1' = D_1d^2 - 2D_1cd + Ec^2 , \qquad \bar{A}_1 = D_2 ,$$

$$A_2' = c + 4Ec\|Q(z^*)\| - 2D_1d\,\|Q(z^*)\| , \qquad \bar{A}_2 = 2D_2\|z^*\| ,$$

$$A_3' = \|\bar{T}(Q(z^*)) - Q(z^*)\| , \qquad \bar{A}_3 = 2d\left[\frac{1}{n}\right] - R ,$$

$$\bar{T}(Q(z^*)) = \underline{y} + \bar{Q}_n(Q(z^*)) .$$

Let us assume that either:

Case 1:

(a) $A_3 < 0$

(b) $1 - 2D_2\|Q(z^*)\| > 0$

(c) $D_1 \leq E$

(d) there exists d such that $A_2' < 0$, $(A_2')^2 - 4(A_1')(A_3)' \geq 0$

(e) $n \geq \max\left[\dfrac{2cD_2}{1-2D_2\|Q(z^*)\|} , \dfrac{2D_1c}{1-2D_1\|\underline{v}^*\|-2D_1\|Q(x_0)-\underline{v}^*\|} ,\right.$

$\left. n_1 , n_2 , N_1 , \dfrac{2(d+c)D_2}{1-2D_2\|Q(z^*)\|} , \dfrac{2d}{R}\right] = N , R > 0$

(f) $1 - 2D_1\|\underline{v}^*\| - 2D_2\|Q(x_0) - \underline{v}^*\| > 0$ if $\underline{v}^*$ is a solution of (8)

and

(g) $0 \leq \|x_0 - z^*\| \leq r^+$, where r^+ is the positive solution of the equation

$$f_4(r) = 0 .$$

where n_1, n_2 are the large solutions of the equations

$$f_1(n) = 0$$

and

$$f_2(n) = 0$$

respectively.

Or,

Case 2:

(a') $A_3 < 0$,

(b') $1 - 2D_2\|Q(z^*)\| > 0$,

(c') $D_1 > E$,

(d') $d \in (d_1, d_2)$, where d_1, d_2 are the solutions of the equation

$$f_3(d) = 0 ,$$

(e') $n \geq N$, where N is as defined above,

(f'), (f)

and

(g'), (g) .

Or finally:

Case 3:

(a'') $A_3 < 0$,

(b'') $1 - 2D_2\|Q(z^*)\| > 0$,

(c'') $D_1 > E$

(d'') there exists d such that: $d < d_1$, or $d > d_2$,

$$A_2' < 0$$

and

$$(A_2')^2 - 4(A_1')(A_3') \geq 0 ,$$

(e'') $n \geq N$, where N is as defined above ,

(f''), (f)

and

(g''), (g) .

Then we can easily verify that in all the above cases:

$$0 < \frac{1}{2\|\bar{B}_n\|} - (\|Q(z^*) - Q_n(z^*)\| + \|Q(z^*)\|) \leq a(n) , \tag{26}$$

$$f_1(n) \leq 0 , \tag{27}$$

$$f_2(n) \leq 0 . \tag{28}$$

Note that since:

$$\|\bar{T}(Q_n(z^*)) - Q_n(z^*)\| = \|(\bar{T}(Q(z^*)) - Q(z^*)) + \bar{B}_n(Q_n(z^*) - Q(z^*))$$

$$- 2\bar{B}_n(Q_n(z^*) - Q(z^*), Q(z^*)) + (Q(z^*) - Q_n(z^*))\|$$

$$\leq \|\bar{T}(Q(z^*)) - Q(z^*)\| + Ec^2\left[\frac{1}{n}\right]^2 + 4Ec\left[\frac{1}{n}\right]\|Q(z^*)\| + c \cdot \left[\frac{1}{n}\right] \tag{29}$$

(The first norm immediately above is assumed in X by identifying $\underline{y}$, $\bar{Q}_n(Q(z^*)) \in \mathbb{R}^n$ with an element in $X_m \subset X$) .

To show that the quantity under the radical in the definition of $b(n)$ is nonnegative it is enough to show that

$$f_1(n) \leq 0 . \tag{30}$$

Also, to show that

$$b(n) \leq d\left[\frac{1}{n}\right] , \tag{31}$$

evidently it is enough to verify that

$$a(n) > d\left[\frac{1}{n}\right] \tag{32}$$

and

$$f_2(n) \leq 0 \tag{33}$$

which are true in all the above cases.

Note that in all the above cases

$$(a)\ \ \bar{a}(n) > 0 , \ (b)\ \ \|Q_n(x_0) - \underline{v}^*\| < \bar{a}(n) \tag{34}$$

and

$$\bar{U}(\underline{v}^* , \|Q_n(x_0) - \underline{v}^*\|) \subset \bar{U}(Q_n(z^*), R) . \tag{35}$$

Inequality (34)(a) is immediate from (f) whereas (35) holds if

$$\|\underline{v}^* - Q_n(z^*)\| + \|Q_n(x_0) - \underline{v}^*\| \leq d\left[\frac{1}{n}\right] + \|Q_n(x_0) - Q_n(z^*)\| + \|Q_n(z^*) - \underline{v}^*\|$$

$$\leq 2d\left[\frac{1}{n}\right] + \|B_n(x_0-z^*) - 2B_n(x_0-z^*,z^*)\|$$

$$\leq 2d\left[\frac{1}{n}\right] + D_2\|x_0-z^*\|^2 + 2D_2\|x_0-z^*\|\|z^*\| \leq R \qquad (36)$$

with

$$b(n) \leq d\left[\frac{1}{n}\right] \leq R \qquad (37)$$

or

$$f_4(r) \leq 0 \qquad (38)$$

which is true by (g) .

Finally note that (34)(b) will be true if

$$\|Q_n(x_0) - \underline{v}^*\| \leq \|Q_n(x_0) - Q(x_0)\| + \|Q(x_0) - \underline{v}^*\| \leq \frac{1}{2D_1} - \|\underline{v}^*\| \leq \bar{a}(n)$$

or

$$c \cdot \frac{1}{n} + \|Q(x_0) - \underline{v}^*\| \leq \frac{1}{2D_1} - \|\underline{v}^*\|$$

or

$$n \geq \frac{2D_1 c}{1-2D_1\|\underline{v}^*\|-2\|D_1\|\|(Q(x_0)-\underline{v}^*\|} ,$$

provided that

$$1 - 2D_1\|\underline{v}^*\| - 2D_1\|Q(x_0)-\underline{v}^*\| > 0$$

which is true by the choice of n and (f).

We now state the main result.

Theorem 2. Let z^* be a solution of (1). Assume that (15) holds and that the hypotheses in any one of the cases 1, 2 or 3 hold.

Then

(i) there exists a locally unique solution $\underline{v}^*$ of (18) such that

$\underline{v}^* = Q_n(z^*) + 0\left[\frac{1}{n}\right]$, for all $n \geq N$ with N as defined in (e).

(ii) the sequence $\underline{v}^k = \underline{y} + \bar{Q}_n(\underline{v}^k)$, $k = 0, 1, 2, \ldots$ (39)

converges to $\underline{v}^*$ for $v^0 \approx Q_n(x_0)$ ($0\left[\frac{1}{n}\right]$ the usual order of $\frac{1}{n}$).

Proof. We prove the theorem using the following steps:

Step 1. We show that there exists a solution $\underline{v}^*$ of (4) such that $\underline{v}^*$ is unique in $U(Q_n(z^*), a(n))$. Moreover $\underline{v}^* \in \overline{U}(Q_n(z^*), b(n))$.

We apply Theorem 1 to equation (4). Then if

$$a(n) > 0$$

and the quantity under the radical in the definition of $b(n)$ is nonnegative the conclusion of Step 1 holds. But this is true by (26), (29) and (30).

Step 2. We show that:

$$b(n) \leq d \, O\left[\frac{1}{n}\right] , \quad n \geq N .$$

But this is true by (32) and (33). That proves Step 2 and part (i) of the theorem.

Step 3. We now show that iteration (39) converges to $\underline{v}^*$. We apply the corollary to (39). Then (39) will converge to $\underline{v}^*$ if

$$(a) \quad \overline{a}(n) > 0 , \quad (b) \quad \|Q_n(x_0) - \underline{v}^*\| < \overline{a}(n) \tag{34}$$

and

$$\overline{U}(\underline{v}^* , \|Q_n(x_0) - \underline{v}^*\|) \subset \overline{U}(Q_n(z^*), R) \tag{35}$$

But (34) follows from (f) ((f') or (f'')) and (35) follows from (36), (37), (38) and (g) ((g') or (g'')). That proves (ii) and completes the proof of the theorem.

References

1. Allgower, E. L. and Prenter, P. M. On the branching of solutions of quadratic differential equations. Aequationes Mathematicae, 10, (1974), 81-96.

2. Ames, W. Ordinary differential equations in transport processes. Academic Press, New York, 1968.

3. Anselone, P. M. and Moore, R. H. An extension of the Newton-Kantorovich method for solving nonlinear equations with an application to elasticity. J. Math. Anal. and Applic., 13, (1966), 476-501.

4. Argyros, I. K. Quadratic equations and applications to Chandrasekhar's and related equations. Bull. Austral. Math. Soc., 32, (1985), 275-292.

5. _____. On the approximation of solutions of compact operator equations. Revista Proyecciones, No. 14, (1988), 29-46.

6. _____. Quadratic finite rank operator equations in Banach space. Tamkang J. Math., Vol. 18, No. 4, (1987), 8-19.

7. Chandrasekhar, S. Radiative transfer. Dover Publ., New York, 1960.

8. Davis, H. T. Introduction to nonlinear differential and integral equations. Dover Publ., New York, 1962.

9. Kelley, C. T. Approximation of some quadratic integral equations in transport theory. J. Integral Equations, 4, (1982), 221-237.

10. Ortega, J. M. and Rheinboldt, W. C. Iterative solutions of nonlinear equations in several variables. Academic Press, New York, 1970.

11. Rall, L. B. Quadratic equations in Banach space. Rend. Circ. Math., Palermo, 10, (1961), 314-332.

12. _____. Computational solution of nonlinear operator equations. John Wiley Publ., New York, 1968.

13. _____. Nonlinear functional analysis and applications. Academic Press, New York, 1971.

Rev. Academia de Ciencias. Zaragoza **47** *(1992)*

GENERALIZATION OF TWO RESULTS ON BANACH SPACES WITH BASE

by

F. García–Castellón

Departamento de Matemáticas
Universidad de Zaragoza

Abstract: Some properties in Banach spaces with base are also accomplished for separable Banach spaces. We give here two generalizations, where the base is replaced by a complete sequence.

1. Introduction.

Bessaga and Pelczinski show in [1] (see also [6], p.20) that "every infinite dimensional closed subspace of a Banach space with base (a_n), contains a subspace with a base, which is equivalent to a block–sequence of (a_n)".

On the other hand, I. Singer (s.[9], p.76) shows that "if B is a Banach space with a base (a_n), then every sequence (y_n) verifying $dim[(y_n)] = \infty$, admits a basic block–sequence equivalent to another block–sequence of (a_n)".

We generalize the previous results to every separable Banach space.

2. Definitions, notations and previous results.

Let B be an infinite dimensional Banach space and let B' be the topological dual space.

If (a_i) denotes a sequence in B, we call *nucleus* of (a_i), $K(a_i) := \bigcap_{n=1}^{\infty} [a_n, a_{n+1}, \ldots]$, where $[\ldots]$ denotes the closed linear spand. If $[(a_i)] = B, (a_i)$ is said *complete* and therefore B is separable.

A sequence (a_i) is said *minimal* if there exists $(a_i') \subset B'$, called a conjugate sequence of (a_i), so that the pair (a_i, a_i') is a biorthogonal system (that is $a_i'(a_j) = \delta_{ij}$). If (a_i) is complete, there is at most a unique conjugate sequence of (a_i).

An intrinsic characterization of minimality is

Theorem I. Given (a_i) in B, then (a_i) is minimal if and only if for every n is $[(a_i)] = [a_1, \ldots, a_n] \oplus [a_{n+1}, a_{n+2}, \ldots]$

A minimal sequence with zero nucleus is said an *M–basic* sequence. Moreover, if (a_i) is an M–basic complete sequence, it is called and *M–base* of B. It is known

Theorem II(s.[8]). Every separable Banach space has an M–base.

The sequence (a_i) is said *basic* if for every $x \in [(a_i)]$ there exists a unique sequence of real numbers (λ_i) so that $x = \sum_{i=1}^{\infty} \lambda_i a_i$. We have the following

Theorem III(s.[4]). Every M–basic sequence of a Banach space has a basic subsequence.

The sequence (b_i) is said to be a *block–sequence of* (a_i) if there exists an increasing sequence of natural numbers (q_n) with $q_0 = 0$ and $b_n \in [(a_i)_{i=q_{n-1}+1}^{i=q_n}]$

From theorem III and theorem 8 of [3],p.371, it follows

Theorem IV. For every sequence (a_i) in B, it is verifyed $K(a_i) = 0$ if and only if every block–sequence of (a_i) has a basic subsequence.

About the stability of minimal sequences, we have the following

Theorem V(s.[5]). Let (a_i) be a normalized ($\|a_i\| = 1, i = 1, 2 \ldots$) minimal sequence, (a_i') a conjugate sequence of (a_i) and (ε_i) a sequence of positive real numbers verifying $\sum_1^{\infty} \varepsilon_i \|a_i'\| < 1$. Then every sequence (y_i) with $\|y_i - a_i\| < \varepsilon_i$, is equivalent to (a_i). Moreover, if (a_i) is complete, (y_i) is complete also. Consequently (y_i) is minimal, and

if (a_i) is basic, then so is (y_i). If $\sum_1^\infty \varepsilon_i \|a_i'\| < +\infty$, the above conclusions are valid, possibly after removing a finite number of elements from (a_i) and (y_i).

Finally, we remember that two closed linear subspaces M, N are called *quasi-complemented* if $M \cap N = 0$ and $\overline{M+N} = B$. In this relation, we have

Theorem VI(s.[7]). Every closed subspace M of a separable Banach space has a quasi-complement.

3. Two generalizations.

Let us see now our two results.

Theorem 1. Let B be a separable Banach space and (a_i) a complete sequence. Then every infinite dimensional closed subspace Y of B contains a closed subspace Y_0 with a base equivalent to a block-sequence of (a_i).

Proof:

We consider the closed subspace $Y \cap K(a_i)$. There are two cases.

a) $Codim_Y Y \cap K(a_i) = \infty$. Then, by theorem VI, $Y \cap K(a_i)$ has a quasicomplement in Y denoted by Y_1. Obviously $Y_1 \cap K(a_1) = 0$ and moreover, for every n, $Y_1 \cap [a_n, a_{n+1}, \ldots] \neq 0$ since Y_1 is infinite dimensional.

Let $(\varepsilon_i) \in l_1$ fixed, with $\varepsilon_i > 0$.

Because $Y_1 \cap [a_1, a_2, \ldots] \neq 0$, there exists $y_1 \in Y_1, \|y_1\| = 1$ and $b_1 = t_1 a_1 + t_2 a_2 + \ldots + t_{p_1} a_{p_1}$, such that

$$\|y_1 - b_1\| < \varepsilon_1$$

Since $Y_1 \cap K(a_i) = 0$ we have $y_1 \notin K(a_i)$. Thus it is possible to find an index q_1 such that $y_1 \notin [a_{q_1}, a_{q_1+1}, \ldots]$ with $q_1 > p_1$. Because $Y_1 \cap [a_{q_1}, a_{q_1+1}, \ldots] \neq 0$, there exists $y_2 \in Y_1, \|y_2\| = 1$ and $b_2 = t_{q_1} a_{q_1} + \ldots + t_{p_2} a_{p_2}$, where $p_2 = q_1 + m$, such that

$$\|y_2 - b_2\| < \varepsilon_2$$

We may iterate the process indefinitely and thus we find finally two sequences $(y_i) \subset Y_1$ and (b_i), block–sequence of (a_i). Moreover $[y_n, y_{n+1}, \ldots] \subset [a_{q_{n-1}}, a_{q_{n-1}+1}, \ldots]$ and therefore $K(y_i) \subset Y_1 \cap K(a_i)$. We conclude $K(y_i) = 0$.

Let (y_{n_k}) be a basic subsequence of (y_i). It is sufficient to take $(y_{n_k}), (b_{n_k})$ and to apply the theorem V. Here the desired subspace is $Y_0 = [(y_{n_k})]$.

b) $Codim_Y Y \cap K(a_i) = p$. We denote $Y_1 = Y \cap K(a_i)$. By theorem II, Y_1 admits an M–base (y_i) which we can suppose normalized.

We fix $(\varepsilon_i) \in l_1$ where $\varepsilon_i > 0$.

Since $y_1 \in [(a_i)]$ there exists $b_1 = t_1 a_1 + t_2 a_2 + \ldots + t_{p_1} a_{p_1}$ such that

$$\|y_1 - b_1\| < \varepsilon_1$$

From $y_2 \in Y_1$ and $Y_1 \subset K(a_i)$, it follows $y_2 \in [a_{p_1+1}, a_{p_1+2}, \ldots]$. Hence there exists $b_2 = t_{p_1+1} a_{p_1+1} + \ldots + t_{p_2} a_{p_2}$ so that

$$\|y_2 - b_2\| < \varepsilon_2$$

We can iterate this process, and we find finally a block–sequence (b_i) of (a_i). Now, by theorem III, there is a basic subsequence (y_{n_k}) of (y_i). It is sufficient to take (y_{n_k}), (b_{n_k}) and to apply the theorem V. Here the requested subspace is $Y_0 = [(y_{n_k})]$. $\Diamond$

Observe that the result of Dean–Singer–Sternbach (s.[2]) "every sequence (a_n), $dim[(a_n)] = \infty$, of a Banach space admits a basic block–sequence" is a direct consequence of theorem 1.

Theorem 2. Let B be a separable Banach space and (a_i) a complete sequence. Then every sequence $(y_n), dim[(y_n)] = \infty$, has a basic block–sequence equivalent to a block–sequence of (a_i).

Proof:

By theorem 1, (y_n) contains a basic block–sequence (b_j). We fix $(\varepsilon_i) \in l_1$ with $\varepsilon_i > 0$.

First $[b_1, b_2] \cap [a_2, a_3, \ldots] \neq 0$ since the second space is at least of codimension 1. Hence there are $z_1 = t_1 b_1 + t_2 b_2$, $\|z_1\| = 1$ and $c_1 = s_2 a_2 + s_3 a_3 + \ldots + s_{n_1} a_{n_1}$ such that

$$\|z_1 - c_1\| < \varepsilon_1$$

It is true also $[b_3, b_4, \ldots, b_{n_1+3}] \cap [a_{n+1}, a_{n+2}, \ldots] \neq 0$ since the second space is at least of codimension n_1. So it is possible to take $z_2 = t_3 b_3 + t_4 b_4 + \ldots + t_{n_1+3} b_{n_1+3}$, $\|z_2\| = 1$, and $c_2 = s_{n_1+1} a_{n_1+1} + \ldots + s_{n_2} a_{n_2}$ such that

$$\|z_2 - c_2\| < \varepsilon_2$$

We can iterate the process indefinitely and we find finally two sequences $(z_i), (c_i)$ verifying:

i) (z_i) is a basic block–sequence of (y_n) and $\|z_i\| = 1$

ii) (c_i) is a block–sequence of (a_n).

By theorem V we obtain the desired result. $\diamondsuit$

BIBLIOGRAPHY

[1] Bessaga C. and Pelczynski A., *"On bases and unconditional convergence of series in Banach spaces"*, Studia Math., **17**(1958), 151–164.

[2] Dean D., Singer I., Sternbach L., *"On shrinking basic sequences in Banach spaces"*, Studia Math., **40**(1971), 23–33.

[3] García–Castellón y Plans A., *"On linear simplification in Banach spaces II"*, Boll. Un. Mat Italiana, Serie **VII 1-A**(1987), 367–373.

[4] Kadets M.I., Pelczynski A., *Basic sequences, biorthogonal and norming sets in Banach and Frechet spaces"*, Studia Math., **125**(1965), 297–333.

[5] Krein M.G., Mil'man D.P., Rutman M.A., *"A note on a basis in Banach space"*, Comm. Inst. Sci. Math., Univ. Kharkoff (Zap. Inst. Mat. Mech.) **16**(1940), 106–110.

[6] Lindenstrauss J. and Tzafriri, *"Classical Banach spaces"*, Berlin–Heidelberg New York, Springer–Verlag (1973).

[7] Mackey G.W., *"Note on a theorem of Murray"*, Bull. Amer. Math. Soc. **52**(1946), 322–325.

[8] Markushevich A.I., *"Sur les bases (au sens large) dans les espaces lineaires"*, Doklady Akad. Nauk., SSSR **41**(1943), 227–229.

[9] Singer I., *"Bases in Banach Spaces II"* Springer–Verlag, Berlin, New York (1981).

Rev. Academia de Ciencias. Zaragoza 47 *(1992)*

On the Holonomy Bundle of the Sphere

by

Fernando Etayo and Ujué R. Trías

Departamento de Matemáticas, Estadística y Computación

Facultad de Ciencias. Universidad de Cantabria

Avda. de los Castros, s/n. 39071 SANTANDER

ABSTRACT. Given a connection in a manifold M we can define a locally injective map from the horizontal distribution in a point of the frame bundle into the holonomy bundle. Using this map, we can study the holonomy bundle of the sphere, proving that it is the real projective 3-space, and that it can be decomposed in real projective planes which are the image of the above map when we move the initial point.

A.M.S. Math. Subj. Class. 53C05, 57N10.

§1. Introduction.

Let M be a real manifold, $\pi: FM \longrightarrow M$ the frame bundle. Assume that a connection Γ is given in the bundle. Then, following [3], for each $u \in FM$ we obtain H_u and $\mathcal{H}_u$ which are the following sets: H_u is the *horizontal distribution* defined by Γ, i. e., $T_u FM$ is decomposed in $H_u \oplus Q_u$, Q_u being the tangent space to the fibre at $p=\pi(u)$, and $\mathcal{H}_u$ is the *holonomy bundle* at u, i. e., the set of every point of FM that can be joined with u by an horizontal path. We want to define a map $\alpha_u : H_u \longrightarrow \mathcal{H}_u$.

Let $v \in H_u$. Then, $\pi_*(v) \in T_p M$. Let exp the *exponential map*, $\exp: T_p M \longrightarrow M$, $\exp(\pi_*(v)) = \gamma_v(1)$, γ_v being the unique geodesic on M such that $\gamma_v(0)=p$, $\dot{\gamma}_v(0)=\pi_*(v)$. Finally, let γ_v^h be the horizontal lift of γ_v with $\gamma_v^h(0)=u$. We define $\alpha_u(v) = \gamma_v^h(1)$, which belongs to $\mathcal{H}_u$. The map α_u is smooth and locally injective, because the exponential map is a local difeomorphism.

Remark. If (M,g) is a Riemannian manifold, we can give a Riemannian structure on FM with the Sasaki-Mok metric G, [4]. In this case, it is easy to see that α_u coincides with the restriction to the horizontal distribution of the exponential map of the Levi-Civita connection of G, because γ_v^h is a G-geodesic, [1].

§2. The holonomy bundle of the sphere.

Let S^2 be the 2-sphere with the canonical connection, and let $u \in FS^2$, $p=\pi(u)$, π being the bundle projection. In this case, dim S^2 = 2, dim FS^2 = 6, dim $\mathcal{H}_u$ = dim S^2 + dim Φ_u = 3, Φ_u being the holonomy group of the connection, $\Phi_u = SO(2) = S^1$. Let $\tilde{\pi}: \mathcal{H}_u \longrightarrow M$ the holonomy bundle.

Let p be the north pole of S^2 and q the south pole. It is obvious that $\alpha_u: U \longrightarrow \alpha_u(U)$ is a difeomorphism, where $U = (\pi_*|_{\mathcal{H}_u})^{-1}(\exp^{-1}(V))$ and $V = S^2 - \{q\}$.

Using the above notation, we are going to prove the following

THEOREM. *let* S^2 *be the 2-sphere with the canonical connection, and let* $u \in FS^2$, $p=\pi(u)$. *Then:*

(i) The set $\overline{\alpha_u(U)}$ *is a real projective plane. Moreover, the projection* $\tilde{\pi}: \alpha_u(U) \longrightarrow S^2 - \{q\} = V$ *is a difeomorphism and* $\tilde{\pi}^{-1}(q) \cap \overline{\alpha_u(U)} = \tilde{\pi}^{-1}(q) = S^1$.

(ii) The holonomy bundle $\mathcal{H}_u$ *can be decomposed as* $\mathcal{H}_u = \bigcup_{i \in I} P_i$, P_i *being* $\overline{\alpha_{u_i}(U_i)}$, *and* $u_i \in \mathcal{H}_u \cap \pi^{-1}(p)$. *Moreover,* $P_i \cap P_j = \tilde{\pi}^{-1}(q) = S^1$.

(iii) The holonomy bundle $\mathcal{H}_u$ *is a real projective 3-space.*

Proof.

(i) In the south pole q, all the radial geodesics begining at p coincide, but each one of them moves the given reference u in a different

way. Of course, two opposit geodesics move u in the same way. Topologically, $\alpha_u(U)$ is an open disk and $\overline{\alpha_u(U)}$ is a disk with antipodal boundary points identified: a real projective plane. The fiber $\tilde{\pi}^{-1}(q)$ is the set of all the references at q that can be obtained by moving u from p to q along any path. This set is the fiber of the holonomy bundle, S^1.

(ii) It is easy, from (i).

(iii) Using (ii) we know that $\mathcal{H}_u = \bigcup_{i\in I} P_i$, where P_i is the projective plane corresponding to the point $u_1 \in \mathcal{H}_u \cap \pi^{-1}(p) = S^1$, because $\tilde{\pi}:\mathcal{H}_u \longrightarrow S^2$ is a S^1-principal bundle. Then, we have $\mathcal{H}_u$ as a S^1-family of real projective planes with a circunference S^1 belongning to any of these planes; this cicunference is the $\tilde{\pi}$-fiber of the south pole q. Then, we can visuallize $\mathcal{H}_u$ as in fig. 1:

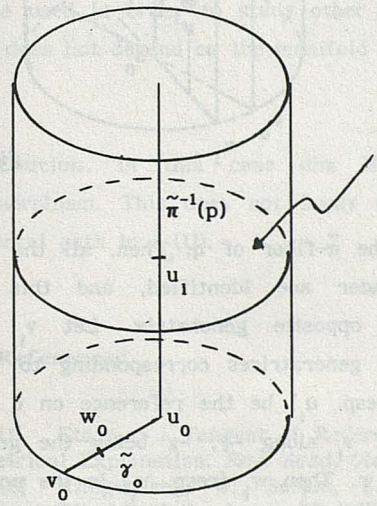


fig. 1

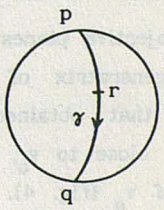


fig. 2

Let r be any point of $S^2-\{p,q\}$, and let γ be the unique geodesic joining p and q through r (fig. 2). Let u_0 be a reference on p, belongning to $\mathcal{H}_u$, v_0 be the reference obtained on q by parallel transport moving u_0 along γ. Let $\tilde{\gamma}_0$ be the horizontal lift of γ begining at u_0 and w_0 be the reference on r which belongs to the image of $\tilde{\gamma}_0$. Then, $\tilde{\gamma}_0$ is the radius of the disk corresponding to u_0 through w_0.

Now, we are going to study the $\tilde{\pi}$-fiber of p, q and r, in order to prove that $\mathcal{H}_u$ is the real projective 3-space. We can see $\mathcal{H}_u$ in the following way: we have a family of disks with identified antipodal boundary points

(these are the real projective planes). This family is parametrized by the $\tilde{\pi}$-fiber of p, which is S^1. This fiber can be viewed as the unit interval with identified edges. Then, $\mathcal{H}_u$ is a solid cylinder, with some identifications (fig. 1 and 3).

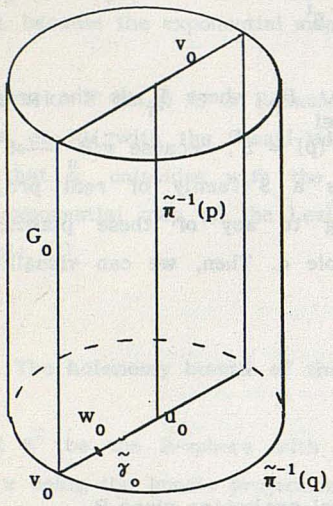


fig. 3

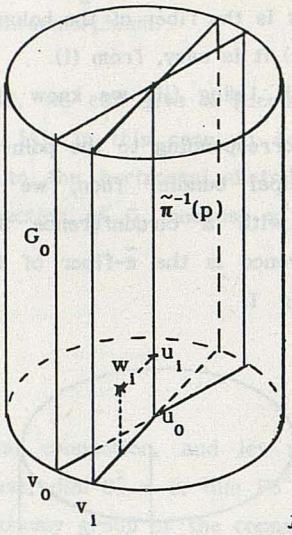


fig. 4

All the projective planes contain the $\tilde{\pi}$-fiber of q. Then, all the points in the same generatrix of the cylinder are identified, and this point coincides with that obtained in the opposite generatrix. Let v_1 be a reference on q, close to v_0, i. e., the generatrices corresponding to v_1 are near to those of v_0 (fig. 4). Let w_1 (resp. u_1) be the reference on q (resp. on p) obtained by parallel transport of v_1 along $-\gamma$, $-\gamma$ being the geodesic begininig at q and ending at p, through r. Then w_i (resp. u_1) is the point of the $\tilde{\pi}$-fiber of r (resp. of p) belongning to the image of $\tilde{\gamma}_1$, the horizontal lift of $-\gamma$ begining at v_1. Then, u_1 is close to u_0 (actually, the angle between the references u_1 and u_0 is the same that the angle between v_1 and v_0). Then, $w_1 = \tilde{\pi}^{-1}(r) \cap P_1$, and w_1 belongs to the radius of center u_1 aiming to the generatrix over v_1 which is the closest to G_0, G_0 being the generatrix over v_0 such that w_0 is in the radius aiming G_0 (see fig 3 and 4).

When v_1 runs $\tilde{\pi}^{-1}(q)$, then u_1 runs $\tilde{\pi}^{-1}(p)$, and v_1 and u_1 are in the top plane of the cylinder. And the points w_1 draw a helix, in such a way that $w_1 \in P_1$ and $w_0 \in P_0$ are in opposite place. Then we have to glue the top and the bottom planes after a rotation of 180°. Now it is easy to see that the quotient space is the solid sphere with identified antipodal boundary points,

i. e., the real projective 3-space.

QED

Remark. There exists an easier proof of (iii). The holonomy bundle $\mathcal{H}_u$ is the set of all the references over the sphere that can be obtained by moving the reference u along any path. Then, $\mathcal{H}_u$ is the set of all the references over the sphere given by two vectors with the same angle and norm of that those of u, and we need only the first of those vectors in order to obtain the other one. So, $\mathcal{H}_u$ is diffeomorphic to the spherical tangent bundle of S^2. And it is well known that the spherical tangent bundle of S^2 is the real projective 3-space (see [2], [5]).

The main interest of our long proof is the following: we have obtained this projective space as a family of "natural" projective planes. This idea can be used in order to study other holonomy bundles, because the definition of α_u does not depend on the manifold M.

Caution. In this case dim H_u < dim $\mathcal{H}_u$ and $\tilde{\pi}:\alpha_u(U) \longrightarrow V$ is a difeomorphism. This does not imply that any path in V can be lifted to a horizontal path in $\alpha_u(U)$.

References.

[1] Etayo, F.: Tangent of References embeds in References of Tangent: A Geometrical Explanation. *Rev. Acad. Ciencias Zaragoza* **45** (1990) 55-67.

[2] Klingerberg, W., Sasaki, S.: On the Tangent Sphere Bundle of a 2-Sphere. *Tôhoku Math. Journ.* **27** (1975), 49-56.

[3] Kobayashi, S., Nomizu, K.: *Foundations of Differential Geometry I.* J. Wiley, N. York, 1963.

[4] Mok, K. P.: On the Differential Geometry of Frame Bundle of Riemannian Manifolds. *Jour. reine Angew. Math.* **302** (1978) 16-31.

[5] Montesinos, J. M.: *Classical Tessellations and Three-Manifolds.* Springer, 1985.

Rev. Academia de Ciencias. Zaragoza 47 (1992)

APPLICATIONS ÉCONOMIQUES DES SUITES DANS LES ESPACES VECTORIELS TOPOLOGIQUES

par

Juan Carlos CANDEAL[1] , Carlos HERVÉS[2] et Esteban INDURAIN[3]

1) Universidad de Zaragoza. Departamento de Análisis Económico.
2) Universidad Carlos III de Madrid. Departamento de Economía.
3) Universidad Pública de Navarra. Departamento de Matemática e Informática.

Dirección postal : Esteban INDURAIN . Universidad Pública de Navarra. Departamento de Matemáticas e Informática. El Sadar s/n . 31006. PAMPLONA

Abstract : We introduce here a survey of applications to General Equilibrium Theory of concepts and ideas related to sequences in Topological Vector Spaces .

A.M.S. Subject Class. (1980) : 90 a 14 .
Key words : General Equilibrium. Sequences in topological vector spaces.

1. Avant-propos .

-Avec ce travail, on veut répondre à la question suivante : ... *À quoi bon les suites en Économie ?* On montrera quelques exemples, et un faisceau de résultats, choisis des plus célèbres branches de l' Économie Mathématique (où l' Analyse Fonctionnelle des *suites* joue un rôle fondamental) : La notion d' Équilibre et les Économies à une infinité de biens. Notre intention générale est d' y faire patent , aux mathématiciens (analystes) , le rôle de l' Économie Mathématique : Clarifier (et les munir de rigueur mathématique) les hypothèses utilisées et déduire exactement leurs conséquences , étant ouverte sur les applications.

Ce travail correspond à la communication << À QUOI BON LES SUITES EN ÉCONOMIE >> présentée par les auteurs à l' occasion du "COLLOQUE SUR GÉOMETRIE DES SUITES ET THÉORIE DES OPÉRATEURS" (Luminy-Marseille , FRANCE , juin 1991) .

2. La notion d' équilibre en économies finies .

-D ' abord, nous allons étudier le cas d ' une économie *finie* . (*Le modèle Arrow-Debreu* , paru en 1954 : voir Arrow et Debreu, [1954] Arrow, [1963] , Debreu, [1959], [1974] ou Ekeland, [1979]). La considération des économies infinies sera une généralisation naturelle.

Soit I = {1,2,...,n} un ensemble fini d'agents économiques (individus) . Chaque agent (i) est muni d' une *dotation initiale de ressources* $\omega_i = (\omega_{i,1}, \ldots, \omega_{i,k}) \in \mathbb{R}^{k,+}$ ($\mathbb{R}^{k,+} = \{(y_1,\ldots,y_k) ; y_j > 0 , j = 1,2,\ldots,k\}$) . C' est à dire : $\omega_{i,j}$ est la quantité du bien "j" que l' agent "i" possède au début, et il y a "k" biens .

Un *panier de biens* est un vecteur de $\mathbb{R}^{k,+}$.

$\omega = \omega_1 + \ldots + \omega_n \in \mathbb{R}^{k,+}$ dénotera *le vecteur des ressources totales* de l' économie, c' est à dire : ω représente tout ce dont les agents pourront disposer.

On appelle *allocation* tout vecteur $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^{kn,+}$ ($\alpha_j \in \mathbb{R}^{k,+} ; j = 1,2,\ldots,n$). (Équivalentement, , une allocation peut être considérée comme une fonction $f : I \dashrightarrow \mathbb{R}^{k,+}$) . L' allocation $(\omega_1, \ldots, \omega_n)$ est dite *allocation initiale*. On dit qu' une *allocation* $(\alpha_1, \ldots, \alpha_n)$ est *réalisable* si $\alpha_1 + \ldots + \alpha_n = \omega$. (Intuitivement, une allocation réalisable est un mode quelconque de *partager les disponibilités globales* entre les agents).

Chaque agent "i" a définie une relation $\preceq_i$ de *préférence individuelle* (un préordre total) sur $\mathbb{R}^{k,+}$. $\mathfrak{X} \preceq_i \mathfrak{Y}$ (ou, avec une notation équivalente, $\mathfrak{Y} \succeq_i \mathfrak{X}$) dénotera : "pour l'agent "i" , $\mathfrak{X}$ est plus préféré ou indifférent à $\mathfrak{Y}$ " ($\mathfrak{X}$, $\mathfrak{Y} \in \mathbb{R}^{k,+}$) ; $\mathfrak{X} \prec_i \mathfrak{Y}$ "pour l'agent "i" , $\mathfrak{X}$ est strictement plus préféré que $\mathfrak{Y}$ " (on doit remarquer que $\mathfrak{X} \prec_i \mathfrak{Y}$ est le contraire de $\mathfrak{Y} \preceq_i \mathfrak{X}$) ; et $\mathfrak{X} \sim_i \mathfrak{Y}$ "pour l'agent "i" , les allocations $\mathfrak{X}$ et $\mathfrak{Y}$ sont indifférentes ". Chaque relation $\preceq_i$ sera *continue* : Cela veut dire que pour tout $\mathfrak{Y}_0 \in \mathbb{R}^{k,+}$, les ensembles $\{\mathfrak{X} \in \mathbb{R}^{k,+} ; \mathfrak{X} \succeq_i \mathfrak{Y}_0\}$ et $\{\mathfrak{X} \in \mathbb{R}^{k,+} ; \mathfrak{X} \preceq_i \mathfrak{Y}_0\}$ sont fermés en $\mathbb{R}^{k,+}$. Et , finalement, chaque relation $\preceq_i$ sera aussi *convexe* (c ' est à dire : quand on a $\mathfrak{X} \succeq_i \mathfrak{Y}$, on a aussi, pour tout $t \in [0,1]$, que $t\mathfrak{X} + (1-t)\mathfrak{Y} \succeq_i \mathfrak{Y}$).

Une *économie* $\mathcal{E}$ représentera l ' ensemble suivante :

$\mathcal{E} = \{ I = \{1, 2, \ldots, n\} ; (\preceq_i)_{i=1,2\ldots,n} ; (\omega_i)_{i=1,2\ldots,n} \}$.

Le modèle antérieur est appelé "*Modèle d' Arrow - Debreu d' une économie finie* ".

On appelle *équilibre* en $\mathcal{E}$ tout élément $e = \{(\mathfrak{X}_1, \ldots, \mathfrak{X}_n , p)\}$ où $(\mathfrak{X}_1, \ldots, \mathfrak{X}_n) \in \mathbb{R}^{kn,+}$ est une allocation ($\mathfrak{X}_j \in \mathbb{R}^{k,+} ; j = 1,2,\ldots,n$) , et $p \in \mathbb{R}^{k,+}$, vérifiant :

(i) $(\mathfrak{X}_1, \ldots, \mathfrak{X}_n)$ est réalisable ,

(ii) $\langle p , \mathfrak{X}_i \rangle \leq \langle p , \omega_i \rangle$, pour tout $i \in I$ (l' expression $\langle p, \mathfrak{X} \rangle$ dénotera le *produit scalaire* en $\mathbb{R}^k$) .

(iii) Chaque $\mathfrak{X}_i$ est maximal: si $\mathfrak{X}_i <_i \mathfrak{X}$, alors $\langle p , \mathfrak{X} \rangle < \langle p , \omega_i \rangle$.

Le vecteur "p" est appelé *système des prix d' equilibre* , et l' allocation $(\mathfrak{X}_1,\ldots,\mathfrak{X}_n)$ est dite *allocation d' équilibre.* L' ensemble $Bi(p) = \{\mathfrak{X} \in \mathbb{R}^{k,+} ; \langle p , \mathfrak{X} \rangle \leq \langle p , \omega_i \rangle\}$ est appelé *ensemble de budget de l' agent "i"* .

-Maintenant , le problème principal à résoudre est ... *Y a -t- il un équilibre dans l' économie ?* Le problème de l' existence d' un équilibre peut être considéré comme une question d' *optimisation*, mais on doit remarquer que ce n'est pas une optimisation en sens classique , parce qu' on ne sait pas si les préférences $(\preceq_i)_{i=1,2\ldots,n}$ pourront être "évaluées" à l' aide de *représentations d' utilité.* Une *préférence* $\preceq_i$ est représentable s' il existe une fonction (appelée *fonction d' utilité*) $u_i : \mathbb{R}^{k,+} \dashrightarrow \mathbb{R}$ telle que $\mathfrak{X} <_i \mathfrak{Y}$ si et seulement si $u_i(\mathfrak{X}) < u_i(\mathfrak{Y})$. Quand toutes les préférences soient représentables, et les fonctions d' utilité soient connues , on aura un "vrai" problème d' optimisation.

-On dit qu' une allocation $(\mathfrak{X}_1,\ldots,\mathfrak{X}_n)$:

i) Représente une situation *individuellement rationnelle* si $\omega_i \preceq_i \mathfrak{X}_i$ $(i = 1,2,\ldots,n)$.

ii) Est un *optimum de Pareto pour l' économie* $\mathcal{E}$ si elle est réalisable et il n' existe pas une autre allocation réalisable $(\mathfrak{Y}_1,\ldots,\mathfrak{Y}_n)$ avec $\mathfrak{X}_i <_i \omega_i$ $(i = 1,2,\ldots,n)$.

iii) *Appartient au noyau de l' économie* $\mathcal{E}$ si

(a) elle est réalisable ,

(b) $(\mathfrak{X}_1,\ldots,\mathfrak{X}_n)$ n' est pas *bloquée* pour aucune *coalition* d' agents $S \subset I$

(On dit qu' une coalition $S \subset I$ *bloque* $(\mathfrak{X}_1,\ldots,\mathfrak{X}_n)$ s' il existent $\mathfrak{Y}_i \in \mathbb{R}^{k,+}$ $(i \in S)$ tels que $\mathfrak{X}_i <_i \mathfrak{Y}_i$ $(i \in S)$ et $\Sigma_{i\in S}\, \mathfrak{Y}_i = \Sigma_{i\in S}\, \mathfrak{X}_i$).

THÉOREME : (Voir Varian, [1986] , Chap. VI , ou Segura, [1986] , Chap. 7)

Soit $\mathcal{E}$ une économie , et $\mathcal{E} = \{(\mathfrak{X}_1, \ldots, \mathfrak{X}_n , p)\}$ un équilibre en $\mathcal{E}$. Alors , l' allocation $(\mathfrak{X}_1, \ldots, \mathfrak{X}_n)$ représente une situation individuellement rationnelle, est un optimum de Pareto pour $\mathcal{E}$, et appartient au noyau de l' économie.

3. Quelques modèles d' économies infinies .

-Arrivés à ce point , nous allons considérer la possibilité de *généraliser.* *(Dans ce paragraphe et les suivants, nous verrons quelques tentatives de*

justificatition des généralisations qu' on va faire).

D' aprés Ekeland (voir Ekeland, [1979] , p. 16) : << Que le nombre d' agents soit fini n' étonnera personne. Il faudra qu' il soit très grand si l' on veut avoir quelque prétention à décrire la réalité contemporaine. *On peut s' étonner* davantage *que le nombre de biens soit fini , surtout si l' on tient compte du fait qu' ils sont localisés et datés.* >>. Cependant , il ne travaille qu' avec modèles finis. Il rajoute (p. 17) : << De toutes façons , prendre infini le nombre de biens économiques conduirait à des difficultés mathématiques hors de proportion avec le maigre avantage économique qu'on pourrait en rétirer >>. Néanmoins , rajoutons nous, les modèles infinis sont , pour le moment, à la mode, dépuis 1970 . (*Voici quelques réferences* : Hildenbrand, [1970] ; Peleg et Yaari, [1970] ; Bewley, [1972] ; Bojan, [1974] ; Mas-Colell, [1974], [1975] ; Gale et Mas-Colell, [1975],[1979] ; Elbarkuki, [1977] ; Brown et Lewis, [1981] ; Kreps, [1981] ; Magill, [1981] ; Aliprantis et Brown, [1983] ; Florenzano, [1983] ; Jones, [1983] ; Khan et Vohra, [1984] ; Khan, [1984] ; Ostroy, [1984] ; Toussaint, [1984] ; Araujo, [1985a] , [1985b] ; Barbolla, [1985] ; Duffie et Huang, [1985] ; Mas-Colell, [1986] ; Yannelis et Zame, [1986] ; Zame, [1987] ; Besada, Estévez et Hervés, [1988 a] , [1988 b] , [1988c] ; Araujo et Monteiro, [1989] ; Kehoe et al. , [1989] ; Monteiro, [1990] ; Laine, [1991] ; Mas-Colell, [1991] ; Rustichini et Yannelis, [1991])

-Aumann ([1964] et [1966]) est considéré le premier qui prit une économie infinie pour modèle. À son avis, étant donné un système de prix d' équilibre pour une économie $\mathcal{E}$, comme dans le modèle d' Arrow-Debreu , les agents agissent comme si ces prix étaient donnés à l' avance . Cette idée ne satisfait pas à R. Aumann. Pour lui, la notion de compétence parfaite est fondamentale dans l' étude de l' équilibre en économie. L' idée essentielle sous-jacente est que l' économie est composée par un "très grand" nombre de participants, et *l' influence individuelle d'un agent sur l' économie est négligeable*. Il dit que les modèles finis pour étudier l' équilibre ne satisfaient pas cette considération sur les influences individuelles. Alors, il pense que : << Un modèle mathématique d' accord avec la notion intuitive de compétence parfaite doit avoir un *nombre infini de participants* . Et le modèle le plus naturel pour cette considération contient un *continuum de participants*, comme les points d' une ligne droite ou les particles d' un fluide.

La raison mathématique pour prendre un continuum est la possibilité d' utililiser de l' *intégration*. Notamment, si l'on change l' intégrand dans un seul point, cela n' a pas d' influence sur le total. Autrement dit : les actions individuelles d' un agent son négligeables. >> .

[1987], [1989] considérent la possibilité de travailler avec des ensembles "enrichis" du point de vue mathématique. Ils construisent des modèles d' économies avec n agents , dont l' espace des marchandises (E) représente un espace vectoriel topologique avec quelque structure "riche" (espace de Riesz, espace de Fréchet, espace de Banach, ...), et , s' il est possible, avec une rélation d' ordre. En conséquence, il semble intéressant le travail avec des ensembles (les préférences, les agents, les biens , etc.) munis d' une *structure* mathématique. Dans le cas des espaces de Banach, les intégrales qui semblent être les plus convenables sont l' intégrale de Pettis (qu' on appelle aussi "intégrale faible" ou "deuxième intégrale de Dunford") et l' intégrale de Bochner (qu' on appelle aussi "intégrale de Dunford et Schwartz" ou "intégrale première de Dunford") . (Voir Dunford et Schwartz, [1958] ; Diestel et Uhl, [1977] , Diestel, [1984] , Ch. IV , ou bien Lusternik et Sobolev, [1989] , pp. 252 et ss.). Ces deux intégrales coïncident lorsqu' elles coexistent.

-Voici un modèle infini, mais qui met en rélation les cas d ' économies finies, séquentiellement infinies , et continues. Ce modèle a été construit par J. García-Cutrín et C. Hervés , [1991] : Il y a *"n" types différentes d' agents, mais avec un continuum d' agents de chaque type* (cela renferme l' idée de représenter dans le modèle les différentes conduites , en nombre fini "n", que les agents économiques puissent avoir. Par exemple : chaque type d' agent pourrait représenter les lignes essentielles de conduite d' un certain parti politique). Ainsi, on a l ' économie $\mathcal{E}_{n\mathcal{C}}$, avec :

(i) $I = [0,1] = I_1 \cup I_2 \cup ... \cup I_n$, avec $Ij = [\ (j-1)/n\ ,\ j/n\ )$.

(ii) Pour chaque "agent" $t \in Ij$, on a $\omega_t = \omega_j$ et $\preceq_t = \omega_j$. (Autrement dit : Les agents de chaque groupe ont les mêmes ressources et préférences) .

(iii) Les allocations seront des opérateurs (ou fonctions mesurables) $f : [0,1] \dashrightarrow E_+$.

Associée à chaque ensemble de "n" biens $(\mathfrak{X}_1,...,\mathfrak{X}_n)$ on prend une fonction en escalier $f : [0,1] \dashrightarrow E_+$, telle que $g(t) = \mathfrak{X}_j$, pour tout $t \in Ij$.

On appelle intégrale de "f" à $\int_I f = (\mathfrak{X}_1+...+\mathfrak{X}_n) / n \in E_+$, où l' addition $\mathfrak{X}_1+...+\mathfrak{X}_n$ est faite en E_+ , lequel doit être un espace vectoriel topologique (ou, au moins , un cône convexe).

-Les principaux résultats obtenus, par rapport au modèle, sont :

THÉOREME : Soit $\mathcal{E}$ une économie finie avec "n" agents , sur E_+ , et $e = \{ (\mathfrak{X}_1, ... , \mathfrak{X}_n , p)\}$ un équilibre en $\mathcal{E}$ ("p" est un élément qui appartient à E^*

, le dual topologique de E) . Soit "f" la fonction en escalier associée à $(\mathfrak{X}_1, \ldots, \mathfrak{X}_n)$. *Alors , (f , p) représente aussi un équilibre pour l ' économie infinie* $\mathcal{E}_{n\mathcal{C}}$.

THÉOREME : Soit $\mathcal{E}$ *une économie finie avec "n" agents , sur* E_+, *dont les préférences sont convexes et continues. Soit* $k\mathcal{E}$ *l' économie finie obtenue par répétition , "k" fois, de l' économie* $\mathcal{E}$. *Alors, on a que :*

(i) Si f appartient au noyau de l' économie infinie $\mathcal{E}_{n\mathcal{C}}$, *l' allocation* $(\mathfrak{X}_1, \ldots, \mathfrak{X}_n)$, *où* $\mathfrak{X}_j = n \cdot \int_{I_j} f$, *appartient aussi au noyau de l' économie* $\mathcal{E}$.

(ii) Si l' allocation $(\mathfrak{X}_1, \ldots, \mathfrak{X}_n, \ldots^{(k\ fois)} \ldots, \mathfrak{X}_1, \ldots, \mathfrak{X}_n)$ *appartient au noyau de* $k\mathcal{E}$, *pour tout "k" naturel , la fonction en escalier "f" associée à* $(\mathfrak{X}_1, \ldots, \mathfrak{X}_n)$, *appartient aussi au noyau de l' économie infinie* $\mathcal{E}_{n\mathcal{C}}$.

4. Dualité , équilibre , ... et, enfin , les suites .

-Pour une économie $\mathcal{E}$, finie ou non , les systèmes de prix "p" sont pris, en général, comme les éléments d' un espace dual topologique .

Lorsque l' économie est définie sur $\mathbb{R}^{k,+}$, (il y aura "k" biens) , un système de prix , $p \in \mathbb{R}^{k,+}$, peut être considéré comme un élément de l' espace dual $(\mathbb{R}^{k,+})^*$, lequel , bien sûr, coïncide avec $\mathbb{R}^{k,+}$. En fait, étant donnés un système de prix $p = (p_1,\ldots,p_k)$, et un panier de biens $\mathfrak{X} = (x_1,\ldots,x_k)$, la valeur (ou "prix totale") du panier est le produit scalaire $\langle p,\mathfrak{X}\rangle = p_1x_1 + p_2x_2 + \ldots + p_kx_k$. Lorsque l' économie est définie sur un espace différent, E_+ , (au lieu d' être définie sur $\mathbb{R}^{k,+}$) , espace qui est muni , au moins, d' une topologie, la situation se complique, parce que l' espace dual topologique E^* ne coïncide plus (en général) avec E. Mais, il y a des autres problèmes dans cette situation : Par exemple, dans le cas de travailler sur $\mathbb{R}^{k,+}$, quand on considére deux *suites* $(\mathfrak{X}_i)_{i\in\mathbb{N}}$ et $(p_i)_{i\in\mathbb{N}}$ ($\mathfrak{X}_i$, $p_i \in \mathbb{R}^k$, pour tout "i") convergentes , respectivement, aux éléments $\mathfrak{X}$, $p \in \mathbb{R}^{k,+}$, la suite (sur $\mathbb{R}$) définie par $(\langle p_i,\mathfrak{X}_i\rangle)_{i\in\mathbb{N}}$ est, aussi, convergente á $\langle p,\mathfrak{X}\rangle$. Sur d' autres espaces topologiques , la situation est très différente : D' abord, *il faut préciser la topologie des espaces E et E* , pour parler de convergence.* Mais, en outre, on a que : << La convergence , séparément, des suites $(\mathfrak{X}_i)_{i\in\mathbb{N}}$ et $(p_i)_{i\in\mathbb{N}}$ ($\mathfrak{X}_i \in E$, $p_i \in E^*$, pour tout "i") n' implique pas (en général) la convergence de la suite $(\langle p_i,\mathfrak{X}_i\rangle)_{i\in\mathbb{N}}$ >> .

-Voici une *question ouverte fondamentale*, qui met en rélation les suites , la convergence et la dualité (et de laquelle on pourrait dire que l' origine se trouve dans l' Économie) : *Y a -t- il une forme optime de choisir les topologies en E et E* pour "maintenir les convergences" ? (Autrement dit : que la*

-*Le modèle de R. Aumann consiste à* : "k" biens dans l' économie ("continue") $\mathcal{E}_C$. I = [0,1] , muni de la topologie usuelle . Ressources totales : $\{\omega_t \in \mathbb{R}^{k,+} ; t \in [0,1]\}$, ou équivalentement , $\omega : I \dashrightarrow \mathbb{R}^{k,+}$ ($t \mapsto \omega_t$) , *opérateur initial de réssources* (que l' on prendra mesurable) . Pour chaque $t \in [0,1]$ il y a une rélation de préférence $\preceq_t$ (continue et convexe) définie sur $\mathbb{R}^{k,+}$. Une allocation pour ce modèle sera une fonction mesurable $\phi : I \dashrightarrow \mathbb{R}^{k,+}$. Elle sera dite *réalisable* si $\int_I \phi = \int_I \omega \in \mathbb{R}^{k,+}$. Un équilibre sera un couple (f, p) $(f : I \dashrightarrow \mathbb{R}^{k,+} ; p \in \mathbb{R}^{k,+})$ avec $f(t) \in B_t(p)$, ensemble de budget de l' agent "t" , pour chaque $t \in I$. $(B_t(p) = \{\mathfrak{X} \in \mathbb{R}^{k,+} ; <p, \mathfrak{X}> \leq <p, \omega_t>\})$ qui vérifie la condition de maximalité : " si $f(t) \prec_t \mathfrak{X}$, alors $<p, \omega_t> < <p, \mathfrak{X}>$ " .

Ici, le problème principal est de prouver l' existence d' un équilibre. Le problème fut établi par R. Aumann en 1964, et résolu en 1966. Il est un fait remarquable que dans le modèle "continu" de R. Aumann les allocations d' équilibre coïncident avec les allocations du noyau de l' économie $\mathcal{E}_C$. (Voir aussi Bewley, [1973]) .

-W. Hildenbrand , [1970] considére aussi un modèle avec un continuum d' agents, si bien *il trouve une concordance entre les modèles finis et les infinis.* D' abord , il prend des *suites* d ' économies finies , égales deux à deux , $(\mathcal{E}_i)_{i \in \mathbb{N}}$. ($\mathcal{E}_i = \mathcal{E}_1$, pour tout i . $\mathcal{E}_1$ est une économie finie , avec "n" agents et "k" biens). La suite sera ainsi, en fait, une économie *infinie dénombrable* ("séquentielle") $\mathcal{E}_S = \prod_{i \in \mathbb{N}} \mathcal{E}_i$. Il démontre que *une suite d ' économies* égales deux à deux *est convergente* (dans un sens topologique et analytique, que W. Hildenbrand exprime dans son article : Il considére les économies comme mesures sur un certain espace mis en rélation avec les agents, et là , l' économie limite apparaîtra comme la limite d' une suite de mesures) *à une économie avec un continuum d' agents.*

-On trouve un point de vue différent dans le travail de Peleg et Yaari, [1970] . Pour ces auteurs , on doit prendre une économie avec *un nombre fini d' agents , mais avec une infinité dénombrable de biens (dont l' étude sera fait parmi des suites*) , vu que les bien sont localisés et datés.

-Une autre approximation est dûe à T. F. Bewley, [1972] . À son avis, on peut bien prendre une économie avec un nombre fini d' agents, mais *les biens doivent être pris en quantité infinie.* Il considère un *continuum de biens* (un espace de *mesure*) .

-Dans une optique plus modèrne , Aliprantis, Brown et Burkinshaw

convergence , séparément, des suites $(\mathfrak{X}i)_{i\in\mathbb{N}}$ *et* $(pi)_{i\in\mathbb{N}}$ *implique la convergence de la suite* $(\langle pi, \mathfrak{X}i \rangle)_{i\in\mathbb{N}}$)

-Une proprieté, bien connue, de quelques espaces de Banach est la *propriété Dunford-Pettis* (voir Diestel, [1984], p. 113) : << On dit qu' un espace de Banach X satisfait la propriété Dunford-Pettis lorsque, pour tout couple de suites $((\mathfrak{X}i)_{i\in\mathbb{N}}, (\mathfrak{X}^*i)_{i\in\mathbb{N}})$ en X x X* , tel que $(\mathfrak{X}i)_{i\in\mathbb{N}}$ et $(\mathfrak{X}^*i)_{i\in\mathbb{N}}$ sont faiblement convergents à zéro, on a aussi que la suite $(\mathfrak{X}^*i(\mathfrak{X}i))_{i\in\mathbb{N}}$ (dans le corps scalaire , K , associé à X) est convergente à zéro >>. Alors, on peut interpréter la propriété Dunford-Pettis en termes d' une économie sur un espace de Banach , par exemple : << Si l' espace de Banach X vérifie la propriété Dunford-Pettis , alors pour toute économie sur X, toute suite de biens dont la quantité converge à zéro, et toute suite de prix "en déflation" (le prix unitaire de chaque bien converge à zéro) ... le panier limite ne vaudra rien >>. Donc, ici, la question la plus naturelle est : *Quelles sont les conséquences économiques de la propriété Dunford-Pettis ?* Il s' agit d' une question encore ouverte.

-La possibilité de récrire une propriété de l' Analyse Fonctionnelle en termes économiques va plus loin que le seul "change d' encadrement" . D' abord , la question suivante surgit : *Est-ce-que la Théorie Économique sert pour mieux comprendre les concepts de l ' Analyse ?* Les applications économiques serviront pour "nous approcher" des concepts abstraits. Ainsi, il nous semble intéressante la recherche de faits économiques qui réinterprétent des résultats, peut-être bien connus , de l' Analyse Fonctionnelle (en particulier, de la théorie de suites dans les espaces de Banach) . Par exemple , les suites caractérisent la refléxivité d' un espace de Banach (voir Induráin, Plans et Reyes, [1987] , p. 18 .) : << Un espace de Banach X est réflexif si et seulement si toute suite Schauder-basique normalisée est faiblement convergente à zéro >> .

Maintenant , en termes économiques : Cela voudrait dire qu' un espace E est réflexif si et seulement si pour toute suite Schauder-basique $(\mathfrak{X}i)_{i\in\mathbb{N}}$ de quantités "unitaires" (parce que $\|\mathfrak{X}i\| = 1$, pour tout "i" , où $\|.\|$ représente la norme en X) de biens économiques en E, et pour tout système de prix "p" on a que la suite de valeurs $(\langle p,\mathfrak{X}i\rangle)_{i\in\mathbb{N}}$ est convergente à zéro. (Autrement dit : Dans toute suite , Schauder-basique, de biens unitaires, on donne "une plus grande rélévance économique" aux premiers éléments de la suite).

-Dans un espace de Banach séparable et réflexif , on a la suivante caractérisation de la convergence faible de suites (voir Schaeffer , [1974], p. 195 , ou Induráin, Plans et Reyes, [1987], p. 114) : << Soit X un espace de Banach séparable et réflexif, et $(\mathfrak{X}i, \mathfrak{X}i^*)_{i\in\mathbb{N}} \subset X \times X^*$ un système biorthogonal , avec

$(\mathfrak{X}_i)_{i\in\mathbb{N}}$ base de Markousevitch de l'espace X. Soit $(\mathcal{Z}_i)_{i\in\mathbb{N}}$ une suite quelconque dans X. Alors, la suite $(\mathcal{Z}_i)_{i\in\mathbb{N}}$ est faiblement convergente si et seulement si elle est bornée, et, pour tout $j\in\mathbb{N}$ on a que $(\mathfrak{X}_j^*(\mathcal{Z}_i))_{i\in\mathbb{N}}$ converge à une valeur réelle b_j. (Il se suit, en outre, l'existence d'un vecteur $b \in \hat{X}$ avec $\mathfrak{X}_j^*(b) = b_j$, pour tout $j\in\mathbb{N}$). >>

Maintenant, en termes économiques : << Soit X un espace de Banach séparable et réflexif, et $(\mathfrak{X}_i, \mathfrak{X}_i^*)_{i\in\mathbb{N}} \subset X \times X^*$ un système biorthogonal, avec $(\mathfrak{X}_i)_{i\in\mathbb{N}}$ base de Markousevitch de l'espace X. Les éléments de la suite $(\mathfrak{X}_i)_{i\in\mathbb{N}}$ seront des biens économiques, tandis que les éléments de la suite $(\mathfrak{X}_i)_{i\in\mathbb{N}}$ seront des systèmes de prix. Alors, étant donnée une suite de biens $(\mathcal{Z}_i)_{i\in\mathbb{N}}$, pour n'importe quel système de prix "p" on a que $(p(\mathcal{Z}_i))_{i\in\mathbb{N}}$ tend vers p(b), pour un certain $b \in X$ si et seulement si la suite est bornée et pour tout $j \in \mathbb{N}$ on a que $(\mathfrak{X}_j^*(\mathcal{Z}_i))_{i\in\mathbb{N}}$ tend vers une valeur réelle b_j. >> (*Il se suit, en plus, l'existence d'un vecteur de quantités* $b \in X$ *dont la valeur selon les systèmes fondamentaux de prix est déterminée* par $\mathfrak{X}_j^*(b) = b_j$, pour tout $j\in\mathbb{N}$).

-Une caractérisation différente de la réflexivité dans les espaces de Banach séparables munis d'une base de Schauder s'appuie sur le concept de base "*bornément complète*" : Soit X un espace de Banach séparable, et $(\mathfrak{X}_i, \mathfrak{X}_i^*)_{i\in\mathbb{N}} \subset X \times X^*$ un système biorthogonal, avec $(\mathfrak{X}_i)_{i\in\mathbb{N}}$ base de Schauder de l'espace X. On dit qu'une base de Schauder $(\mathfrak{X}_i)_{i\in\mathbb{N}}$ dans un espace X est bornément complète si pour toute suite borné $(\mathcal{Z}_i)_{i\in\mathbb{N}}$ telle que pour tout $f \in [(\mathfrak{X}_i^*)_{i\in\mathbb{N}}]$ (adhérence lineaire de $(\mathfrak{X}_i^*)_{i\in\mathbb{N}}$) la suite $(f(\mathcal{Z}_i))_{i\in\mathbb{N}}$ est convergente, il éxiste un vecteur $\mathcal{Z} \in X$ (indépendamment des "f") tel que $f(\mathcal{Z})$ est la limite de $(f(\mathcal{Z}_i))_{i\in\mathbb{N}}$.

En termes économiques, l'idée sous-jacente est que pour toute suite bornée de biens $(\mathcal{Z}_i)_{i\in\mathbb{N}}$, si pour tout système de prix, f, "dépendant" (c'est à dire : "dans l'adhérence linéaire de ...") de $(\mathfrak{X}_i^*)_{i\in\mathbb{N}}$, $(f(\mathcal{Z}_i))_{i\in\mathbb{N}}$ est convergente, alors, *il existe un bien économique* $\mathcal{Z}$ (indépendamment des "f") *dont la valeur* $f(\mathcal{Z})$ *est la limite* de la suite $(f(\mathcal{Z}_i))_{i\in\mathbb{N}}$.

5. Les suites, la dualité, la myopie, et l'impatience.

-Pour conclure notre exposé, nous allons considérer dans cette dernière section quelques modèles qui opérent aussi avec des suites, et dont les conséquences économiques sont remarquables. Nous montrerons que, pour l'existence d'un équilibre, il faudra que les agents économiques soient, dans un certain sens qu'on précisera, "myopes" ou "impatients".

-Faisons une généralisation du modèle d' Arrow -Debreu . Désormais, une *marchandise* (M) sera un bien économique (B) , localisé dans un certain endroit (L) , qu' on délivre un jour ou date spécifiés d' avance (T), lequel, en outre, peut dépendre des états de la Nature (N) . ($\mathfrak{M} = \mathfrak{B} \times \mathfrak{L} \times \mathfrak{T} \times \mathfrak{N}$) . Par état de la Nature on prend une mesure de probabilité qui réprésente le risque de change brusque , par exemple, dans la situation politique ou économique d' un pays. Ainsi, le nombre de marchandises devient *infini* (on doit rappeler ici que , dans le modèle d'Arrow-Debreu, il était fini), surtout si l' on tient compte des *dates* (T) .

-Les modèles que nous allons présenter dans cette section seront *dénombrables*.

-Prénons $k \in \mathbb{N}$ fixe. Le modèle consiste à :

(i) k biens spécifiés dans la premier période,

(ii) k biens (les mêmes) spécifiés dans la deuxième période ,

(iii) k biens (les mêmes) spécifiés dans la troisième période, ...

Une *spécification* dans l' économie sera une *suite* : $(x_1, x_2, ..., x_k , x_{k+1}, x_{k+2}, ... , x_{2k} , x_{2k+1}, ..., x_{3k}, ...)$, où x_{nk+p} sera le bien p-ième dans la n-ième période. $I = \{1,2,...,n\}$ (ensemble fini d'agents économiques) . Chaque agent "i" est muni d' une dotation initiale de ressources $\omega_i = (\omega_{i,1} ,..., \omega_{i,k}) \in E$, et a définie une relation $\preceq_i$ de préférence individuelle continue et convexe sur E. E sera un sous-ensemble de $\mathbb{R}^{\mathfrak{M}} = \{ f : \mathfrak{M} \text{ -----> } \mathbb{R}, \text{ fonction}\}$.

-Nous étudierons quelques exemples d' espaces que pourraient servir de modèle , et montrerons que le choix de la topologie sur E détermine la conduite des agents : Les modèles les plus fréquentes dans la littérature économique sont l_∞ , $\mathfrak{L}_\infty$, $\mathfrak{C}_0$, $\mathfrak{L}_1$ et $\Lambda(P)$, *espace parfait de Köthe* , voir Köthe, [1951] , [1969] ou Schaeffer, [1974] , p. 198). Voici un exemple (construit par M. Besada , J. García-Cutrín et C. Hervés) , de ces modèles : $I = \{1,2,...,n\}$ (ensemble fini d'agents économiques) . Chaque agent "i" est muni d' une dotation initiale de ressources $\omega_i \in \mathbb{R}^{\mathbb{N}}$ (suites) . Soit $\omega = \omega_1 + ... + \omega_n$ le vecteur des ressources totales de l' économie. Pour nous, un *système de prix* sera une suite $(p_i)_{i \in \mathbb{N}}$ telle que $\Sigma_{i \in \mathbb{N}} |p_i| . |\omega_i| < +\infty$. L' ensemble $P = \{p = (p_i)_{i \in \mathbb{N}} ; \Sigma_{i \in \mathbb{N}} |p_i| . |\omega_i| < +\infty\}$ est un espace vectoriel. Maintenant, soit $\Lambda(P) = \{(\mathfrak{X}_i)_{i \in \mathbb{N}} ; \Sigma_{i \in \mathbb{N}} |p_i| . |\mathfrak{X}_i| < +\infty$, pour tout $p \in P\}$. On peut munir à $\Lambda(P)$ avec une topologie naturelle, avec laquelle, il devient un espace de Köthe. On doit remarquer la *dualité* existente entre P et $\Lambda(P)$. (Par exemple, si $\omega = (1,1,1,1,...)$, alors $P = l_1$, et $\Lambda(P) = l_\infty$) .

-Nous allons introduire ici quelques définitions de concepts topologiques utiles pour notre étude :

1) Brown et Lewis , [1981] appellent *myope* une rélation de préférence " $\precsim$ " définie sur l_∞ , si pour chaque couple $(x,y) \in l_\infty \times l_\infty$, avec $y \prec x$, il existe $n_0 \in \mathbb{N}$ (n_0 sera fonction de (x,y)) tel que $y + (0,\ldots^{n \text{ zéros}}\ldots,0,1,1,1,1,1,\ldots) \prec x$, pour tout $n \geq n_0$. *(En termes économiques la myopie d' une rélation de préférence répond à l' idée d' impatience des agents)* .

2) Étant donnée une suite $\mathfrak{X} = (\mathfrak{X}_i)_{i \in \mathbb{N}}$, on dénotera par $\pi_n(\mathfrak{X})$ la suite $(\mathfrak{X}_1, \ldots, \mathfrak{X}_n, 0,\ldots,0,\ldots)$, et par $S_n(\mathfrak{X})$ la suite $(0,\ldots^{n \text{ zéros}}\ldots, 0, \mathfrak{X}_1, \ldots, \mathfrak{X}_n, \ldots)$.

3) Une rélation de préférence " $\precsim$ " définie sur l_∞ , est *fortement myope* si pour chaque trio $(x,y,z) \in l_\infty \times l_\infty \times l_\infty$, avec $y \prec x$, il existe $n_0 \in \mathbb{N}$ (n_0 sera fonction de (x,y, z)) tel que $y + S_n(z) \prec x$, pour tout $n \geq n_0$.

4) Une rélation de préférence " $\precsim$ " définie sur l_∞ , est *une myopie supérieure* si pour chaque trio $(x,y,z) \in l_\infty \times l_\infty \times l_\infty$, avec $y \prec x$, il existe $n_0 \in \mathbb{N}$ (n_0 sera fonction de (x,y, z)) tel que $y+\pi_n(x) +S_n(z) \prec x$, pour tout $n \geq n_0$.

5) Une topologie sur un ensemble E de $\mathbb{R}^{\mathbb{N}}$ est dite *myope* si toute rélation de préférence continue sur E est myope (la définition de myopie pour les rélations sur E est analogue à celle donnée pour l_∞).

-Voici les principaux résultats sur ce point :

THÉOREME : << Une topologie est fortement myope si et seulement si elle est une myopie supérieure >> .

THÉOREME : << La norme usuelle, en l_∞ , n'est pas myope >> .

THÉOREME : << Une topologie est fortement myope si et seulement si pour tout $\mathfrak{X}$, la suite $(S_i(\mathfrak{X}))_{i \in \mathbb{N}}$, tend vers zéro (dans ladite topologie) >> .

THÉOREME (Brown et Lewis, [1981]) : << La topologie fortement myope, la plus fine, sur l_∞ , est la topologie de Mackey >>

THÉOREME (Araujo, [1985 b]) : << Soit une économie dont les préférences sont continue et convexes, définie sur l_∞ , où la topologie qu'on prend est plus forte que la topologie de Mackey. Alors, il n' y a pas d' équilibre pour l' économie . >>

6. References .

ALIPRANTIS, Ch. D. et D. J. BROWN : "Equilibria in markets with a Riesz space of commodities". Journal of Mathematical Economics 11, 189-207 . (1983).

ALIPRANTIS, Ch. D. , BROWN, D. J. et O. BURKINSHAW : "Edgeworth equilibria". Econometrica 55. (1987).

ALIPRANTIS, Ch. D. , BROWN, D. J. et O. BURKINSHAW : "Existence and optimality of competitive equilibria". Springer Verlag. Berlin. 1989

ARAUJO, A. : "Regular economies and sets of measure zero in Banach spaces". Journal of Mathematical Economics 14, 61-66. (1985 a).

ARAUJO, A. : "Lack of Pareto optima and equilibria for economies with infinitely many goods: The need for impatience". Econometrica 53, 455-462. (1985 b).

ARAUJO, A. et P. K. MONTEIRO : "Equilibrium without uniform conditions". Journal of Economic Theory 48 (2) , 416- 427. (1989).

ARROW, K. : " Social choice and individual value" . John Wiley. New York. 1963.

ARROW, K. et G. DEBREU : "Existence of an equilibrium for a competitive economy". Econometrica 22, 265-290. (1954).

AUMANN, R. J. : "Markets with a continuum of traders". Econometrica 32 (1-2) 39-50. (1964).

AUMANN, R. J. : "Existence of competitive equilibria in markets with a continuum of traders". Econometrica 34, 1- 17. (1966).

BARBOLLA, R. : "Equilibrium and quasi-equilibrium existence theorems for a general model without ordered preferences". Economics Letters 19, 205-210. (1985).

BESADA, M. , ESTÉVEZ , M. et C. HERVÉS : "Una generalización del teorema de Scarf a economías con infinitas mercancías". Revista Española de Economía 5 (1-2). (1988 a).

BESADA, M. , ESTÉVEZ , M. et C. HERVÉS : "Existencia de equilibrio en una economía con producción e infinitas mercancías". Investigaciones Económicas 12 (1) 69-81. (1988 b).

BESADA, M. , ESTÉVEZ , M. et C. HERVÉS : "Equilibria in economies with countably many commodities". Economics Letters 26, 203-207. (1988 c) .

BEWLEY, T. : "Existence of equilibria with infinitely many commodities". Journal of Economic Theory 4, 514-540. (1972).

BEWLEY, T. F. : "Equality of the core and set of equilibria in economies with infinitely many commodities and a continuum of agents" International Economic Review 14. (1973).

BOJAN, P. : "A generalization of theorems on the existence of competitive economic equilibrium in the case of infinitely many commodities". Math. Balkanica 4, 491-494. (1974).

BROWN, D. J. et L. LEWIS : "Myopic economic agents ". Econometrica 49 (2), 359-368. (1981).

DEBREU, G. : "Theory of value : An axiomatic analysis of Economic Equilibrium". Cowles foundation Monograph 17 . Cowles Foundation. Yale University. 1959.

DEBREU, G. : "Mathematical theory of economic equilibrium" . Proceedings of the International Congress of Mathematicians. Vancouver, Canada. 1974.

DIESTEL, J. : "Sequences and series in Banach spaces". Springer Verlag. New York. 1984.

DIESTEL, J. et J. J. UHL Jr. : "Vector measures". American Mathematical Society, Providence, RI. 1977.

DUFFIE, D. et C. HUANG : "Implementing Arrow-Debreu equilibria by continuous tradings of a few long-lived securities". Econometrica 53, 1337-1356. (1985).

DUNFORD, N. et J. T. SCHWARTZ : "Linear Operators I". John Wiley. New York. 1958.

EKELAND, I. : " Élements d' Économie Mathématique". Hermann (Collection Méthodes) . Paris. 1979.

ELBARKUKI, R.A. : " The existence of an equilibrium in economic stuctures with a Banach space of commodities". Akad. Nauk. Azerbaijan SSR Doklady 33, 8-12 . (1977).

FLORENZANO, M . : "On the existence of equilibria in economies with an infinite dimensional commodity space". Journal of Mathematical Economics 12, 207-220. (1983).

GALE, D. et A. MAS-COLELL : "An equilibrium existence theorem for a general model without ordered preferences". Journal of Mathematical Economics 2, 9-15. (1975).

GALE, D. et A. MAS-COLELL : "Correction to an equilibrium existence theorem for a general model without ordered preferences". Journal of Mathematical Economics 6, 297-298. (1979).

GARCIA-CUTRIN, J. et C. HERVÉS : "Un modelo de economía con núcleo no vacío" (Prépublication) . Universidad de Vigo. (1991).

HILDENBRAND, W. : "On economies with many agents". Journal of Economic Theory 2, 161-188. (1970).

INDURAIN, E. , PLANS, A. et A. REYES : "Notas sobre geometría de sucesiones en espacios de Banach" . Monografías del Departamento de Matemáticas. Universidad de Extremadura. Badajoz. 1987.

JONES, L. : "Existence of equilibria with infinitely many consumers and infinitely many commodities" Journal of Mathematical Economics 12, 119-138. (1983).

KHAN, M. Ali : "A remark on the existence of equilibria in markets without ordered preferences and with a Riesz space of commodities". Journal of Mathematical Economics 13, 165-169. (1984).

KHAN, M. Ali et R. VOHRA : "Equilibrium in abstract economies without ordered preferences and with a measure space of agents". Journal of Mathematical Economics 13, 133-142. (1984).

KEHOE, T. J. , LEVINE , D. K. , MAS-COLELL, A. et ZAME, W. R. : "Determinacy of equilibrium in large-scale economies". Journal of Mathematical Economics 18, 231-262. (1989).

KÖTHE , G . : "Neubegründung der Theorie der vollkommenen Raüme". Mathematische Nachrichten 4, 70-80. (1951).

KÖTHE, G. : "Topological vector spaces I". Springer Verlag, New York. 1969.

KREPS, D. M. : "Arbitrage and equilibrium in economies with infinitely many commodities" . Journal of Mathematical Economics 8, 15-35. (1981).

LAINE, J. : "Finitely optimal allocation in countably large-square exchange economies". Economics Letters 35 , 1-4. (1991).

LUSTERNIK, L. et V. SOBOLEV : " Précis d' analyse fonctionnelle". Mir. Moscou. 1979.

MAGILL, M. : "An equilibrium existence theorem". Journal of Mathematical Analysis and Applications 84 (1), 162-169. (1981).

MAS-COLELL, A. : "An equilibrium existence theorem without complete or transitive preferences". Journal of Mathematical Economics 1, 237-246. (1974).

MAS-COLELL, A. : "Model of equilibrium with differentiated commodities". Journal of Mathematical Economics 2, 263-295. (1975).

MAS-COLELL, A. : "The price equilibrium existence problem in topological vector lattices". Econometrica 54, 1039- 1054. (1986).

MAS-COLELL, A. : "Comments to session on Infinite Dimensional General Equilibrium", to be published in "Advances in Economic theory, 5th World Congress" (J.J. Laffont, ed.) Cambidge University Press. (1991).

MONTEIRO, P. K. : "The decomposition of excess demand functions on Banach spaces". (Preprint). Communicated in the 6th World Congress of the Econometric Society. Barcelona (Spain), August 1990.

OSTROY, J. M. : "On the existence of Walrasian equilibrium in large-square economies". Journal of Mathematical Economics 13, 143-163. (1984).

PELEG, B. et M. YAARI : "Markets with countably many commoditites". International Economic Review 11, 369-377. (1970).

RUSTICHINI, A. et N. C. YANNELIS : " Edgeworth' s conjecture in economies with a continuum of agents and commodities" . Journal of Mathematical Economics 20, 307-326. (1991).

SCHAEFFER, H. H. : "Espacios vectoriales topológicos". Teide. Barcelona. 1974.

SEGURA, J. : "Análisis microeconómico". Alianza Editorial. Madrid. 1986.

TOUSSAINT, S. : "On the existence of equilibria in economies with infinitely many commodities and ordered preferences". Journal of Economic Theory 33, 98-115. (1984).

VARIAN, H. R. : "Analisis Microeconómico". Bosch. Barcelona. 1986.

YANNELIS, N. C. et ZAME, W. : "Equilibria in Banach lattices without ordered preferences". Journal of Mathematical Economics 15, 85-110. (1986).

ZAME, W. : "Competitive equilibria in production economies with an infinite-dimensional commodity space". Econometrica 55, 1075-1108. (1987).

Rev. Academia de Ciencias. Zaragoza **47** *(1992)*

A Generalized "K-Attainable" Inequality Related to the Weak Convergence of Probability Measures

by

George A. Anastassiou

Department of Mathematical Sciences
Memphis State University
Memphis, TN 38152, U.S.A.

1980 Subject Classification (1985 Revision): Primary 41A17, 41A25, 41A50, Secondary 41A80, 60E15.

Keywords and phrases: convexity, degree of approximation, extended complete Tchebycheff system, generalized polynomial, K-functional, probability measure, rate of weak convergence, sharp inequality.

ABSTRACT

The rate of weak convergence of a sequence of probability measures to the unit measure at a fixed point is given by a sharp generalized inequality in the presence of an extended complete Tchebycheff system. This involves a very flexible new measure of smoothness, the generalized reduced K-functional, which has all the basic properties of the usual K-functional.

1. INTRODUCTION

The following introductory notions come from [5], which is an essential reference for this article.

Let the functions $g, u_0, u_1, \ldots, u_n \in C^{n+1}([a,b])$, $n \geq 0$, $[a,b] \subset \mathbb{R}$, such that $u_0(x) = c > 0$ and $u_1(x)$ is a concave function for $x \leq t$ and a convex function for $x \geq t$, where t is a fixed point in (a,b).

Let the Wronskians

$$W_i(x) = W(u_0(x),u_1(x), \ldots, u_i(x)) = \begin{vmatrix} u_0(x) & u_1(x) & \cdots & u_i(x) \\ u_0'(x) & u_1'(x) & \cdots & u_i(x) \\ \vdots & & & \\ u_0^{(i)}(x) & u_1^{(i)}(x) & \cdots & u_i^{(i)}(x) \end{vmatrix},$$

$$i = 0,1, \ldots, n$$

and assume that all $W_i(x)$ are positive everywhere on [a,b]. Obviously the functions

$$\phi_0(x) = W_0(x) = u_0(x), \quad \phi_1(x) = \frac{W_1(x)}{(W_0(x))^2},$$

$$\phi_i(x) = \frac{W_i(x)W_{i-2}(x)}{(W_{i-1}(x))^2}, \quad i = 2,3, \ldots, n$$

are positive everywhere on [a,b].

Consider the linear differential operator of order $i \geq 0$

$$L_i g(x) = \frac{W[u_0(x),u_1(x), \ldots, u_{i-1}(x),g(x)]}{W_{i-1}(x)} \tag{1}$$

$i = 1, \ldots, n+1$; $L_0 g(x) = g(x)$ all $x \in [a,b]$. Here $W[u_0(x),u_1(x), \ldots, u_{i-1}(x),g(x)]$ denotes the Wronskian of $u_0,u_1, \ldots, u_{i-1},g$.

Consider also the functions

$$g_i(x,t) = \frac{1}{W_i(x)} \cdot \begin{vmatrix} u_0(t) & u_1(t) & \cdots & u_i(t) \\ u_0'(t) & u_1'(t) & \cdots & u_i'(t) \\ \vdots & & & \\ u_0^{(i-1)}(t) & u_1^{(i-1)}(t) & \cdots & u_i^{(i-1)}(t) \\ u_0(x) & u_1(x) & \cdots & u_i(x) \end{vmatrix} \tag{2}$$

$i = 1,2, \ldots, n$; $g_0(x,t) = 1$, all $x,t \in [a,b]$. Note that $g_i(x,t)$ as a function of x, it is a linear combination of $u_0(x), u_1(x), \ldots, u_i(x)$ and furthermore it holds

$$g_i(x,t) = \frac{1}{\phi_1(t)\cdots\phi_i(t)} \int_t^x \phi_1(x_1) \int_t^{x_1} \cdots \int_t^{x_{i-2}} \phi_{i-1}(x_{i-1}) \int_t^{x_{i-1}} \phi_i(x_i) dx_i dx_{i-1} \cdots dx_1,$$

all $i = 1,2, \ldots, n$.

From [5], p. 138, Theorem II we have that

$$g(x) = g(t) + \sum_{i=1}^{n} L_i g(t) \cdot g_i(x,t) + R_n(x), \tag{3}$$

where

$$R_n(x) = L_{n+1} g(\xi) \cdot \int_t^x g_n(x,s) ds,$$

ξ is between x,t; all $x \in [a,b]$, t is a fixed point in (a,b), $n \geq 0$.

Here we shall call

$$\begin{aligned} N_n(x,t) &= \int_t^x g_n(x,s) ds, \\ \tilde{N}_n(x,t) &= |N_n(x,t)|; \quad n \geq 0. \end{aligned} \tag{4}$$

Our work is mainly motivated by the following result (see [2], p. 376).

Theorem. Let $u_0, u_1, \ldots, u_n \in C^n([a,b])$, $n \geq 0$. Then $\{u_i\}_{i=0}^n$ is an extended complete Tchebycheff (E.C.T.) system on $[a,b]$ iff $W_i(x)$ are positive everywhere on $[a,b]$, all $i = 0,1, \ldots, n$.

Next we would like to introduce the following new measure of smoothness, which plays a central role in this article.

Definition 1. Let $f \in C([a,b])$. Let $n \geq 0$ and $t \in (a,b)$ be fixed. We define the generalized reduced K-functional as:

$$\tilde{K}_{n+1}^{(t)}(f,h) := K(f,h) := K(h) = \inf_g (\|f-g\|_\infty + h\|L_{n+1} g\|_\infty), \quad h \geq 0. \tag{5}$$

Here g ranges through the functions $g \in C^{n+1}([a,b])$ satisfying

$$L_i g(t) = 0, \quad i = 1, \ldots, n$$

where $\|\cdot\|_\infty$ denotes the sup norm.

The notion of K(f,h) generalizes to a wide class of E.C.T. systems the notion of reduced K-functional (see [1]), where there we have the special case of $u_i(x) = x^i$, $i = 0,1, \ldots, n$. E.g. consider any E.C.T. system in $C^{n+1}([a,b])$ containing $u_0(x) = c > 0$, $u_1(x) = \sinh x$; where $t = o \in (a,b)$. And K(f,h) is similar to the K-functional of Peetre (see [3] ; [4] , p. 60), where we do not impose the restrictions $L_i g(t) = 0$, $i = 1, \ldots, n$; there is

$$L_{n+1} = \frac{d^{n+1}}{dx^{n+1}} .$$

The quantity K(f,h) measures how well f can be approximated by a smooth function g. It is the suitable, very flexible measure for the approximation of small functions f with large derivatives in the sup norm.

In Theorem 1 we present the basic properties of the introduced generalized reduced K-functional, proving that it behaves nicely like the other K-functionals.

In Theorem 2 we present a generalized sharp inequality involving $\tilde{K}^{(t)}_{n+1}(f,h)$, which is the main goal of this article. This inequality gives an estimate in the very general E.C.T. setting for the rate of weak convergence of a sequence of probability measures to the unit measure at a fixed point.

Finally, Theorem 3 contains an independent result in the uniform approximation of continuous functions by generalized polynomials.

2. MAIN RESULTS

Theorem 1. The generalized reduced K-functional has the following properties: (i) Subadditive in terms of f. And in terms of h is (ii) continuous, (iii) nonnegative, (iv) monotonely increasing, (v) concave, hence (vi) subadditive. (vii) $\tilde{K}^{(t)}_{n+1}(f,0) = 0$.

Remark 1. Thus $\tilde{K}^{(t)}_{n+1}(f,h)$ has all the basic properties of the usual Peetre functional $K_{n+1}(f,h)$. In particular $\tilde{K}^{(t)}_1 = K^{(t)}_1 = K_1$, by $u_0(x)$ constant and $L_1 = d/dx$.

Also, since $L_i c = 0$, $c \in \mathbb{R}$, all $i \geq 1$ we get that $\tilde{K}^{(t)}_{n+1}(c,h) = 0$, all $h \geq 0$.

Proof of Theorem 1. (i) Let $g_1, g_2 \in C^{n+1}([a,b])$ such that $L_i g_1(t) = L_i g_2(t) = 0$, $i = 1, \ldots, n$, i.e., $L_i(g_1 + g_2)(t) = 0$, $i = 1, \ldots, n$. Hence

$$K(f_1 + f_2, h) \leq \|(f_1 + f_2) - (g_1 + g_2)\|_\infty + h\|L_{n+1}(g_1 + g_2)\|_\infty$$
$$\leq (\|f_1 - g_1\|_\infty + h\|L_{n+1} g_1\|_\infty) + (\|f_2 - g_2\|_\infty + h\|L_{n+1} g_2\|_\infty),$$

for all g_1, g_2 of the above class. Obviously one has

$$K(f_1 + f_2, h) \leq K(f_1, h) + K(f_2, h).$$

(ii) (a) K is an upper-semicontinuous function in h i.e.,

$$\limsup_{h \to c} K(h) \leq K(c).$$

Proof. Observe that

$$\limsup_{h \to c} K(h) \leq \limsup_{h \to c} (\|f - g\|_\infty + h\|L_{n+1} g\|_\infty)$$
$$= \lim_{h \to c} (\|f - g\|_\infty + h\|L_{n+1} g\|_\infty) = (\|f - g\|_\infty + c\|L_{n+1} g\|_\infty),$$

for all $g \in C^{n+1}([a,b])$ satisfying $L_i g(t) = 0$, $i = 1, \ldots, n$. The result is now obvious.

(b) K is right continuous in h.

Proof. Since K is increasing in h we have $K(h) \leq K(h+\epsilon)$, $\epsilon > 0$. From (a) for $\epsilon \to 0$ we get

$$\limsup_{h+\epsilon \to h} K(h+\epsilon) \leq K(h).$$

Furthermore it holds

$$K(h) \leq \liminf_{h+\epsilon \to h} K(h+\epsilon)$$

along with

$$\liminf_{h+\epsilon \to h} K(h+\epsilon) \leq \limsup_{h+\epsilon \to h} K(h+\epsilon).$$

Hence

$$\liminf_{h+\epsilon \to h} K(h+\epsilon) = \limsup_{h+\epsilon \to h} K(h+\epsilon).$$

$$= \lim_{h+\epsilon \to h} K(h+\epsilon) = K(h),$$

which proves the claim.

(c) It holds $K(h-) \leq K(h) \leq K(h-)$, i.e., K is left continuous in h.

Proof. Obviously $K(h-) \leq K(h)$, by K being increasing in h. If $0 < \epsilon < h$ then $h/h-\epsilon > 1$. Now see that

$$K(h) \leq (\|f-g\|_\infty + h\|L_{n+1}g\|_\infty) \leq \left[\frac{h}{h-\epsilon}\right] \cdot (\|f-g\|_\infty + (h-\epsilon)\|L_{n+1}g\|_\infty),$$

for all $g \in C^{n+1}([a,b])$ such that $L_i g(t) = 0$, $i = 1, \ldots, n$. Thus

$$K(h) \leq \left[\frac{h}{h-\epsilon}\right] K(h-\epsilon),$$

i.e.,

$$K(h) \leq K(h-),$$

which proves the claim.

We have established that K is continuous in h.

(iii) obvious.

(iv) For $h_1 \leq h_2$ we get

$$K(h_1) \leq (\|f-g\|_\infty + h_1\|L_{n+1}g\|_\infty) \leq (\|f-g\|_\infty + h_2\|L_{n+1}g\|_\infty),$$

for all $g \in C^{n+1}([a,b])$ such that $L_i g(t) = 0$, $i = 1, \ldots, n$. Thus $K(h_1) \leq K(h_2)$.

(v) Let $0 \leq \lambda \leq 1$. We observe that

$$\lambda K(h_1) + (1-\lambda)K(h_2) \leq \lambda(\|f-g\|_\infty + h_1\|L_{n+1}g\|_\infty) + (1-\lambda)(\|f-g\|_\infty + h_2\|L_{n+1}g\|_\infty)$$
$$= \|f-g\|_\infty + (\lambda h_1 + (1-\lambda)h_2)\|L_{n+1}g\|_\infty,$$

for all $g \in C^{n+1}([a,b])$ such that $L_i g(t) = 0$, $i = 1, \ldots, n$. Therefore

$$\lambda K(h_1) + (1-\lambda)K(h_2) \leq K(\lambda h_1 + (1-\lambda)h_2),$$

proving the claim.

(vii) From Lemmas 2,3 we get that

$$\tilde{g}(x) := N_n(x,t) = \int_t^x g_n(x,s)ds \in C^{n+1}([a,b]), \quad n \geq 0,$$

is a strictly monotone function.

Thus for $x \neq y$ we get that $\tilde{g}(x) \neq \tilde{g}(y)$, that is, $\tilde{g}(x)$ separates points. Furthermore from [5], p. 138 we have that

$$L_{n+1}\tilde{g}(x) = 1,$$
$$\tilde{g}^{(i)}(t) = 0, \quad i = 0,1,2, \ldots, n.$$

So that $L_i\tilde{g}(t) = 0$, $i = 1, \ldots, n$.

Consider now the algebra A generated by $\{1, \tilde{g}(x)\}$ i.e., $A = \langle\{1, \tilde{g}(x)\}\rangle$. A typical element of A is a finite linear combination of non-negative powers of $\tilde{g}$. Note that A is a linear subspace of $C^{n+1}([a,b]) \subset C([a,b])$, which contains the constant functions and separates points. Therefore by the Stone-

Weierstrass Theorem we obtain that

$$\bar{A} = C([a,b])$$

i.e., for $f \in C([a,b])$: $\forall \epsilon > 0$ there exists $e \in A$ such that $\|f - e\|_\infty < \epsilon$. Since $u_0(x)$ is constant we get that $L_i(1) = 0$, all $i = 1, \ldots, n$.

And since $\tilde{g}^{(i)}(t) = 0$, all $i = 0,1,2, \ldots, n$ we have that $(\tilde{g}^{\ell})^{(i)}(t) = 0$, $i = 0,1,2, \ldots, n$ ($\ell \geq 1$). That is, $L_i(\tilde{g}^{\ell})(t) = 0$, $i = 1, \ldots, n$. These are leading to $L_i e(t) = 0$, $i = 1, \ldots, n$ for all $e \in A$. Hence A is contained in the class of functions $g \in C^{n+1}([a,b])$ such that

$$L_i g(t) = 0, \quad i = 1, \ldots, n.$$

Consequently,

$$0 \leq \hat{K}^{(t)}_{n+1}(f,0) = \inf_{g} (\|f - g\|_\infty) \leq \inf_{e \in A} (\|f - e\|_\infty) = 0.$$

We have established that

$$\hat{K}^{(t)}_{n+1}(f,0) = 0. \qquad \square$$

Theorem 2. Let μ be a probability measure on $[a,b] \subset \mathbb{R}$ such that

$$\int_{[a,b]} |x - t| \mu(dx) = d_1(t) > 0, \tag{6}$$

where t is a fixed point of (a,b).

Let the functions $u_0(x), u_1(x), \ldots, u_n(x)$ belong to $C^{n+1}([a,b])$, $n \geq 0$, and let the Wronskians $W_0(x), W_1(x), \ldots, W_n(x)$ be positive throughout [a,b].

Assume that $u_0(x) = c > 0$ and $u_1(x)$ is a concave function for $x \leq t$ and a convex function for $x \geq t$.

Call

$$\bigwedge_n(t) = \max\left\{ \frac{\hat{N}_n(b,t)}{b - t}, \frac{\hat{N}_n(a,t)}{t - a} \right\}, \tag{7}$$

where

$$\hat{N}_n(x,t) = \left| \int_t^x g_n(x,s)ds \right|, \quad \text{all } x \in [a,b]; \ n \geq 0.$$

Consider $f \in C([a,b])$. Then

$$\left| \int_{[a,b]} f \, d\mu - f(t) \right| \leq 2\hat{K}_{n+1}^{(t)}\left[f; \frac{\Lambda_n(t) \cdot d_1(t)}{2}\right], \quad n \geq 0. \tag{8}$$

For special choices of $d_1(t)$ the above inequality is *attained* (*i.e., it is sharp*) by the function $N_n(x,t)$ and the probability measure μ_0 described as follows:

(i) If $\dfrac{\hat{N}_n(b,t)}{b-t} \geq \dfrac{\hat{N}_n(a,t)}{t-a}$ and $d_1(t) \leq b - t$, then the optimal probability measure μ_0 is supported at $\{t,b\}$ with masses $\left[1 - \dfrac{d_1(t)}{b-t}\right]$, $\left[\dfrac{d_1(t)}{b-t}\right]$, respectively.

(ii) If $\dfrac{\hat{N}_n(b,t)}{b-t} \leq \dfrac{\hat{N}_n(a,t)}{t-a}$ and $d_1(t) \leq t - a$, then the optimal probability measure μ_0 is supported at $\{t,a\}$ with masses $\left[1 - \dfrac{d_1(t)}{t-a}\right]$, $\left[\dfrac{d_1(t)}{t-a}\right]$, respectively.

Remark 2. According to [5], p. 139 when $u_i(x) = x^i$, $i = 0,1,2, \ldots, n$ we have that

$$N_n(x,t) = \frac{(x-t)^{n+1}}{(n+1)!} \quad \text{and} \quad \hat{N}_n(x,t) = \frac{|x-t|^{n+1}}{(n+1)!}$$

Thus

$$\hat{N}_n(b,t) = \frac{(b-t)^{n+1}}{(n+1)!}, \quad \hat{N}_n(a,t) = \frac{(t-a)^{n+1}}{(n+1)!}$$

and

$$\Lambda_n(t) = \max\left\{ \frac{(b-t)^n}{(n+1)!}, \frac{(t-a)^n}{(n+1)!} \right\}.$$

That is,

$$\Lambda_n(t) = \frac{(c(t))^n}{(n+1)!},$$

where $c(t) = \max(b-t, t-a)$; $n \geq 0$. Here note that $\hat{K}_{n+1}^{(t)} = K_{n+1}^{(t)}$: the reduced

K-functional introduced in [1]. In this special case inequality (8) reduces to the known attained inequality (see [1]),

$$\left|\int_{[a,b]} f\,d\mu - f(t)\right| \leq 2\cdot K_{n+1}^{(t)}\left[f;\ \frac{d_1(t)\cdot(c(t))^n}{2(n+1)!}\right],$$

$t \in (a,b)$; $n \geq 0$.

Proof of Theorem 2. For $f \in C([a,b])$ and $g \in C^{n+1}([a,b])$ we have $f(x) - f(t) = (f(x) - g(x)) + (g(x) - g(t)) + (g(t) - f(t))$. Integrating relative to μ we obtain

$$\begin{aligned}\left|\int_{[a,b]} f\,d\mu - f(t)\right| &\leq \int_{[a,b]} |f(x) - g(x)|\mu(dx)\\ &\quad + |g(t) - f(t)| + \int_{[a,b]} |g(x) - g(t)|\mu(dx)\\ &\leq 2\|f - g\|_\infty + \int_{[a,b]} |g(x) - g(t)|\mu(dx).\end{aligned}$$

Now assume $L_i g(t) = 0$ for $i = 1, \ldots, n$, then from (3) we get

$$g(x) - g(t) = L_{n+1}g(\xi)\cdot\int_t^x g_n(x,s)ds,$$

where ξ is between x,t. Hence $|g(x) - g(t)| \leq \|L_{n+1}g\|_\infty\cdot\tilde{N}_n(x,t)$. And from Lemma 5 we obtain

$$|g(x) - g(t)| \leq \|L_{n+1}g\|_\infty\cdot\max\left\{\frac{\tilde{N}_n(b,t)}{b-t},\ \frac{\tilde{N}_n(a,t)}{t-a}\right\}\cdot|x-t| \text{ for all } x \in [a,b].$$

Consequently,

$$\left|\int_{[a,b]} f\,d\mu - f(t)\right| \leq 2\|f-g\|_\infty + \|L_{n+1}g\|_\infty\cdot\Lambda_n(t)\cdot d_1(t),$$

where $\Lambda_n(t)$, $d_1(t)$ as in (7), (6), respectively. Therefore we have proved (8)

$$\left|\int_{[a,b]} f\,d\mu - f(t)\right| \leq 2\cdot\tilde{K}_{n+1}^{(t)}\left[f;\ \frac{\Lambda_n(t)\cdot d_1(t)}{2}\right].$$

Furthermore, consider $\tilde{f}(x) = N_n(x,t) \in C^{n+1}([a,b])$. This satisfies $L_{n+1}\tilde{f}(x) = 1$ with $\tilde{f}^{(i)}(t) = 0$, $i = 0,1,2, \ldots, n$, so that $L_i\tilde{f}(t) = 0$, $i = 1, \ldots, n$. Hence

$$\hat{K}_{n+1}^{(t)}(\hat{f},h) \leq h\|L_{n+1}\hat{f}\|_\infty = h.$$

That is,

$$2\cdot\hat{K}_{n+1}^{(t)}\left(\hat{f};\frac{\Lambda_n(t)\cdot d_1(t)}{2}\right) \leq \Lambda_n(t)\cdot d_1(t).$$

I.e., inequality (8) for $f = \hat{f}$ leads to

$$\left|\int_{[a,b]} \hat{f}\, d\mu\right| \leq \Lambda_n(t)\cdot d_1(t). \tag{9}$$

Here we see that

$$\text{(i)}\quad \left|\int_{[a,b]} \hat{f}\, d\mu_0\right| = \left|\int_{[a,b]} N_n(x,t)\mu_0(dx)\right|$$

$$= \left(\frac{N_n(b,t)}{b-t}\right)\cdot d_1(t) = \left(\frac{\hat{N}_n(b,t)}{b-t}\right)\cdot d_1(t)$$

$$= \Lambda_n(t)\cdot d_1(t),$$

i.e., inequality (9) is attained.

$$\text{(ii)}\quad \left|\int_{[a,b]} \hat{f}\, d\mu_0\right| = \left|\left(\frac{N_n(a,t)}{t-a}\right)\cdot d_1(t)\right|$$

$$= \left(\frac{\hat{N}_n(a,t)}{t-a}\right)\cdot d_1(t) = \Lambda_n(t)\cdot d_1(t),$$

i.e., inequality (9) is again attained.

We have proved in both cases that inequality (8) is attained. □

From Lemmas 2, 3 and the proof of Theorem 1(vii) we get the following independent results.

Theorem 3. Assume that $u_0(x) = c > 0$ and $u_1(x)$ is a concave function for $x \leq t$ and a convex function for $x \geq t$; $x,t \in [a,b]$, t is a fixed point. Let

$$\varphi_n(x) := N_n(x,t) = \int_t^x g_n(x,s)ds, \quad n \geq 0.$$

Consider the algebra $A_n = \langle\{1, \varphi_n(x)\}\rangle$, which is a linear subspace of $C^{n+1}([a,b])$. Then $\bar{A}_n = C([a,b])$, $n \geq 0$, in the uniform norm. Furthermore

$$L_i e(t) = 0, \quad i = 1, \ldots, n \quad \text{for all} \quad e \in A_n;\ n \geq 1.$$

3. AUXILIARY RESULTS

The next results are by themselves independently interesting.

Lemma 1. Let

$$N_n(x,t) = \int_t^x g_n(x,s)ds, \quad \text{all} \quad x,t \in [a,b].$$

Then

$$\frac{\partial N_n(x,t)}{\partial x} = \int_t^x \frac{\partial g_n(x,s)}{\partial x}\, ds, \quad n \geq 1 \tag{10}$$

and

$$\frac{\partial^2 N_n(x,t)}{\partial x^2} = \int_t^x \frac{\partial^2 g_n(x,s)}{\partial x^2} ds, \quad n \geq 2. \tag{11}$$

Proof. See [5], p. 132(6) and apply Leibnitz's formula once/twice. □

The previous result is used in

Lemma 2. Assume that $u_0(x) = c > 0$ and $u_1(x)$ is a convex function for $x \geq t$. Let

$$N_n(x,t) = \int_t^x g_n(x,s)ds, \quad \text{all} \quad x,t \in [a,b], \quad n \geq 0.$$

Then $N_n(x,t) > 0$ for $x > t$, $N_n(t,t) = 0$, and it is a strictly increasing function in $x \geq t$, also $N_n(x,t)$ is a continuous function in $x \in [a,b]$.

Furthermore, $N_n(x,t)$ is a strictly convex function in $x \geq t$, $n \geq 1$, and $N_0(x,t) = x - t$ is a trivially convex function in $x \geq t$.

Proof. From $W_0(x) = \phi_0(x) = u_0(x) = c > 0$ and $W_1(x) = cu_1'(x) > 0$, $u_1(x)$ is a strictly increasing function everywhere on [a,b]. Hence $\phi_1(x) = W_1(x)/(W_0(x))^2 = u_1'(x)/c > 0$. By assumption $u_1(x)$ is a convex function in $x \geq t$ implying that $u_1'(x)$ is an increasing function there, that is, $\phi_1(x)$ is increasing in $x \geq t$.

Note that $g_n(x,t) > 0$ $(x > t)$, $g_n(t,t) = 0$; $n \geq 1$ with $g_0(x,t) = 1$. We have

$$\frac{\partial g_n(x,t)}{\partial x} = \frac{\phi_1(x)}{\phi_1(t)\cdots\phi_n(t)} \int_t^x \phi_2(x_1)\cdots\int_t^{x_{n-2}} \phi_n(x_{n-1})dx_{n-1}\cdots dx_1.$$

From $\phi_i(x) > 0$, $i = 1, \ldots, n$, $n \geq 2$ and $\phi_1(x)$ being an increasing function we have that $\partial g_n(x,t)/\partial x$ is a strictly increasing function in $x \geq t$, note that $\partial g_n(x,t)/\partial x > 0$ $(x > t)$, $\partial g_n(t,t)/\partial x = 0$.

Thus $g_n(x,t)$ is a strictly convex function in $x \geq t$, $n \geq 2$ and clearly $g_1(x,t)$ is convex in $x \geq t$.

Obviously $N_n(x,t)$ is a continuous function in $x \in [a,b]$, $n \geq 0$.

From Lemma 1

$$\frac{\partial^i N_n(x,t)}{\partial x^i} = \int_t^x \frac{\partial^i g_n(x,s)}{\partial x^i} ds, \quad (x \geq t,\ n \geq 2),\ i = 1,2.$$

It is clear that $N_n(x,t)$ is a strictly increasing function in $x \geq t$, $n \geq 2$.

By strict convexity of $g_n(x,s)$ in $x \geq s$ we get $\partial^2 g_n(x,s)/\partial x^2 > 0$ $(x > s)$, which leads to

$$\frac{\partial^2 N_n(x,t)}{\partial x^2} > 0 \quad (x > t), \quad \frac{\partial^2 N_n(t,t)}{\partial x^2} = 0.$$

Hence $N_n(x,t)$ is a strictly convex function in $x \geq t$, $n \geq 2$.

Since $g_0(x,t) = 1$, all $x,t \in [a,b]$, $N_0(x,t) = x - t$, $(x \geq t)$. Obviously $N_0(x,t)$ is a trivially convex and strictly increasing function in $x \geq t$. Note that

$$g_1(x,t) = \phi_1^{-1}(t)\int_t^x \phi_1(s)ds.$$

From $\partial g_1(x,t)/\partial x = \phi_1(x)/\phi_1(t)$ and ϕ_1 an increasing function, we have that $\partial g_1(x,t)/\partial x$ is increasing in $x \geq t$. Obviously $\partial g_1(x,t)/\partial x > 0$ for all $x \in [a,b]$.

Let s be such that $t \leq s \leq x_1 < x_2$, then

$$\frac{\partial g_1(x_2,s)}{\partial x} \geq \frac{\partial g_1(x_1,s)}{\partial x}$$

Adding

$$\int_t^{x_1} \frac{\partial g_1(x_2,s)}{\partial x} ds \geq \int_t^{x_1} \frac{\partial g_1(x_1,s)}{\partial x} ds$$

and

$$\int_{x_1}^{x_2} \frac{\partial g_1(x_2,s)}{\partial x}\, ds > 0,$$

one has

$$\int_t^{x_2} \frac{\partial g_1(x_2,s)}{\partial x}\, ds > \int_t^{x_1} \frac{\partial g_1(x_1,s)}{\partial x}\, ds .$$

The last inequality and Lemma 1 (10) imply that $\partial N_1(x,t)/\partial x$ is strictly increasing in $x \geq t$, which in turn implies that $N_1(x,t)$ is a strictly convex function in $x \geq t$. Since

$$\frac{\partial N_1(x,t)}{\partial x} > 0 \quad (x > t), \qquad \frac{\partial N_1(t,t)}{\partial x} = 0$$

we conclude that $N_1(x,t)$ is a strictly increasing function in $x \geq t$. □

The counterpart of Lemma 2 has as follows:

Lemma 3. Assume that $u_0(x) = c > 0$ and $u_1(x)$ is a concave function for $x \leq t$. Let

$$N_n(x,t) = \int_t^x g_n(x,s)ds, \quad \text{all} \quad x,t \in [a,b]: \quad x \leq t, \quad n \geq 0.$$

If n is odd, then $N_n(x,t)$ is a strictly decreasing and a strictly convex function in $x \leq t$, $N_n(x,t) > 0$ for $x < t$.

If n is even, then $N_n(x,t)$ is a strictly increasing and a strictly concave function in $x \leq t$. Furthermore $N_0(x,t) = x - t$ is a strictly increasing and a trivially concave function in $x \leq t$.

Also $N_n(x,t) < 0$ $(x < t)$ for n zero or even; with $N_n(t,t) = 0$ all $n \geq 0$.

Proof. By assumption $u_1(x)$ is a concave function in $x \leq t$ implying that $u_1'(x)$ is a decreasing function there, that is, $\phi_1(x)$ is decreasing in $x \leq t$.

See that for $n \geq 1$

$$\frac{\partial g_n(x,t)}{\partial x} = \frac{\phi_1(x)}{\phi_1(t)\cdots\phi_n(t)} \int_t^x \phi_2(x_1)\int_t^{x_1}\cdots\int_t^{x_{n-2}} \phi_n(x_{n-1})dx_{n-1}\cdots dx_1$$

$$= (-1)^{n-1}\,\frac{\phi_1(x)B(x,t)}{\phi_1(t)\cdots\phi_n(t)},$$

where

$$B(x,t) = \int_x^t \phi_2(x_1)\int_{x_1}^t \cdots \int_{x_{n-2}}^t \phi_n(x_{n-1})dx_{n-1}\cdots dx_1 > 0 \quad (x < t),$$

$$B(t,t) = 0.$$

Since $B(x,t)$ is a strictly decreasing function in $x \leqslant t$, we get that $\phi_1(x)B(x,t)$ is also strictly decreasing in $x \leqslant t$. When $n > 1$ is odd then $\partial g_n(x,t)/\partial x > 0$ $(x < t)$, $\partial g_n(t,t)/\partial x = 0$ and it is a strictly decreasing function in $x \leqslant t$.

When n is even then $\partial g_n(x,t)/\partial x < 0$ $(x < t)$, $\partial g_n(t,t)/\partial x = 0$ and it is a strictly increasing function in $x \leqslant t$.

We have proved that for n odd, $g_n(x,t) < 0$ $(x < t)$, $g_n(t,t) = 0$ and it is strictly concave in $x \leqslant t$ for $n > 1$; clearly $g_1(x,t)$ is concave in $x \leqslant t$. Also for n even, $g_n(x,t) > 0$ $(x < t)$, $g_n(t,t) = 0$ and it is strictly convex in $x \leqslant t$.

From Lemma 1,

$$\frac{\partial^i N_n(x,t)}{\partial x^i} = \int_t^x \frac{\partial^i g_n(x,s)}{\partial x^i}ds, \quad (x \leqslant t,\ n \geqslant 2),\ \ i = 1,2.$$

It is clear that when $n > 2$ is odd, then $N_n(x,t)$ is a strictly decreasing and a strictly convex function in $x \leqslant t$.

And when n is even, then $N_n(x,t)$ is a strictly increasing and a strictly concave function in $x \leqslant t$. Note that for $n \geqslant 1$ odd, $N_n(x,t) > 0$ and for n zero or even, $N_n(x,t) < 0$, where $x < t$, with $N_n(t,t) = 0$ all $n \geqslant 0$. Obviously $N_0(x,t) = x - t$ is a trivially concave and a strictly increasing function in $x \leqslant t$.

From $\partial g_1(x,t)/\partial x = \phi_1(x)/\phi_1(t)$ and ϕ_1 a decreasing function, we have that $\partial g_1(x,t)/\partial x$ is decreasing in $x \leqslant t$. Obviously, $\partial g_1(x,t)/\partial x > 0$ for all $x \in [a,b]$.

Let s be such that $x_1 < x_2 \leqslant s \leqslant t$, then

$$\frac{\partial g_1(x_1,s)}{\partial x} \geqslant \frac{\partial g_1(x_2,s)}{\partial x}.$$

Adding

$$\int_{x_2}^{t} \frac{\partial g_1(x_1,s)}{\partial x}\,ds \geqslant \int_{x_2}^{t} \frac{\partial g_1(x_2,s)}{\partial x}\,ds$$

and

$$\int_{x_1}^{x_2} \frac{\partial g_1(x_1,s)}{\partial x}ds > 0$$

one has

$$\int_{x_1}^{t} \frac{\partial g_1(x_1,s)}{\partial x}\,ds > \int_{x_2}^{t} \frac{\partial g_1(x_2,s)}{\partial x}\,ds$$

or

$$\int_{t}^{x_1} \frac{\partial g_1(x_1,s)}{\partial x}ds < \int_{t}^{x_2} \frac{\partial g_1(x_2,s)}{\partial x}ds.$$

The last inequality and Lemma 1(10) imply that $\partial N_1(x,t)/\partial x$ is strictly increasing in $x \leqslant t$, which says that $N_1(x,t)$ is a strictly convex function in $x \leqslant t$. Since

$$\frac{\partial N_1(x,t)}{\partial x} < 0 \ (x < t), \quad \frac{\partial N_1(t,t)}{\partial x} = 0$$

we conclude that $N_1(x,t)$ is a strictly decreasing function in $x \leqslant t$. □

Lemmas 2 and 3 enable us to conclude

Lemma 4. Assume that $u_0(x) = c > 0$ and $u_1(x)$ is a concave function for $x \leqslant t$ and a convex function for $x \geqslant t$.

Let $\tilde{N}_n(x,t) = |N_n(x,t)|$, where

$$N_n(x,t) = \int_{t}^{x} g_n(x,s)ds, \quad \text{all} \quad x,t \in [a,b], \quad n \geqslant 0.$$

Then for $n \geq 1$, $\tilde{N}_n(x,t)$ is a strictly decreasing function in $x \leq t$ and a strictly increasing function in $x \geq t$, furthermore it is a continuous and a strictly convex function in $x \in [a,b]$.

Obviously $\tilde{N}_0(x,t) = |x-t|$ possesses the above properties, but is a convex function in $x \in [a,b]$. In particular, $\tilde{N}_n(x,t) > 0$ for $x \neq t$, with $\tilde{N}_n(t,t) = 0$, all $n \geq 0$.

Lemma 4 implies the next result which is used in the proof of Theorem 2.

Lemma 5. Under the assumptions of Lemma 4, for fixed $t \in (a,b)$, we have that

$$\tilde{N}_n(x,t) \leq \max\left\{ \frac{\tilde{N}_n(b,t)}{b-t}, \frac{\tilde{N}_n(a,t)}{t-a} \right\} \cdot |x-t|,$$

all $x \in [a,b]$; for all $n \geq 1$. Equality can be true only at $x = t$ and at $x = a$ or b.

The above inequality, for $n = 0$, becomes identity.

Proof. When $t < x < b$ by strict convexity of $\tilde{N}_n(x,t)$, $n \geq 1$; $\tilde{N}_n(t,t) = 0$ we get

$$\frac{\tilde{N}_n(x,t)}{x-t} < \frac{\tilde{N}_n(b,t)}{b-t} .$$

Thus

$$\tilde{N}_n(x,t) < \left[\frac{\tilde{N}_n(b,t)}{b-t} \right] \cdot (x-t)$$

$$\leq \max\left\{ \frac{\tilde{N}_n(b,t)}{b-t}, \frac{\tilde{N}_n(a,t)}{t-a} \right\} \cdot (x-t).$$

And when $a < x < t$, again by strict convexity of $\tilde{N}_n(x,t)$ we get

$$\frac{\tilde{N}_n(a,t)}{a-t} < \frac{\tilde{N}_n(x,t)}{x-t} .$$

Thus

$$\tilde{N}_n(x,t) < \left[\frac{\tilde{N}_n(a,t)}{t-a} \right] \cdot (t-x)$$

$$\leq \max\left\{ \frac{\tilde{N}_n(b,t)}{b-t}, \frac{\tilde{N}_n(a,t)}{t-a} \right\} \cdot (t-x). \qquad \square$$

References

[1] G.A. ANASTASSIOU, "A "K-Attainable" Inequality Related to the Convergence of Positive Linear Operators", J. Approx. Theory **44** (1985), 380-383.

[2] S. KARLIN and W.J. STUDDEN, "Tchebycheff Systems: With Applications in Analysis and Statistics", Interscience, New York, 1966.

[3] J. PEETRE, "A Theory of Interpolation of Normed Spaces", Notes Universidade de Brasilia, 1963.

[4] L.L. SCHUMAKER, "Spline Functions. Basic Theory", John Wiley and Sons, Inc., New York, 1981.

[5] D.V. WIDDER, "A Generalization of Taylor's Series", Trans. A.M.S., 30 (1928), 126-154.

Rev. Academia de Ciencias. Zaragoza 47 (1992)

ESTIMACIÓN DE LA VARIANZA CON MUESTREO SISTEMÁTICO

por

M. Ruiz Espejo

Departamento de Organización de Empresas
Facultad de Ciencias Económicas y Empresariales
Universidad Complutense de Madrid

Summary. Based on Zinger's (1980) and Ranà and Singh's (1989) papers, we are giving a new proof of the unbiasedness of the non-negative estimator of the variance of the estimator $\bar{x}^*$ proposed by Zinger (1980) for population mean $\bar{X}$, using other techniques for the analysis in finite populations, when a systematic sample s of size m is selected of the population and consecutively a simple random sample without replacement r of size n of the rest of units of U-s, being U the finite population of size N=mk.

1. Introducción

Sea una población finita de tamaño N=mk, denotada por $U = \{1,2,\ldots,N\}$. Tomamos una muestra sistemática, s, de tamaño m, de la población U. Sea $S = \{ s \subset U:$ s es una muestra sistemática de tamaño $m\}$. En estas condicicnes el estimador clásico media muestral, $\bar{X}_s$, es insesgado para estimar la media poblacional $\bar{X}$. Su varianza, $V(\bar{X}_s)$, no es posible ser estimada insesgadamente (Ruiz, 1986, o Rana y Singh, 1989). Con el objeto de obtener un estimador insesgado de la varianza de un estimador que aproveche la información suministrada por la muestra, Zinger (1980) propuso el estadístico

$$\bar{x}^* = \beta\bar{X}_s + (1-\beta)\bar{X}_r ,$$

siendo β una constante y $\bar{X}_r$ la media de una muestra aleatoria simple sin reemplazamiento, r, de tamaño n ($n+m \leq N$) obtenida de U-s, una vez observada s previamente.

En la sección 2 veremos que la varianza $V(\bar{x}^*)$ es estimable si y sólo si $\beta = m/N$, y calcularemos su expresión explícitamente. Finalmente, en la sección 3, obtendremos un estimador no negativo e insesgado de $V(\bar{x}^*)$, co-

rroborando los resultados de Rana y Singh (1989), aunque con diferentes técnicas.

2. Varianza del estimador de Zinger

En primer lugar comprobamos que $\bar{x}^*$ es insesgado para $\bar{X}$, independientemente del valor asignado a β.

$$E(\bar{x}^*) = \beta E_1[E_2(\bar{X}_s|s)] + (1-\beta)E_1[E_2(\bar{X}_r|s)] \quad . \tag{1}$$

Ahora, $E_2(\bar{X}_s|s) = \bar{X}_s$, y por tanto $E_1(\bar{X}_s) = \bar{X}$.

$$E_2(\bar{X}_r|s) = \frac{1}{N-m} \sum_{i \in U-s} x_i \ ,$$

y como para toda muestra sistemática $s \in S$, $p(s) = m/N$,

$$E_1\left(\frac{1}{N-m} \sum_{i \in U-s} x_i\right) = \frac{1}{N-m} E_1\left(\sum_{i \in U-s} x_i\right) = \frac{1}{N-m} \sum_{s \in S}\left(\sum_{i \in U-s} x_i\right)p(s) =$$

$$= \frac{1}{N-m}\frac{m}{N} \sum_{i \in U} x_i \, \mathrm{card}\{s \in S:\ i \in U-s\} =$$

$$= \frac{1}{N-m}\frac{m}{N} \sum_{i \in U} x_i(k-1) = \frac{m(k-1)}{N-m} \bar{X} = \bar{X} \ .$$

De (1), tenemos sustituyendo

$$E(\bar{x}^*) = \beta\bar{X} + (1-\beta)\bar{X} = \bar{X} \ ,$$

es decir, $\bar{x}^*$ es insesgado para estimar $\bar{X}$, para todo β.

A continuación calculamos $V(\bar{x}^*)$, usando el teorema de Madow,

$$V(\bar{x}^*) = V_1[E_2(\bar{x}^*|s)] + E_1[V_2(\bar{x}^*|s)] \quad . \tag{2}$$

Ahora,

$$E_2(\bar{x}^*|s) = \beta\bar{X}_s + (1-\beta)\left(\frac{1}{N-m} \sum_{i \in U-s} x_i\right) \ .$$

$$V_1[E_2(\bar{x}^*|s)] = \beta^2 V_1(\bar{X}_s) + (1-\beta)^2 \frac{1}{(N-m)^2} V_1(N\bar{X}-m\bar{X}_s) +$$

$$+ 2\beta(1-\beta)\mathrm{Cov}_1(\bar{X}_s,\bar{X}_s)\,\frac{-m}{N-m} =$$

$$= V_1(\bar{X}_s)\left[\beta^2 + (1-\beta)^2\frac{m^2}{(N-m)^2} + 2\beta(1-\beta)\frac{-m}{N-m}\right] =$$

$$= V_1(\bar{X}_s)\left[\beta - (1-\beta)\frac{m}{N-m}\right]^2 = V_1(\bar{X}_s)\left(\frac{\beta N - m}{N-m}\right)^2 \qquad (3)$$

Luego veremos que $E_1[V_2(\bar{x}^*|s)]$ es estimable, pero $V_1[E_2(\bar{x}^*|s)]$ no es estimable (pues $V_1(\bar{X}_s)$ no lo es) salvo que la constante que le afecta en (3) sea nula. En efecto, $V_1[E_2(\bar{x}^*|s)]$ es estimable insesgadamente si y sólo si

$$\left(\frac{\beta N-m}{N-m}\right)^2 = 0,$$

o bien si y sólo si $\beta = m/N$.

Continuando,

$$E_1[V_2(\bar{x}^*|s)] = E_1\left\{V_2[\beta\bar{X}_s + (1-\beta)\bar{X}_r|s]\right\} = E_1[(1-\beta)^2 V_2(\bar{X}_r|s)] \ .$$

Pero,

$$V_2(\bar{X}_r|s) = V_2\left(\frac{1}{n}\sum_{i\in r} x_i|s\right) = \frac{(N-m)-n}{(N-m)n}\,S^2_{U-s}$$

donde S^2_{U-s} es la cuasivarianza de la variable x en la población U-s. Por tanto,

$$E_1[V_2(\bar{x}^*|s)] = E_1\left[(1-\beta)^2\,\frac{N-m-n}{(N-m)n}\,S^2_{U-s}\right] = (1-\beta)^2\frac{N-m-n}{(N-m)n}\,E_1(S^2_{U-s}) \ , (4)$$

y sustituyendo (4) en (2) para $\beta = m/N$, tenemos

$$V(\bar{x}^*) = E_1[V_2(\bar{x}^*|s)] = \frac{(N-m)(N-m-n)}{nN^2}\,E_1(S^2_{U-s}) \ . \qquad (5)$$

3. Estimador insesgado no negativo

Para estimar insesgadamente $V(\bar{x}^*)$ proponemos el estimador cs^2_r, donde c es una constante a determinar y s^2_r la cuasivarianza muestral para la muestra r, es decir

$$s_r^2 = \frac{1}{n-1} \sum_{i \in r} (x_i - \bar{X}_r)^2 = \frac{1}{n(n-1)} \sum_{i<j \in r} \sum (x_i - x_j)^2 .$$

Ahora bien,

$$E(s_r^2) = E_1[E_2(s_r^2|s)] = E_1(S_{U-s}^2) , \qquad (6)$$

y de aquí, si obligamos a que $E(cs_r^2) = V(\bar{x}^*)$, obtenemos que de (5) y (6)

$$c = \frac{V(\bar{x}^*)}{E(s_r^2)} = \frac{(N-m)(N-m-n)}{nN^2} . \qquad (7)$$

Conclusión: cs_r^2 es un estimador insesgado y no negativo de $V(\bar{x}^*)$, donde c viene dada en (7). De este modo queda explicada otra demostración del resultado de Rana y Singh (1989).

Referencias

|1| Rana, R.S. y Singh, R. (1989). Note on systematic sampling with supplementary observations. Sankhyā Ser. B 51, 205-211.

|2| Ruiz, M. (1986). Funciones paramétricas estimables en teoría de muestras. Estadística Española 28, nº 112-113, 69-73.

|3| Zinger, A. (1980). Variance estimation in partially systematic sampling. J. Amer. Statist. Assoc. 75, 206-211.

Rev. Academia de Ciencias. Zaragoza 47 (1992)

OPTIMALIDAD DE LA MEDIA MUESTRAL

por

M. Ruiz Espejo

Departamento de Estadística
Facultad de Ciencias Económicas y Empresariales
Universidad Complutense de Madrid

Summary. In this paper, we prove the optimality of sample mean of size m+n, as integrator unbiased statistic "linear combination" of the sample means of sizes m and n respectively, and not only for simple random sampling with replacement of any population, but also without replacement with disjoint samples of sizes m and n of a finite population.

1. Introducción

En un contexto general de inferencia en poblaciones finitas (véase Ruiz, 1987) se sabe que no existe estimador UMV (uniformemente de mínima varianza) en la clase de todos los estimadores insesgados de funciones paramétricas que dependen funcionalmente de todos los argumentos del parámetro, para todos los diseños no censales. Sin embargo, en condiciones más restringidas puede encontrarse a veces un estimador UMV en ciertas clases de estimadores insesgados.

En el presente trabajo presentamos uno de estos casos que completa aún más otros resultados en la misma dirección como puede verse en Cassel et al. (1977, sección 3.5).

Sea una población finita o infinita, con varianza finita σ^2. Si se toman dos muestras independientes de tamaños m y n respectivamente, el estimador integrador de sus medias muestrales parciales, $\bar{x}_s$ y $\bar{x}_r$, como combinación lineal de ambas (de modo similar al usado en Ruiz, 1992)

$$\bar{x}^* = k\bar{x}_s + (1-k)\bar{x}_r \ , \qquad (1)$$

alcanza su máxima precisión cuando $k = m/(m+n)$, siendo siempre insesgado $\bar{x}^*$

para la media poblacional $\bar{x}$,

$$E(\bar{x}^*) = kE(\bar{x}_s) + (1-k)E(\bar{x}_r) = k\bar{x} + (1-k)\bar{x} = \bar{x} \text{ para todo } k \in \mathbb{R}.$$

En efecto,

$$V(\bar{x}^*) = k^2 V(\bar{x}_s) + (1-k)^2 V(\bar{x}_r) = k^2 \frac{\sigma^2}{m} + (1-k)^2 \frac{\sigma^2}{n} =$$

$$= \sigma^2 \left[\frac{k^2}{m} + \frac{(1-k)^2}{n}\right] .$$

Al variar $k \in \mathbb{R}$, $V(\bar{x}^*)$ se minimiza si y sólo si $k = m/(m+n) = k^*$. Para este valor de k óptimo, $\bar{x}^*$ es la media muestral de las m+n observaciones que constituían las muestras independientes de tamaños m y n.

Esto, que es razonable para diseños de muestreo aleatorio simple con reemplazamiento, ¿seguirá siendo válido para muestreo aleatorio simple sin reemplazamiento en poblaciones finitas? La respuesta es sí, como veremos a continuación aunque su justificación es sensiblemente más laboriosa.

2. Planteamiento del problema

Sea $U = \{1,2,\ldots,N\}$ una población finita de tamaño N. De U se selecciona una muestra $s \subset U$ de tamaño m por muestreo aleatorio simple sin reemplazamiento. De esta muestra se obtiene la media muestral $\bar{x}_s$. Seguidamente se selecciona otra muestra aleatoria simple sin reemplazamiento $r \subset U-s$ de tamaño n ($m \geq 2$, $n \geq 2$, $m+n \leq N$), de la que obtenemos la estimación media muestral $\bar{x}_r$. La muestra $s \cup r$ es aleatoria simple sin reemplazamiento de tamaño m+n (Hedayat y Sinha, 1991, teorema 4.2).

El estimador $\bar{x}^*$ dado en (1) es ahora también insesgado para la media poblacional $\bar{x}$, independientemente del valor asignado a k,

$$E(\bar{x}^*) = E_1[E_2(\bar{x}^*|s)] = E_1[k\bar{x}_s + (1-k)E_2(\bar{x}_r|s)] =$$

$$= E_1[k\bar{x}_s + (1-k)\bar{x}_{U-s}] = k\bar{x} + (1-k)\bar{x} = \bar{x} ,$$

donde

$$E_1(\bar{x}_{U-s}) = \frac{1}{N-m} E_1(N\bar{x} - m\bar{x}_s) = \frac{1}{N-m}(N\bar{x} - m\bar{x}) = \bar{x} .$$

3. Varianza

Haciendo uso del teorema de Madow,

$$V(\bar{x}^*) = V[k\bar{x}_s + (1-k)\bar{x}_r] = E_1[(1-k)^2 V_2(\bar{x}_r|s)] +$$

$$+ V_1[k\bar{x}_s + (1-k)E_2(\bar{x}_r|s)] =$$

$$= (1-k)^2 E_1\left[\frac{(N-m)-n}{(N-m)n} S^2_{U-s}\right] +$$

$$+ k^2 V_1(\bar{x}_s) + (1-k)^2 V_1(\bar{x}_{U-s}) + 2k(1-k)\mathrm{Cov}_1(\bar{x}_s, \bar{x}_{U-s}) . \quad (2)$$

Ahora

$$E_1(S^2_{U-s}) = \frac{N}{N-1}\sigma^2 , \quad (3)$$

$$V_1(\bar{x}_s) = \frac{N-m}{(N-1)m}\sigma^2 , \quad (4)$$

$$V_1(\bar{x}_{U-s}) = \frac{N-(N-m)}{(N-1)(N-m)}\sigma^2 \quad (5)$$

y por (4) tenemos

$$\mathrm{Cov}_1(\bar{x}_s, \bar{x}_{U-s}) = \mathrm{Cov}_1\left[\bar{x}_s, \frac{1}{N-m}(N\bar{x} - m\bar{x}_s)\right] = \frac{-m}{N-m} V_1(\bar{x}_s) = -\frac{\sigma^2}{N-1} . \quad (6)$$

Sustituyendo (3), (4), (5) y (6) en (2) y operando tenemos

$$V(\bar{x}^*) = (1-k)^2 \frac{(N-m-n)N}{(N-1)(N-m)n}\sigma^2 + \frac{(kN-m)^2}{(N-1)(N-m)m}\sigma^2 =$$

$$= \frac{\sigma^2}{(N-1)(N-m)mn} f(k) \quad (7)$$

donde $f(k) = (1-k)^2(N-m-n)Nm + (kN-m)^2 n$.

Para minimizar $V(\bar{x}^*)$, basta minimizar $f(k)$. Derivando $f(k)$ con respecto a k, e igualando a cero, obtenemos $k = k^* = m/(m+n)$, que no depende de N, al cual le corresponde un mínimo pues $f''(k) = 2[N(N-m-n)m + N^2 n] > 0$.

Particularizando k^* en $\bar{x}^*$, tenemos que $\bar{x}^*$ es la media muestral de m+n

unidades, concretamente de la muestra $s \cup r$. Sustituyendo $k = k^*$ en $f(k)$ y a su vez ésta en (7), deducimos simplificando que

$$f(k^*) = \frac{(N-m-n)(N-1)mn}{m+n},$$

y de aquí,

$$V(\bar{x}^*) = \frac{N-m-n}{(N-1)(m+n)} \sigma^2 ,$$

que coincide con la varianza de la media muestral de tamaño m+n, bajo muestreo aleatorio simple sin reemplazamiento.

4. Conclusión

La media muestral de tamaño m+n se comporta como el estimador óptimo en el sentido de "mínima varianza", en la clase de estimadores insesgados combinación lineal de medias muestrales de tamaños m y n respectivamente, bajo los diseños de muestreo aleatorio simple con y sin reemplazamiento. De este modo queda explicado otro teorema de existencia de estimadores insesgados UMV bajo condiciones restringidas.

Referencias

|1| Cassel, C.M., Särndal, C.E. y Wretman, J.H. (1977). Foundations of Inference in Survey Sampling, Wiley, Nueva York.

|2| Cochran, W.G. (1977). Sampling Techniques (3ª edición), Wiley, Nueva York.

|3| Grosbras, J.M. (1987). Méthodes Statistiques des Sondages, Économica, París.

|4| Hedayat, A.S. y Sinha, B.K. (1991). Design and Inference in Finite Population Sampling, Wiley, Nueva York.

|5| Ruiz, M. (1987). Sobre estimadores UMV y UMECM en poblaciones finitas. Estadística Española 29, nº 115, 105-111.

|6| Ruiz, M. (1992). Estimación de la varianza con muestreo sistemático. Rev. Acad. Cienc. Zaragoza (bajo recensión).

Rev. Academia de Ciencias. Zaragoza 47 (1992)

Corrección de órbitas de pares visuales por medio de series de Fourier de la anomalía media

R. Cid y C. Longás

Departamento de Física Teórica

Facultad de Ciencias

50009 - Universidad de ZARAGOZA

Abstract

In this article, using Fourier series of coefficients (θ_0, a_k, b_k) and the variable M (mean anomaly), a method for the correction of previous orbits of visual binary stars is developed. For this purpose and by means of the initial elements $(a_0, e_0, T_0, P_0, \Omega_0, \omega_0, I_0)$ and residual differences $\Delta\theta$ in position angles, the computation of the increments $(\Delta\theta_0, \Delta a_k, \Delta b_k)$ and $(\Delta a_0, \Delta e_0, \Delta T_0, \Delta P_0, \Delta\Omega_0, \Delta\omega_0, \Delta I_0)$, is accomplished. In order to determine this increments, the method of least squares is applied, but this determination can be obtained by means another numerical methods.

1. Introducción

En nuestros trabajos (R. Cid, 1950, 1952, 1960 y 1989) se ha desarrollado un método de corrección de órbitas elípticas de estrellas dobles visuales, utilizando desarrollos en serie de Fourier para las distancias y el tiempo, de los tipos $\rho^2 = f(a_k, b_k; \theta)$ y $t = F(k_0, A_k, B_k; \theta)$, es decir, donde la variable que interviene en dichos desarrollos es el ángulo de posición θ.

En el presente artículo, de manera semejante, la variable que interviene en los desarrollos es la anomalía media.

En lo sucesivo y como es habitual, si E_0 designa la estrella principal, E la estrella satélite y E' la proyección cilíndrica de E sobre el plano tangente a la esfera celeste en E_0, utilizaremos las siguientes notaciones:

En la órbita relativa

- a semieje mayor
- e excentricidad
- P periodo orbital
- T época de paso por el periastro A

Ω	ángulo del nodo
ω	argumento del periastro
I	inclinación
$r = dist(E_0E)$	radio vector
f	anomalía verdadera
$n = \frac{2\pi}{P}$	movimiento medio
$C = na^2\sqrt{1-e^2}$	constante de las áreas
$p = a(1-e^2)$	parámetro
$M = n(t-T)$	anomalía media
ξ	anomalía excéntrica

En la órbita aparente

$\rho = dist(E_0E')$	distancia angular observada
$\theta = ang(NE_0E')$	ángulo de posición
$C' = C\cos I$	constante de las áreas en la aparente

2. Desarrollos en series de Fourier de la anomalía media

Se comprueba facilmente que la función

$$f(M) = -\frac{2C'}{n\rho^3}\frac{d\rho}{dM}$$

es continua y periódica, de periodo 2π, y por tanto admite un desarrollo en serie de Fourier, uniformemente convergente, de la forma

$$f(M) = \alpha_0 + \sum_{k=1}^{\infty}(\alpha_k \cos kM - \beta_k \operatorname{sen} kM) \tag{1}$$

donde los coeficientes α_k , β_k , están acotados y $\alpha_0 = 0$, puesto que

$$\alpha_0 = \frac{1}{2\pi}\int_0^{2\pi} f(M)dM = \left[\frac{C'}{n\rho^2}\right]_0^{2\pi} = 0$$

En efecto, la continuidad y el carácter periódico de la función $f(M)$ depende de que lo sean las funciones ρ^{-3} y $\frac{d\rho}{dM}$, una vez excluida la posibilidad de una colisión aparente ($\rho = 0$). Ahora bien, se tiene

$$\frac{d\rho}{dM} = \frac{d\rho}{df}\frac{df}{d\xi}\frac{d\xi}{dM} = \frac{ab}{r^2}\frac{d\rho}{df} \tag{2}$$

Por otra parte, si (A, B, F, G), son las constantes de Innes (1926 y 1932), definidas por las igualdades

$$\begin{aligned} A &= \cos\Omega\cos\omega - \operatorname{sen}\Omega\operatorname{sen}\omega\cos I \\ B &= \operatorname{sen}\Omega\cos\omega + \cos\Omega\operatorname{sen}\omega\cos I \\ F &= -\cos\Omega\operatorname{sen}\omega - \operatorname{sen}\Omega\cos\omega\cos I \\ G &= -\operatorname{sen}\Omega\operatorname{sen}\omega + \cos\Omega\cos\omega\cos I \end{aligned} \tag{3}$$

podemos establecer las bien conocidas ecuaciones

$$\begin{aligned} \rho\operatorname{sen}\theta &= r(B\cos f + G\operatorname{sen} f) \\ \rho\cos\theta &= r(A\cos f + F\operatorname{sen} f) \end{aligned} \tag{4}$$

Haciendo para abreviar

$$A^* = -A \operatorname{sen} f + F \cos f, \qquad B^* = -B \operatorname{sen} f + G \cos f$$

y derivando las igualdades (4) con respecto a f, podemos obtener facilmente la fórmula

$$\frac{d\rho}{df} = \frac{\rho}{r}\frac{dr}{df} + r(A^* \cos\theta + B^* \operatorname{sen}\theta)$$

y finalmente

$$\frac{d\rho}{dM} = \frac{ab}{r}\left[\frac{e\rho}{p} \operatorname{sen} f + A^* \cos\theta + B^* \operatorname{sen}\theta\right] \tag{5}$$

En resumen, excluidos los casos de colisión orbital y aquellos en los que es $I = \frac{\pi}{2}$ (órbita de plano perpendicular al plano de la órbita aparente), se demuestra la existencia del desarrollo en serie dado por la fórmula (1)

Integrando la fórmula (1), con respecto a M, en el intervalo $(0, M)$, se obtiene

$$\frac{C'}{n\rho^2} = a_0 + \sum_{k=1}^{\infty}(a_k \operatorname{sen} kM + b_k \cos kM) \tag{6}$$

siendo constantes los coeficientes

$$a_0, \qquad a_k = \frac{\alpha_k}{k}, \qquad b_k = \frac{\beta_k}{k}$$

Por otra parte, si tenemos en cuenta la ley de las áreas, podemos establecer la igualdad

$$d\theta = \frac{C'}{n\rho^2} dM \tag{7}$$

y por consiguiente, integrando esta ecuación en los intervalos correspondientes (θ_0, θ) , $(0, M)$, tendremos

$$\theta - \theta_0 = a_0 M + \sum_{k=1}^{\infty}\left[\frac{a_k}{k}(1 - \cos kM) + \frac{b_k}{k} \operatorname{sen} kM\right] \tag{8}$$

Evidentemente, esta igualdad aplicada a los ángulos de posición $\theta(M)$ y $\theta(M + 2\pi)$ nos demuestra la condición $a_0 = 1$ y por tanto los desarrollos (6), (8) se pueden escribir en sus definitivas formas. Esto es,

$$\frac{C'}{n\rho^2} = 1 + \sum_{k=1}^{\infty}(a_k \operatorname{sen} kM + b_k \cos kM) \tag{9}$$

$$\theta = \theta_0 + M + \sum_{k=1}^{\infty}\left[\frac{a_k}{k}(1 - \cos kM) + \frac{b_k}{k} \operatorname{sen} kM\right] \tag{10}$$

Una vez demostrada la posibilidad de estos desarrollos, podemos hacer la siguiente observación:

Por ser acotados los coeficientes α_k , β_k , y por verificarse las igualdades

$$\frac{a_k}{k} = \frac{\alpha_k}{k^2}, \qquad \frac{b_k}{k} = \frac{\beta_k}{k^2}$$

cualquier serie parcial del desarrollo

$$F_k(M) = \sum_{k=1}^{\infty}\left[\frac{a_k}{k}(1 - \cos kM) + \frac{b_k}{k} \operatorname{sen} kM\right] \tag{11}$$

es uniformemente convergente, por tratarse de series armónicas de orden $(\frac{1}{k^2})$. En consecuencia, si i , p , designan, respectivamente, números naturales impares y pares, los desarrollos $F_i(M)$ y $F_p(M)$, son también uniformemente convergentes.

3. Método de corrección de órbitas

Los desarrollos (9) y (10) dependen de un conjunto infinito de constantes (a_k, b_k) cuyo cálculo directo solamente podría ser efectuado por medio de las integrales

$$a_k = \frac{C'}{n}\int_{-\pi}^{\pi}\frac{1}{\rho^2}\operatorname{sen} kM\,dM,$$
$$b_k = \frac{C'}{n}\int_{-\pi}^{\pi}\frac{1}{\rho^2}\cos kM\,dM$$

No obstante, si podemos suponer que los errores de observación $\Delta\theta$ se traducen en errores $(\Delta\theta_0, \Delta a_k, \Delta b_k)$ de las constantes (θ_0, a_k, b_k) y en errores ΔP, ΔT , de los elementos orbitales P y T, dado que los datos de tiempo se consideran exactos y por tanto $\Delta t = 0$.

Diferenciando la fórmula (10), tendremos

$$\Delta\theta = \Delta\theta_0 + \sum_{k=1}^{\infty}\left[\frac{\Delta a_k}{k}(1-\cos kM) + \frac{\Delta b_k}{k}\operatorname{sen} kM\right] + \left[1 + \sum_{k=1}^{\infty}(a_k \operatorname{sen} kM + b_k \cos kM)\right]\Delta M$$

o bien

$$\Delta\theta = \Delta\theta_0 + \sum_{k=1}^{\infty}\left[\frac{\Delta a_k}{k}(1-\cos kM) + \frac{\Delta b_k}{k}\operatorname{sen} kM\right] - \frac{C'}{n\rho^2}\left[\frac{M\Delta P}{P} + n\Delta T\right] \tag{12}$$

puesto que

$$\Delta M = -\left[\frac{M\Delta P}{P} + n\Delta T\right] \tag{13}$$

Ahora bien, dado un conjunto m de observaciones θ_α de un par visual y suponiendo calculada una órbita previa de elementos $(a_c, e_c, T_c.P_c, \Omega_c, \omega_c, I_c)$, podemos calcular las constantes

$$n_c, \qquad C'_c = n_c a_c^2\sqrt{1-e_c^2}\cos I_c, \qquad (A_c, B_c, F_c, G_c),$$

y las efemérides ρ_c, θ_c, para las distintas épocas de observación, así como el valor $(\theta_0)_c$, correspondiente a la anomalía media $M_c = 0$.

En estas condiciones, aplicando, por ejemplo, un proceso de mínimos cuadrados a la ecuación (12), con un cierto número finito de coeficientes Δa_k, Δb_k, que se extienda al conjunto m de observaciones disponibles, obtendremos, por medio del sistema de ecuaciones normales de Gauss, los incrementos

$$\Delta\theta_0, \qquad \Delta a_k, \qquad \Delta b_k, \qquad \Delta P, \qquad \Delta T.$$

cuyos coeficientes respectivos, son:

$$1, \qquad \frac{1}{k}(1-\cos kM_c), \qquad \frac{1}{k}sen kM_c, \qquad \frac{C'_c M_c}{n_c\rho_c^2 P_c}, \qquad \frac{C'_c}{\rho_c^2}.$$

siendo, en cada caso, $\Delta\theta = \theta_\alpha - \theta_c$. las diferencias (O-C).

Resultarán así directamente los nuevos elementos

$$\theta_0 = (\theta_0)_c + \Delta\theta_0, \qquad P_0 = P_c + \Delta P, \qquad T_0 = T_c + \Delta T$$

juntamente con los incrementos Δa_k, Δb_k, cuyo único interés reside en que sirven de control a los cálculos.

Para obtener la corrección Δe, de la excentricidad, podemos proceder del siguiente modo:

Si consideramos el valor particular de la anomalía verdadera $f_1 = \frac{\pi}{2}$, la anomalía excéntrica correspondiente, es decir ξ_1 , será

$$\operatorname{tag}\frac{\xi_1}{2} = \sqrt{\frac{1-e}{1+e}}\operatorname{tag}\frac{f_1}{2} = \operatorname{tag}\frac{\phi}{2}$$

de donde se deduce la igualdad $\xi_1 = \phi$. siendo $e = \cos\phi$

Por tanto, tendremos

$$\Delta M_1 = \Delta(\phi - e\operatorname{sen}\phi) = \Delta\phi(1 - e\cos\phi) - \Delta e\operatorname{sen}\phi$$

$$\Delta e = -\operatorname{sen}\phi\Delta\phi$$

de donde

$$\Delta M_1 = -\Delta e\left[\frac{1-\cos^2\phi}{sen\phi} + \operatorname{sen}\phi\right]$$

y finalmente

$$\Delta M_1 = -2\sqrt{1-e^2}\Delta e \tag{14}$$

que, según (13), nos permitirá calcular Δe por la igualdad

$$\Delta e = \frac{1}{2\sqrt{1-e^2}}\left(\frac{(\phi - e\operatorname{sen}\phi)}{P}\Delta P + n\Delta T\right) \tag{15}$$

Debemos observar que el valor obtenido por la igualdad (15) viene dado en radianes, por lo cual el resultado ha de ser dividido por 2π para reducirlo a su valor real.

De esta forma se obtiene un nuevo valor de la excentricidad por la igualdad

$$e_0 = e_c + \frac{1}{2\pi}\Delta e$$

Recordemos ahora, que la obtención de los incrementos Δa_k y Δb_k solamente es empleado como control de los resultados, por lo cual la resolución del sistema de ecuaciones de Gauss se simplifica notablemente si se introducen unos incrementos Δa_0 y Δb_0 que verifiquen las igualdades

$$\sum_{k=1}^{\infty}\frac{1}{k}(1-\cos kM)\Delta a_k = \Delta a_0\sum_{k=1}^{\infty}\frac{1}{k}(1-\cos kM)$$

$$\sum_{k=1}^{\infty}\frac{1}{k}\operatorname{sen}kM\,\Delta b_k = \Delta b_0\sum_{k=1}^{\infty}\frac{1}{k}\operatorname{sen}kM$$

ya que en este caso el sistema de ecuaciones normales de Gauss contiene solamente cinco incognitas.

En esencia, dicha transformación equivale a utilizar unos incrementos Δa_0 y Δb_0 que son los promedios ponderados de los incrementos Δa_k y Δb_k.

Una vez calculados los nuevos elementos θ_0, P_0, T_0 y e_0, podemos proceder a calcular los elementos Ω_0, ω_0 e I_0.

En efecto, si consideramos las ecuaciones

$$\rho\cos(\theta-\Omega)=r\cos(\omega+f),\qquad \rho\,\mathrm{sen}(\theta-\Omega)=r\,\mathrm{sen}(\omega+f)\cos I.$$

basta dividirlas para obtener

$$\mathrm{tag}(\theta-\Omega)=\mathrm{tag}(\omega+f)\cos I \tag{16}$$

Ahora bien, puesto que ya suponemos mejorados los elementos orbitales P_0, T_0 y e_0, entonces para cada instante t_i de observación, también habrán sido mejorados los valores de las anomalías M_i, E_i, f_i, dadas por las igualdades

$$M_i=n_0(t_i-T_0),\qquad \xi_i-e_0\,\mathrm{sen}\,\xi_i=M_i,\qquad \mathrm{tag}\,\frac{f_i}{2}=\sqrt{\frac{1+e_0}{1-e_0}}\,\mathrm{tag}\,\frac{\xi_i}{2}$$

En consecuencia, si se incrementa la ecuación (16) con f constante, tendremos

$$\frac{\Delta\theta-\Delta\Omega}{\cos^2(\theta-\Omega)}=\frac{\cos I\Delta\omega}{\cos^2(\omega+f)}-\mathrm{tag}(\omega+f)\,\mathrm{sen}\,I\Delta I$$

o bien

$$\Delta\Omega+R\Delta\omega-S\Delta I=\Delta\theta \tag{17}$$

siendo

$$R=\frac{\cos^2(\theta_i-\Omega_c)\cos I_c}{\cos^2(\omega_c+f_i)},\qquad S=\mathrm{tag}(\omega_c+f_i)\,\mathrm{sen}\,I_c\cos^2(\theta_i-\Omega_c)$$

y la ecuación (17), tratada por mínimos cuadrados nos proporcionará los incrementos $\Delta\Omega$, $\Delta\omega$, ΔI, y por consiguiente

$$\Omega_0=\Omega_c+\Delta\Omega,\qquad \omega_0=\omega_c+\Delta\omega,\qquad I_0=I_c+\Delta I$$

Evidentemente, todo el proceso es iterativo y puede ser repetido cuantas veces se considere necesario.

Los nuevos valores P_0, T_0, e_0, Ω_0, ω_0, I_0, así obtenidos, determinan la órbita corregida en cuanto a ángulos de posición y habrá de calcularse el semieje mayor a, puesto que las distancias no han intervenido en el cálculo.

Para ello, con los elementos orbitales anteriores iremos calculando las distancias $(\rho_i)_c$ para un supuesto semieje mayor $a=1$, y los compararemos con los valores $(\rho_i)_0$ observados. El promedio aritmético de los cocientes

$$\frac{(\rho_i)_0}{(\rho_i)_c}=\alpha_i$$

nos dará el nuevo semieje mayor a_0 en la forma

$$a_0=\frac{1}{m}\sum_1^m\alpha_i$$

con lo cual queda terminado el proceso.

Es indudable que la débil convergencia de las series de Fourier y la imposibilidad de introducir infinitos términos en el cálculo puede originar perturbaciones en la corrección. Por esta razón, en la actualidad estamos efectuando correcciones sobre órbitas simuladas, que publicaremos más adelante, para ver en que casos el método resulta adecuado.

4. Bibliografía

CID, R. (1950): Rev. de Geofísica, 35 y 36, 1-58, Madrid.

CID, R. (1952): Urania, n 232, 201.

CID, R. (1989): Rev. Acad. Ciencias de Zaragoza, 44, 71-88.

INNES, R.T.A. y VAN DEN BOS, W.H. (1926): Union Obs. Cir., 68, 354. y (1932), Union Obs. Cir. 86, 261.

VALBOUSQUET, A. (1979): Ast. and Astroph., 77, 159-164.

Rev. Academia de Ciencias. Zaragoza 47 (1992)

Elementos de un sistema experto para la resolución de problemas de Astronomía de Posición

Abad, A.[1], San Juan, J.F[2].

[1]Grupo de Mecánica Espacial. Universidad de Zaragoza, 50009 Zaragoza

E–mail: abad@cc.unizar.es

[2]Departamento de Matemáticas. Estadística y Computación

Universidad de la Rioja, 26004 Logroño.

Abstract

In this paper we analyze the elements that an expert astronomer takes into account in order to solve a problem of Positional Astronomy. We use this analysis to clarify the methods of solution of problems to the students of Astronomy and to construct an expert system that emulates the Astronomer.

1. Introducción

De unos años a esta parte las nuevas técnicas de Inteligencia Artificial han extendido el campo de aplicación de la informática a dominios donde por falta de un algoritmo claro de resolución se hacía dificil su aplicación. En particular, los sistemas expertos constituyen un conjunto de programas que emulan la actuación de un experto humano, quién a partir de la formulación de un problema, aplica una serie de reglas que permiten extraer el mayor número de consecuencias posibles de la información que posee.

Para pensar sobre la filosofía de un sistema experto pensemos en un ejemplo simple, extraido de [1]: *Reflexione sobre el modo en que usted realiza una suma: cuando se trata de dos enteros pequeños (< 10 o 20), obtenemos el resultado mediante la fusión de los dos números; cuando los dos números son mayores, hemos de plantearnos la suma en nuestra mente, teniendo en cuenta "cuantos llevamos"; pero cuando los dos números son demasiado grandes, hemos de realizar la suma por escrito. De modo que, aún en el caso de una operación tan simple, según los datos efectuaremos la operación de una forma u otra.*

Este ejemplo, aunque de una manera muy simplificada, muestra uno de los posibles campos de aplicación de los sistemas expertos, aquél en que no existe un único algoritmo de resolución sino que el algoritmo depende de la formulación del problema y de los datos o informaciones que poseemos sobre el mismo.

Cuando un alumno de Astronomía se enfrenta por primera vez a la resolución de problemas que no impliquen simples cambios de coordenadas astronómicas, sino que relacionen varias

observaciones diferentes, comienzan para él las primeras dificultades de comprensión de la asignatura. Para el alumno cada problema tiene su propio algoritmo, diferente en cada caso, y es la experiencia la que posteriormente le va indicando una serie de pautas generales para la resolución de dichos problemas.

El objeto del presente trabajo consiste en el análisis de los diferentes elementos que aparecen en un problema de Astronomía de Posición, o lo que es igual el análisis que del mismo problema realizará un experto. Esto nos permitirá, desde un punto de vista docente, clarificar para el alumno el proceso de resolución de problemas y asímismo la construcción de un sistema experto informatizado que sirva de test para las ideas discutidas y de ayuda para el usuario de la Astronomía.

Aunque no es un programa específicamente diseñado para la construcción de sistemas expertos, hemos elegido el procesador algebráico de carácter general *Mathematica* para el diseño de nuestro sistema que hemos llamado `Hiparco.m` en honor del primer gran astrónomo de la antiguedad.

2. La *observación* como elemento fundamental de un problema

Elegiremos para comenzar el enunciado de un sencillo problema extraido de [3] (TP-2, pp 316):

Calcúlense las coordenadas geográficas de un Observatorio para el cual la estrella de coordenadas $\alpha = 19^h45^m35^s.13, \delta = -22°53'42''.2$ *se encuentra en su zenit a las* 5h *de TU en el dia en que* $\theta_0 = 17^h37^m39^s.352$.

Este problema contiene todos los elementos del objeto básico que nos vamos a encontrar en todo problema de Astronomía de Posición: *la observación*. Analizado el enunciado podemos apreciar los tres elementos que aparecen en toda observación: *el observatorio, el astro observado y el instante de la observación*. Toda la información que encontramos en el enunciado de este problema consiste en propiedades o parámetros asociadas a uno o varios elementos de la observación.

En efecto, las coordenadas geográficas que se nos piden caracterizan el observatorio. Las coordenadas ecuatoriales α, δ son dos parámetros asociados al astro observado. La hora de TU es un parámetro que caracteriza exclusivamente al instante de la observación. El hecho de que la estrella esté en ese instante en el zenit nos está informando sobre la distancia cenital de la estrella, que es un parámetro dependiente de los tres elementos de la observación, observatorio, astro e instante.

En todo problema nos encontraremos datos que afectan a alguno de los tres elementos de una observación. Asímismo, combinando ternas de elementos formadas por un observatorio, un astro y un instante tendremos las diferentes observaciones que aparecen en el problema, una en el caso del enunciado.

La forma habitual de resolver el problema sería diseñar el algoritmo que en función de los datos conocidos nos permita responder la pregunta planteada. Sin embargo nuestro propósito

es algo más general y consisite en analizar toda la información poseida y mediante una serie de reglas obtener todas las posibles consecuencias que de ella se deriven, entre las que puede encontrarse la incógnita del problema.

Para conseguir este objetivo hemos de analizar en primer lugar toda la posible información asociada a uno, dos o tres de los elementos de la información. Toda la información sobre los diferentes conceptos que se tratan en los siguientes parrafos y sus relaciones puede verse en [2] y [4].

Un observatorio viene caracterizado esencialmente por su longitud oeste λ_w y su latitud, que puede ser geográfica ϕ, geocéntrica ψ o reducida u. Las relaciones entre estas tres últimas vienen dadas como funciones lineales de sus tangentes con coeficientes dependientes de una constante conocida como es la excentricidad del elipsoide de revolución terrestre. La función `Solve` de *Mathematica* nos permite la obtención de las tres latitudes, conocida una, sin saber a priori cual es el tipo de latitud conocida y sin acudir al uso de la función `if` de los lenguajes clásicos. Se ha tenido en cuenta también, como parámetro que define un observatorio, la longitud media de la zona que sirve para determinar la hora oficial o legal de una zona o pais.

En cuanto al astro, toda su información será relativa a las coordenadas, que pueden ser: horizontales (Acimut A y distancia cenital z), horarias (Angulo horario H y declinación δ), absolutas (ascensión recta α y de nuevo declinación) y eclípticas (longitud eclíptica λ y latitud eclíptica β). Si prescindimos de la precesión, que no ha sido tenida en cuenta en esta primera versión de `Hiparco.m`, las coordenadas $\alpha, \delta, \lambda, \beta$ dependen únicamente del astro observado, mientras que A, z y H dependen no solo del astro sino del observatorio y del instante de la observación.

La conexión entre (A, z) y (H, δ) puede realizarse a través de ϕ por resolución del triángulo esférico PZE, *Polo-Zenit-Astro* que permite no solo transformar unas coordenadas en otras sino obtener dos de estos elementos conocidos los otros tres, cualesquiera que sean.

El triángulo esférico PKE, *Polo-Polo eclíptica-Astro*, nos permite relacionar (α, δ) con (λ, β) a través de la oblicuidad de la eclíptica ϵ que ahora es una constante por despreciar la precesión. El valor de esta constante ϵ se puede introducir como dato en el sistema o bien se puede considerar su valor en la época J2000.0 o dejarlo como una incógnita. Al igual que en el caso anterior la resolución de un triángulo esférico nos permite obtener dos de estos parámetros conocidos los otros tres.

H y α se relacionan por medio del tiempo sidéreo θ que es un parámetro que caracteriza el instante y el observatorio. La relación entre los tres es lineal y la misma función construida para relacionar las latitudes nos permite conocer uno de estos tres parámetros a partir de los otros dos.

El último elemento a considerar es el instante de la observación. El instante de la observación puede descomponerse en otros dos elementos, *el dia y la hora* que pueden aparecer de manera simultánea o separada. La forma de expresar el instante puede ser a través del dia juliano JD, el año besseliano B, el año juliano J o simplemente dando la fecha del calendario y la hora. En

la mayoría de los problemas aparecerá únicamente la hora de la observación pues el dia solo se utilizará al considerar la precesión o bien para el cálculo del tiempo sidéreo medio de Greenwich a las cero horas de TU del correspondiente dia $GMST0$, dato que es necesario para relacionar el tiempo sidéreo con el solar. $GMST0$ puede ser introducido directamente en nuestro sistema o bien puede ser calculado si se conoce el dia.

La hora de una observación es un elemento que puede ser definido, bien como parámetro local, dependiente del instante y el observatorio, o bien como universal, dependiente solo del instante. La relación entre tiempos locales y universales se efectúa a través de la longitud del lugar λ_w por medio de ecuaciones lineales. Se han considerado los siguientes tiempos locales: tiempo sidéreo medio θ, tiempo solar verdadero H_s, tiempo solar medio H_m, hora civil H_c y tiempo oficial o legal. Los tiempos no locales considerados son: tiempo sidéreo medio en Greenwich $GMST$, tiempo sidéreo medio en Greenwich a las cero horas del dia $GMST0$ y tiempo universal TU.

Obviamente éstos no son todos los posibles parámetros o propiedades que caracterizan una observación pero son los más representativos y los que más frecuentemente aparecen en los problemas planteados a los alumnos de Astronomía de Posición. En posteriores versiones del paquete `Hiparco.m` seran añadidos nuevos elementos.

3. Herramientas auxiliares de *Mathematica* para la elaboración del paquete `Hiparco.m`

El objeto matemático más simple que el astrónomo encuentra en sus problemas es el ángulo. Sin embargo, dependiendo del parámetro que describe o del contexto en que se use, aparece expresado en diferentes formatos: grados ("g"), grados y minutos ("gm"), grados, minutos y segundos ("gms"), horas ("h"), horas y minutos ("hm"), horas, minutos y segundos ("hms"), radianes ("rad") y por último segundos ("seg") para ángulos pequeños.

Aunque *Mathematica* utiliza radianes como argumento en sus funciones trigonométricas, sus especiales características hacen posible la definición de un objeto que de manera global trate todo tipo de formato de ángulos. Para ello se define un ángulo como una expresión cuya cabecera puede ser `gms`, `hms`, `rad` o `seg`, con un argumento para `rad` y `seg` y uno, dos o tres para `gms` y `hms` lo que representará respectivamente los formatos "g", "gm", "gms", "h", "hm" y "hms".

Estas expresiones pueden ser introducidas como argumento en las funciones trigonométricas. Las inversas de las trigonométricas devuelven un número que representa radianes, por lo que se ha construido la función `angulo` que devuelve el valor del angulo en uno de los 8 formatos posibles, por defecto en formato "gms".

```
In[1]:=
    alfa = gms[90];
In[2]:=
    angulo[alfa]
```

Out[2]=

```
gms[90,0,0.]
```

In[3]:=

```
angulo[alfa, formato->"hm"]
```

Out[3]=

```
hms[6,0.]
```

In[4]:=

```
Sin[alfa]
```

Out[4]=

```
1.
```

Las relaciones que aparecen entre los distintos parámetros que determinan una observación suelen ser ecuaciones lineales de la forma $\sum_{i=0}^{n} a_i x_i = c$ donde x_i representa n angulos de los que $n-1$ son conocidos y uno desconocido.

Para su resolución hemos definido la función `sumaAngular[]` con tres argumentos. El primero representa una lista con los n coeficientes, el segundo otra lista con los n angulos, donde la incógnita vendrá representada por `gms[]`, el último será un número que contenga el término independiente c. La respuesta es una lista con el valor de los n ángulos. Si hay más de una incógnita esta función no toma ninguna acción.

In[5]:=

Ejemplo : gms[70] + 2 gms[20] - gms[] = 0

In[6]:=

```
sumaAngular[{1,2,-1},{gms[70],gms[20],gms[]},0] // angulo
```

Out[6]=

```
{gms[70,0,0],gms[20,0,0],gms[110,0,0]}
```

El último elemento que nos encontramos en estos problemas es el triángulo esférico, que será representado por una lista de tres listas de dos ángulos que representa respectivamente un lado y el ángulo opuestos: `{{a,A},{b,B},{c,C}}`. La función `resuelveTriangulo[]` intenta resolver este triángulo esférico y devuelve una lista vacía si no tiene solución o una lista con uno o dos triángulos esféricos según el número de soluciones

In[7]:=

```
tri = {{gms[20,40,22],gms[]},
       {gms[40,10,20],gms[]},
       {gms[60,45,13],gms[]}};
```

In[8]:=

```
resuelveTriangulo[tri] // angulo
```

Out[8]=

```
{{gms[20,40,22.],gms[2,33,37.42]},
 {gms[40,10,20.],gms[4,40,56.01]},
 {gms[60,45,13.],gms[173,39,39.88]}}
```

4. Características de Hiparco.m

La mayor parte de los programas de ordenador que tratan problemas astronómicos están formados por diferentes conjuntos de rutinas o funciones con todos los algoritmos necesarios en los diferentes problemas. El usuario de estos programas elige la rutina que resolverá su problema en función de los datos que posee.

Nuestro objetivo es construir un programa al que se puedan introducir los datos de que disponemos. Con estas informaciones el ordenador construye una base de datos que intentará completar relacionando los distintos elementos conocidos por medio de una serie de reglas predefinidas. La máquina, por tanto, no pretenderá con esos datos responder a una pregunta concreta, lo que equivaldría a elegir el algoritmo que la contestara, sino que unicamente buscará añadir el máximo número de datos a los que le proporcione el usuario.

Lo primero que debe hacer el usuario del programa `Hiparco.m` es informar al ordenador sobre las diferentes observaciones que aparecen en su problema a través de los tres elementos básicos que las componen *observatorio, astro e instante.* Las funciones `observatorio[]`, `astro[]` e `instante[]`, con un argumento que constituye un nombre con el que el usuario identifica el correspondiente elemento, crean tres listas que contienen los nombres de todos los observatorios, astros e instantes que aparecen en el problema. El programa `Hiparco.m`, usará la función `listaObservaciones` para crear la lista de observaciones a partir de todas las ternas de elementos formadas tomando un elemento de cada una de las tres listas anteriores.

En el problema planteado aparecen tres elementos: un observatorio que llamaremos `Obs`, una estrella que llamaremos `Est` y un instante en el que se efectua la observación que llamaremos `Ins`. Estos tres elementos constituirán una única observación.

In[1]:=

```
<<Hiparco.m
```

In[2]:=

```
observatorio[Obs];
instante[Ins];
astro[Est];
```

In[3]:=

```
listaObservaciones
```

Out[3]=

```
{{Ast,Ins,Obs}}
```

El siguiente paso es comunicar a Hiparco las diferentes informaciones que poseemos sobre esos elementos. Estas informaciones pueden estar asociadas a uno, dos o tres elementos de la observación. Para cada observación se construye una base de datos con los parámetros que caracterizan cada uno de sus tres elementos. Dichos parámetros son inicialmete definidos como desconocidos por medio de la expresión gms[]. La base de datos de cada observación estará formada por los elementos que aparecen en la tabla

latitudGeocentrica[Obs]	ψ
latitudGeografica[Obs]	ϕ
latitudReducida[Obs]	u
longitudObservatorio[Obs]	λ_w
longitudZona[Obs]	λ_z
ascensionRecta[Est]	α
declinacion[Est]	δ
longitudEcliptica[Est]	λ
longitudEcliptica[Est]	β
epoca[Ins]	
dia[Ins]	
tiempoUniversal[Ins]	TU
tiempoSidereoGreenwich[Ins]	$GMST$
tiempoSidereoGOTU[Ins]	$GMST$ a las $0^h TU$
tiempoSidereo[Ins,Obs]	θ
horaCivil[Ins,Obs]	H_c
tiempoOficial[Ins,Obs]	TZ
acimut[Ast,Ins,Obs]	A
distanciaCenital[Ast,Ins,Obs]	z
anguloHorario[Ast,Ins,Obs]	H
oblicuidadEciptica[]	ϵ

donde las funciones con más de un argumento no precisan que estos sean introducidos en un orden prefijado.

En el problema que nos ocupa tendremos que poner:

```
In[4]:=

    ascensionRecta[Est]            = hms[19,45,35.13];
    declinacion[Est]               = gms[-22,53,42.2];
    tiempoUniversal[Ins]           = hms[5];
    tiempoSidereoGOTU[Ins]         = hms[17,37,39.352];
    distanciaCenital[Ast,Ins,Obs] = hms[0];
    acimut[Ast,Ins,Obs]            = gms[0];
```

donde los dos últimos datos corresponden a la observación en el zenit de la estrella.

Por último, cuando todos los datos han sido introducidos, la función `resuelveProblema[]` intenta completar, para cada observación , la base de datos de ésta , esto es, dar valor a todos los elementos de la tabla usando para ello los datos conocidos y las relaciones mencionadas en la sección 2. En este caso hay una única observación. En otros casos se realizará este proceso para cada observación.

```
In[5]:=

    resuelveProblema[];
```

Ya solo queda al usuario preguntar al ordenador por los diferentes elementos que desea conocer y este contestará con su valor o bien con `gms[]` si no ha podido deducirlo.

```
In[6]:=

    longitudObservatorio[Obs]

Out[6]=

    hms[2,52,53.5]

In[7]:=

    latitudGeografica[Obs]

Out[7]=

    gms[-22,53,42.2]
```

5. Limitaciones y futuras versiones del sistema `Hiparco.m`

Aunque se ha dicho que el sistema Hiparco recorre todas las observaciones del problema completando hasta donde se pueda la base de datos de cada una, la realidad es que la versión actual (versión 1.0) tiene algunas limitaciones cuando tenemos más de una observación.

Con una única observación el sistema completa toda la información posible a partir de los datos poseidos. Sin embargo, cuando aparecen varias observaciones distintas, un experto humano podría extraer más información de los datos que la obtenida por `Hiparco.m`. Esto es así porque en esta versión se analizan por separado las diferentes observaciones, y no se tienen en cuenta las posibles relaciones entre dos observaciones distintas, lo que permitiría extraer información adicional imposible de obtener de otro modo.

Para ilustrar la anterior idea, pensemos por ejemplo en dos observaciones que tengan en común el astro y el instante, esto es, dos observaciones simultáneas de un astro desde dos observatorios. Si se intenta completar por separado cada observación, deben intentar resolverse con los datos disponibles los triángulos PZ_1E y PZ_2E. Si sólo disponemos de dos elementos en cada caso será imposible resolver éstos. Sin embargo, relacionando ambas observaciones aparecen no dos sino cuatro triángulos esféricos combinando de tres en tres los puntos P, Z_1, Z_2, E, por lo que es posible que uno de estos cuatro triángulos nos de la clave para la obtención del tercer elemento necesario para completar los otros triángulos.

De la misma forma podemos combinar observaciones con dos astros o dos instantes comunes, lo que nos dará información adicional en cada caso que completará la ahora disponible.

Es nuestra intención ampliar el conjunto de reglas del sistema `Hiparco.m` en futuras versiones, para contemplar las extensiones ya mencionadas. Asímismo pretendemos incorporar otras que vayan apareciendo al realizar un análisis más detallado de los distintos casos que se pueden presentar en la resolución de estos problemas.

Agradecimientos

Este trabajo ha sido financiado parcialmente por la C.I.C.Y.T. (Proyectos PB90-0921 y ESP91-0919).

Los autores están mencionados por orden alfabético.

Referencias

[1] Aubert, J. P., Schomberg, R.:1986, *Inteligencia Artificial* Ed. Paraninfo. Madrid.

[2] Cid, R.:1970, *Curso de Astronomía.* Universidad de Zaragoza.

[3] Docobo, J.A., Elipe, A.:1983, *ASTRONOMIA. 280 problemas resueltos.* Imprenta de la Universidad de Santiago de Compostela.

[4] Green, R.M.:1985, *Spherical Astronomy* Cambridge University Press. London

[5] Wolfran, S. : 1988, *Mathematica. A System for Doing Mathematics by Computer.* Addison Wesley Publishing Company.

Rev. Academia de Ciencias. Zaragoza 47 (1992)

Scaling hamiltonians in attitude dynamics of two rigid bodies

Elipe, A.[1], Abad, A.[1], Arribas, M.[1,2]

[1]Grupo de Mecánica Espacial. University of Zaragoza, 50009 Zaragoza, Spain

[2]Department of Applied Mathematics, University of Zaragoza, 50014 Zaragoza, Spain.

Abstract

When perturbation methods are applied to the coupled translational-rotational motion of two rigid bodies, there are several possibilities in defining the different terms which constitute the Hamiltonian form; even the zero order may be different depending on the various hypotheses. Usually in the literature, authors choose the model and in most of the cases with no justification of the choice. In this paper, by means of the scaling technique, we present the Hamiltonian depending on several parameters that will determine the different orders for the models considered without hesitation.

1. Introduction

In the problem of the coupled orbital-rotational motion of two rigid bodies, there are several factors, such as the relation among the size of the orbit and the sizes of the rigid bodies, the spining of the bodies (fast or slow rotations) with respect the mean orbital motion, the masses of the bodies, etc., that offer several possibilities for determining the different orders which compose the Hamiltonian, which is essential when perturbation methods are applied.

Usually, in the literature, authors assign the different orders by hypothesis, according with their abilities for solving the problem and in general not in a clear manner (1 to 4). Indeed, denoting by

$\mathcal{H}_k$ the Hamiltonian of the Kepler motion of the centers of mass,

$\mathcal{H}_{Ei}$ the rotational kinetic energy of the rigid body B_i, and

$\mathcal{H}_c$ the remaining coupled terms of the potential,

several possiblities are found, without much explanation, such as:

$$\begin{aligned} &\mathcal{H}_0 = \mathcal{H}_k + \mathcal{H}_{E1} + \mathcal{H}_{E2}, && \text{and} \quad \mathcal{H}_1 = \mathcal{H}_c \\ &\mathcal{H}_0 = \mathcal{H}_k, && \text{and} \quad \mathcal{H}_1 = \mathcal{H}_{E1} + \mathcal{H}_{E2} + \mathcal{H}_c \\ &\mathcal{H}_0 = \mathcal{H}_k, \mathcal{H}_1 = \mathcal{H}_{E1}, && \text{and} \quad \mathcal{H}_2 = \mathcal{H}_{E2} + \mathcal{H}_c. \end{aligned}$$

Within this cumbersome treatment, some exceptions are found in the literature. For instance, Cochran (5) provides a good determination of the perturbation orders of the terms which form the Hamiltonian by him considered: the attitude motion of a satellite moving in a precessing Keplerian orbit around a point mass.

However, after the excellent work of Meyer (**6**) for scaling Hamiltonian systems, the different orders may be obtained in a systematic way. Following his guidelines, when the problem of the coupled orbital–attitude dynamics is considered, we derive several small parameters, which will allow us to determine the order of magnitude of each term without any doubt.

Following Meyer's notation, (**6**)[p. 879], when we make a change of variables or a scaling of the $q_i = \alpha \bar{q}_i$, since the bars over variables are not esthetic, it is convenient to drop them. The operation of first changing the variables $q_i = \alpha \bar{q}_i$, and then dropping the bars, will be denoted by $q_i \longrightarrow \alpha q_i$

It may appear as a paradox the fact of building the scale with Lagrangian instead of Hamiltonian formalism. The reason is simple: it is easier handling operations such as multiply by factors or change the scale of the independent variable with Lagrangians that with Hamiltonians, since in this case, some additional operations must be done with the canonical moments (for details, see Stiefel and Scheifele 1971 (**7**)).

2. Basics

Let us consider the problem of the motion of two rigid bodies S and s attracted by their mutual gravitation forces. Let M, m their respective masses, and P, p their center of mass. In order to describe the rotational and orbital motion, we consider three frames of reference:

Space frame: $O\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3$, a fixed or inertial coordinate system, and

Two body frames: $P\,\boldsymbol{B}_1, \boldsymbol{B}_2, \boldsymbol{B}_3$ and $p\,\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$, the principal frames of inertia of the bodies S and s.

Thereafter, symbols in capital letters stand for the rigid body S, whereas lowercases designate the same concept for the body s.

Let us designate by $\boldsymbol{X}$ and $\dot{\boldsymbol{X}}$ the position and velocity of S with respect to the Space frame. These vectors describe the orbital motion of S in the space frame. The attitude of S is given by means of the Eulerian angles Φ, Θ, Ψ, and the angular velocity vector $\boldsymbol{\Omega} = \Omega_1 \boldsymbol{B}_1 + \Omega_2 \boldsymbol{B}_2 + \Omega_3 \boldsymbol{B}_3$, or the angular momentum of rotation vector $\boldsymbol{G} = I_1\Omega_1\boldsymbol{B}_1 + I_2\Omega_2\boldsymbol{B}_2 + I_3\Omega_3\boldsymbol{B}_3$, where I_1, I_2, I_3 are the principal moments of inertia of the rigid body S.

Taking in consideration the three first terms in the development of the potential function, the Lagrangian $\mathcal{L}$ of the problem is:

$$
\begin{aligned}
\mathcal{L} \;=\; & \tfrac{1}{2}M(\dot{\boldsymbol{X}}\cdot\dot{\boldsymbol{X}}) + \tfrac{1}{2}m(\dot{\boldsymbol{x}}\cdot\dot{\boldsymbol{x}}) + \frac{k^2 M\, m}{\|\boldsymbol{X}-\boldsymbol{x}\|} \\
& + \tfrac{1}{2}\left(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2\right) \\
& + \tfrac{1}{2}\left(i_1\omega_1^2 + i_2\omega_2^2 + i_3\omega_3^2\right) \\
& + \frac{k^2 m}{2\|\boldsymbol{x}_1-\boldsymbol{x}_2\|^3}(I_1 + I_2 + I_3 - 3I) \\
& + \frac{k^2 M}{2\|\boldsymbol{x}_1-\boldsymbol{x}_2\|^3}(i_1 + i_2 + i_3 - 3i),
\end{aligned}
\tag{1}
$$

where I is the moment of inertia of the body S in the direction $\boldsymbol{X} - \boldsymbol{x}$. This direction may be expressed as

$$\boldsymbol{X} - \boldsymbol{x} = \|\boldsymbol{X} - \boldsymbol{x}\| (A_1 \boldsymbol{B}_1 + A_2 \boldsymbol{B}_2 + A_3 \boldsymbol{B}_3)$$

with A_1, A_2, A_3 the direction cosines of this direction with respect to the body frame $P\,\boldsymbol{B}_1, \boldsymbol{B}_2, \boldsymbol{B}_3$.

The moment of inertia I is:

$$I = I_1 A_1^2 + I_2 A_2^2 + I_3 A_3^2$$

and it depends on the Eulerian angles, and on the orbital vectors $\boldsymbol{X}$, $\boldsymbol{x}$, but only through the difference $\boldsymbol{X} - \boldsymbol{x}$,

$$I \equiv I(\boldsymbol{X} - \boldsymbol{x}, \Phi, \Theta, \Psi).$$

Therefore, the Lagrangian function is

$$\begin{aligned} \mathcal{L} \;\equiv\; & \mathcal{L}(\boldsymbol{X}, \boldsymbol{x}, \Phi, \Theta, \Psi, \phi, \vartheta, \psi, \\ & \dot{\boldsymbol{X}}, \dot{\boldsymbol{x}}, \dot{\Phi}, \dot{\Theta}, \dot{\Psi}, \dot{\phi}, \dot{\vartheta}, \dot{\psi}) , \end{aligned}$$

and consequently, the problem is of order 24, or equivalently, the system has 12 degrees of freedom.

2.1 Reduction to the barycenter

In order to reduce the order of the problem, we perform a reduction to the barycenter. Denoting by the vector $\boldsymbol{b}$ the position of the barycenter of the system, we introduce the vectors $\boldsymbol{b}$ and $\boldsymbol{r}$ such that

$$(M + m)\boldsymbol{b} = M\boldsymbol{X} + m\boldsymbol{x}, \quad \boldsymbol{r} = \boldsymbol{x} - \boldsymbol{X}.$$

Inversely, there follows that

$$\boldsymbol{X} = \boldsymbol{b} - \frac{m}{M+m}\boldsymbol{r}, \quad \boldsymbol{x} = \boldsymbol{b} + \frac{M}{M+m}\boldsymbol{r}.$$

With this, the first row of (1) becomes

$$\tfrac{1}{2}(M+m)(\dot{\boldsymbol{b}} \cdot \dot{\boldsymbol{b}}) + \frac{1}{2}\frac{M\,m}{M+m}(\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}}) + \frac{k^2 M\,m}{\|\boldsymbol{r}\|}$$

and because $\mathcal{L}$ does not depend on $\boldsymbol{b}$, it is ignorable, and the reduced Lagrangian, after dropping the constant quantity $(M+m)(\dot{\boldsymbol{b}} \cdot \dot{\boldsymbol{b}})/2$ is

$$\begin{aligned} \mathcal{L} \;=\; & \frac{1}{2}\frac{M\,m}{M+m}(\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}}) + \frac{k^2 M\,m}{\|\boldsymbol{r}\|} \\ & + \tfrac{1}{2}\boldsymbol{\Omega} \cdot \boldsymbol{G} + \frac{k^2 m}{2\|\boldsymbol{r}\|^3}(I_1 + I_2 + I_3 - 3I) \\ & + \tfrac{1}{2}\boldsymbol{\omega} \cdot \boldsymbol{g} + \frac{k^2 M}{2\|\boldsymbol{r}\|^3}(i_1 + i_2 + i_3 - 3i) \end{aligned} \tag{2}$$

and at this point, we recognize that the problem has 9 degrees of freedom.

One of the advantages of the Lagrangian formalism, is that the equations of the motion are not modified when the whole Lagrangian function is multiplied by a constant. This fact allows us to make the change

$$\mathcal{L} \longrightarrow \frac{M+m}{M\,m}\mathcal{L}. \tag{3}$$

With this transformation,

$$\begin{aligned}\mathcal{L} \;=\; & \tfrac{1}{2}(\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}}) + \frac{\mu}{\|\boldsymbol{r}\|} \\ & + \frac{1}{2}\frac{M+m}{M\,m}\boldsymbol{\Omega}\cdot\boldsymbol{G} + \frac{\mu}{2M\|\boldsymbol{r}\|^3}(I_1+I_2+I_3-3I) \\ & + \frac{1}{2}\frac{M+m}{M\,m}\boldsymbol{\omega}\cdot\boldsymbol{g} + \frac{\mu}{2m\|\boldsymbol{r}\|^3}(i_1+i_2+i_3-3i)\end{aligned} \tag{4}$$

where $\mu = k^2(M+m)$ and its dimension is $[\mu] = L^3/T^2$.

Because both masses m and M appear in the denominators, and also in the definition of the moments of inertia, we define the principal *pseudo-moments* of inertia given by the relation

$$i_\nu = m\,j_\nu \text{ and } I_\nu = M\,J_\nu \qquad (1 \le \nu \le 3)$$

and a two *pseudo-moments* of inertia $I = M\,J$ and $i = m\,j$. The dimension of the *pseudo-moments* of inertia is $[J_\nu] = L^2$. The vectors analogous to the rotation angular momentum vectors $\boldsymbol{G}$ and $\boldsymbol{g}$ are now $\boldsymbol{\Gamma}$ and $\boldsymbol{\gamma}$, with $\boldsymbol{G} = M\,\boldsymbol{\Gamma}$ and $\boldsymbol{g} = M\,\boldsymbol{\gamma}$

The Lagrangian is now

$$\begin{aligned}\mathcal{L} \;=\; & \tfrac{1}{2}(\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}}) + \frac{\mu}{\|\boldsymbol{r}\|} \\ & + \frac{1}{2}\frac{M+m}{M}\boldsymbol{\Omega}\cdot\boldsymbol{\Gamma} + \frac{\mu}{2\|\boldsymbol{r}\|^3}(J_1+J_2+J_3-3J) \\ & + \frac{1}{2}\frac{M+m}{m}\boldsymbol{\omega}\cdot\boldsymbol{\gamma} + \frac{\mu}{2\|\boldsymbol{r}\|^3}(j_1+j_2+j_3-3j)\end{aligned} \tag{5}$$

Note that up this point, we have done only one scale: the one given by Eq. (3).

3. Asymptotic scale

Since the independent variable (the time t) does not appear explicitly, in the Lagrangian, the energy $\mathcal{H}$ is an integral of the motion. Using the parameter μ and the integral $\mathcal{H}$, we introduce two constants n and α such that

$$\mu = n^2\,\alpha^3, \qquad \mathcal{H} = \tfrac{1}{2}n^2\,\alpha^2.$$

With this, we are going to define two scales: one of length α, and the other of time $1/n$, given by:

$$\boldsymbol{r} \;\longrightarrow\; \alpha\,\boldsymbol{r}, \tag{6}$$

$$t \;\longrightarrow\; \frac{1}{n}t. \tag{7}$$

In words of Meyer (**6**)[p. 879], scaling and changing units are essentially the same thing. This becomes clear in the preceding scales, since the first one (6), means that we chose as unit of

length the mean semi-major axis or the orbit which describe the centers of mass of the rigid bodies, whereas the second one (7) denotes that the unit of time is the inverse of the mean motion of this orbit, that is to say, almost the orbital period (in fact, the period / 2π).

With these scales, there results:

$$\frac{1}{2}(\dot{r}\cdot\dot{r})+\frac{\mu}{\|r\|}\longrightarrow\alpha^2 n^2\left[\frac{1}{2}(\dot{r}\cdot\dot{r})+\frac{\mu}{\|r\|}\right]$$

On the other hand, we know that the components of the angular velocity are related with the Euler angles and their time derivatives. Therefore, after the scaling, we have

$$\Omega\longrightarrow n\Omega,\quad \omega\longrightarrow n\omega \qquad \text{and} \qquad G\longrightarrow n\alpha^2 G,\quad g\longrightarrow n\alpha^2 g.$$

After the scaling, the Lagrangian becomes:

$$\begin{aligned}\mathcal{L}=n^2\alpha^2\Bigg[&\frac{1}{2}\dot{r}\cdot\dot{r}+\frac{1}{\|r\|}\\&+\frac{1}{2}\frac{M+m}{M}(\Omega\cdot G)+\frac{1}{2\|r\|^3}(J_1+J_2+J_3-3J)\\&+\frac{1}{2}\frac{M+m}{M}(\omega\cdot g)+\frac{1}{2\|r\|^3}(j_1+j_2+j_3-3j),\Bigg]\end{aligned}\tag{8}$$

and consequently, we may divide by the constant coefficient $n^2\alpha^2$ without any modification of the equations of the motion, obtaining

$$\mathcal{L}\longrightarrow\frac{1}{n^2\alpha^2}\mathcal{L}.$$

4. Small parameters

In order to establish the different orders of the terms which compose the Lagrangian, we introduce several dimensionless parameters. Whether or not these parameters should be considered as small parameters is determined by the conditions of the problem considered.

Since the unit of length after the scaling (6) is mean orbital semi-major axis, we introduce two parameters Δ_1, δ_1, given by

$$J\longrightarrow\Delta_1^2 J,\qquad j\longrightarrow\delta_1^2 j,$$

in order to have the size of the rigid bodies of the same magnitude than the unit of length. Roughly speaking, these parameters are a measure of the smallness of the rigid bodies.

Analogously, since the orbital period $(1/n)$ (7) has been chosen as unit of time, we designate two parameters Δ_2, δ_2, as

$$\Omega\longrightarrow\Delta_2^{-1}\Omega,\qquad \omega\longrightarrow\delta_2^{-1}\omega.$$

The parameters Δ_2, δ_2 reflect the speedness of the spining of the bodies (note that for fast rotations of the bodies, we have $\Delta_2,\delta_2\ll 1$).

Finally, we designate by Δ_3, δ_3 the ratio mass

$$\Delta_3 = \frac{M+m}{m}, \qquad \delta_3 = \frac{M+m}{M}.$$

After these definitions, the Lagrangian may be written as

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}} + \frac{1}{\|\boldsymbol{r}\|} \\ & + \frac{1}{2}\frac{\Delta_1^2}{\Delta_2^2}\Delta_3\boldsymbol{\Omega}\cdot\boldsymbol{G} + \Delta_1^2\frac{1}{2\|\boldsymbol{r}\|^3}(J_1+J_2+J_3-3J) \\ & + \frac{1}{2}\frac{\delta_1^2}{\delta_2^2}\delta_3\boldsymbol{\omega}\cdot\boldsymbol{g} + \delta_1^2\frac{1}{2\|\boldsymbol{r}\|^3}(j_1+j_2+j_3-3j). \end{aligned} \tag{9}$$

5. Application to classical cases

In order to illustrate the theory above exposed, let us mention several classical cases.

5.1 One of the masses is much bigger than the other

Let us suppose that the mass of one of the bodies (S) is much bigger than the other ($M \gg m \implies \delta_3 \approx 1$), in such a way that the rotation of the more massive is independent of the other body. This corresponds to the so named *satellite case* (**8**). Under these conditions, the Hamiltonian corresponding to the Lagrangian (9) is

$$\begin{aligned} \mathcal{H} = & \frac{1}{2}\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}} - \frac{1}{\|\boldsymbol{r}\|} \\ & + \frac{1}{2}\frac{\delta_1^2}{\delta_2^2}\boldsymbol{\omega}\cdot\boldsymbol{g} - \delta_1^2\frac{1}{2\|\boldsymbol{r}\|^3}(j_1+j_2+j_3-3j). \end{aligned} \tag{10}$$

N.B. The same result is obtained under the hypothesis that one of the bodies has spherical symmetry of inertia ($J_1 = J_2 = J_3$).

Two cases may be taken into consideration:

- $\delta_1^2 = \delta_2^2 = \epsilon$

  This case corresponds to very fast rotations of the rigid body s (satellite), in such a way that the rotational kinetic energy is of the same order than the Keplerian energy. The Hamiltonian now has the form:

  $$\mathcal{H} = \mathcal{H}_0 + \epsilon\mathcal{H}_1,$$

  with

  $$\mathcal{H}_0 = \frac{1}{2}\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}} - \frac{1}{\|\boldsymbol{r}\|} + \frac{1}{2}\boldsymbol{\omega}\cdot\boldsymbol{g}$$

  $$\mathcal{H}_1 = -\frac{1}{2\|\boldsymbol{r}\|^3}(j_1+j_2+j_3-3j).$$

  With this Hamiltonian, we recover the problem considered by several authors (**3**, **4** and **9**).

$\delta_1 = \delta_2^2 = \epsilon$

This case corresponds to fast rotations of the satellite, and orbital dimensions much bigger that the size of the satellite, such that the rotational kinetic energy is of the same order than the Keplerian energy. The Hamiltonian now has the form:

$$\mathcal{H} = \mathcal{H}_0 + \epsilon\mathcal{H}_1 + \epsilon^2\mathcal{H}_2,$$

with

$$\mathcal{H}_0 = \frac{1}{2}\dot{r}\cdot\dot{r} - \frac{1}{\|r\|}$$

$$\mathcal{H}_1 = \frac{1}{2}\omega\cdot g$$

$$\mathcal{H}_2 = -\frac{1}{2\|r\|^3}(j_1 + j_2 + j_3 - 3j).$$

An usual additional hypothesis is that the rotation of the satellite does not affect the orbital motion, which is considered as a known function of the time (**5, 8** and **10**). In this case, the zero order may be dropped and the Hamiltonian becomes:

$$\mathcal{H} = \mathcal{H}_0 + \epsilon\mathcal{H}_1$$

$$\mathcal{H}_0 = \frac{1}{2}\omega\cdot g$$

$$\mathcal{H}_1 = -\frac{1}{2\|r\|^3}(j_1 + j_2 + j_3 - 3j).$$

5.2 The two masses are similar

When the masses of both bodies are similar, the parameters Δ_3, δ_3 are of the same magnitude than the factors to which they are multiplying, and therefore, their role is not relevant in the equations of the motion. Alike the previous subsection, several cases may be considered, depending on the magnitude of the parameters Δ_ν, δ_ν, $\nu = 1, 2$, which offers more possibilities than in the previous section. Nevertheless, in general, under the enormous complexity of the problem (9 degrees of freedom), most of the authors consider one of the rigid body spherical (**1, 11**). Furthermore, additional hypotheses of axial symmetry or near-axial symmetry for the other body are stablished. Under the consideration of sphericity for one of the rigid bodies, the treatment of this problem does not differ substatially from the considered in the previous subsection.

6. Conclusions

Scaling technique reveals very useful in the coupled rotational-traslational motion problem of rigid bodies. When perturbations theories are applied, the determination of the different orders of the Hamiltonian is essential. By means of the scaling, the orders are determined without ambiguity, and several classical formulations are recovered.

Acknowledgements

Support came in part from the C.I.C.Y.T. (Projects # PB90-0921 and ESP91-0919).

REFERENCES

1. H. KINOSHITA, *Pub. Astron. Soc. Japan* **24** (1972) 423–457.

2. J. M. FERRANDIZ, *Movimiento de dos sólidos sometidos a atracción Newtoniana. Aplicaciones* Ph.D. diss. (Universidad de Zaragoza 1979).

3. M. C. ZANARDI, *Rev. Mexicana Astron. Astrof.* **10** (1985) 355–366.

4. M. C. ZANARDI, *Celes. Mech.* **39** (1986) 147–158.

5. J. COCHRAN, *Celes. Mech.* **6** (1972) 131–152.

6. K. R. MEYER, *SIAM J. Math. Anal.* **15** (1984) pp. 877–889.

7. E. L. STIEFEL and G. SCHEIFELE, *Linear and Regular Celestial Mechanics* (Springer-Verlag, Berlin 1971).

8. E. LEIMANIS, *The General Problem of the Motion of Coupled Rigid Bodies about a Fixed Point.* (Springer-Verlag, Berlin 1965).

9. M. ARRIBAS and A. ELIPE, *Advances in the Astronautical Sciences*, **75** (1991) 333-344.

10. M. ARRIBAS and A. ELIPE, *Celes. Mech.* (1992) submitted

11. H. KINOSHITA, *Pub. Astron. Soc. Japan* **22** (1970) 383–403.

Rev. Academia de Ciencias. Zaragoza 47 (1992)

INTEGRACION ANALITICA DE UN CASO PARTICULAR DEL PROBLEMA DEL MOVIMIENTO ROTACIONAL DE UN GIROSTATO CON UN PUNTO FIJO EN UN CAMPO CENTRAL NEWTONIANO

J.A. Cavas y A. Vigueras
Departamento de Matemática Aplicada y Estadística
Escuela Politécnica Superior de Cartagena. Universidad de Murcia

ABSTRACT. The scope of the present paper is to provide analytic solutions to the problem of the attitude evolution of a symmetric gyrostat about a fixed point in a central Newtonian force field of potential $V^{(2)}$, in an analogous case of that of Lagrange and Poisson when the third component of the total angular momentum is cancelled out ($B_1 = 0$). We consider the possible cases of this problem, which depend upon the nature of the roots of the function g(u) and we obtain the analytic solutions for Euler angles in terms of elliptic functions of the time t. Finally, we give a geometric interpretation of the obtained results.

1. INTRODUCCION.

El problema del movimiento rotatorio de un sólido rígido con un punto fijo, sometido a momentos externos, ha sido estudiado en numerosos artículos por autores como Euler, Jacobi, Poinsot, Lagrange, Poisson, Kowalesky, por citar tan sólo los más clásicos; habiendo obtenido integraciones analíticas en distintos casos particulares.

Con objeto de poner de manifiesto la influencia de los movimientos internos de los cuerpos, que no modifican su distribución de masas, sobre la rotación de los mismos, se ha estudiado también el movimiento rotatorio de giróstatos con un punto fijo. En particular, de giróstatos estacionarios, que son aquellos para los que el momento angular relativo de su parte móvil con respecto a su parte rígida (también denominado momento

girostático) es constante en el sistema móvil, solidario con su parte rígida. El problema del giróstato con un punto fijo es algo más general que el del sólido rígido (al que se reduce si el momento girostático se anula), y ha sido considerado entre otros autores, por G. Peano, V. Volterra, V.A. Steklov, U. Cassina, M.T. Vacca en trabajos dedicados, preferentemente, al estudio del movimiento de los polos terrestres, variación de la latitud sobre la superficie de la tierra, etc.; estos y otros son referenciados en el libro de E. Leimanis (1965).

Refiriéndonos al problema del sólido rígido con un punto fijo en un campo central newtoniano de potencial $V^{(2)}$ (obtenido tomando solamente hasta el término principal del potencial no Kepleriano, es decir, hasta los términos de tercer orden en 1/r), Arkhangelskii (1962, 63) ha probado que, en general, existe una cuarta integral primera algebraica de las ecuaciones del movimiento solamente en los casos de esfericidad, el análogo al de Lagrange - Poisson y el análogo al de Euler - Poinsot.

El problema del giróstato pesado con un punto fijo ha sido considerado por Keis (1964), que confirma la validez del teorema de Poincaré para este caso; y por Vigueras (1983) que integra analíticamente un caso análogo al de Lagrane - Poisson para giróstatos. En cuanto al problema del movimiento de un giróstato con un punto fijo en un campo central newtoniano de potencial $V^{(2)}$, en variables de Andoyer, es considerado en Tsopa (1979, 81) e integrado en casos próximos a regulares por métodos de perturbaciones. Problemas más generales, tales como el problema del movimiento rototraslatorio de n giróstatos, el del movimiento rototraslatorio de un giróstato en un campo central newtoniano, la rotación terrestre usando como modelo de tierra un giróstato simétrico han sido considerados en R. Cid y A. Vigueras (1985, 90) y M.E. Sansaturio y A. Vigueras (1988), donde se dan resultados cualitativos e integraciones aproximadas.

Volviendo al problema del movimiento de un giróstato con un punto fijo en un campo central newtoniano de potencial $V^{(2)}$, en Vigueras (1987) se prueba la integrabilidad en el sentido de Liouville en tres casos, que se reducen a los dados por Arkhangelskii si el momento girostático se anula.

En este trabajo, se integra analíticamente uno de los casos integrables encontrados, el caso análogo al de Lagrange - Poisson, con las condiciones iniciales:

$$\psi(t_0) = \psi_0 \qquad \theta(t_0) = \theta_0 \qquad \phi(t_0) = \phi_0$$
$$\omega_1(t_0) = 0 \qquad \omega_2(t_0) = 0 \qquad \omega_3(t_0) = \omega_3^0$$

y suponiendo que la tercera componente del momento angular total se anula (lo cual podría lograrse eligiendo convenientemente la tercera componente del momento girostático, pensamos, por ejemplo, en satélites artificiales dotados de rotores simétricos). Se consideran todos los subcasos posibles, que se caracterizan en función de la distancia r del centro de atracción P al punto fijo O, así como de la geometría del cuerpo y de las condiciones iniciales, dando la expresión analítica explícita de la solución en cada uno de ellos. Con independencia de las soluciones de equilibrio, aparecen ahora más posibilidades de movimiento que en el caso análogo para un giróstato pesado, lo cual es debido a los efectos derivados de los momentos gravitatorios correspondientes al nuevo término incluido en el potencial $V^{(2)}$.

2. PLANTEAMIENTO DEL PROBLEMA.

El problema que nos planteamos consiste en obtener las **ecuaciones del movimiento de un giróstato fijo por uno de sus puntos O**, perteneciente a su parte rígida, sometido a la atracción newtoniana de otro punto fijo P, cuando el movimiento relativo de S_2 con respecto a S_1 se supone conocido y el potencial V es aproximado por $V^{(2)}$. Más concretamente, suponemos que el giróstato es simétrico y tanto el momento girostático $\vec{l}_r$ como el vector de posición del c.d.m. en el sistema móvil, $\vec{r}_0$, son vectores constantes que están sobre el eje de simetría, es decir:

$$I_1 = I_2 \quad ; \quad a_1 = a_2 = 0 \quad ; \quad x_1^0 = x_2^0 = 0$$

como sabemos, las ecuaciones del movimiento en el sistema móvil son:

$$\dot{I}_1\omega_1 + (I_3-I_1)\omega_2\omega_3 + \omega_2 a_3 = m_0 x_3^0 k_2 + m_1(I_3-I_1)k_2k_3$$

$$\dot{I}_1\omega_2 + (I_1-I_3)\omega_1\omega_3 - \omega_1 a_3 = -m_0 x_3^0 k_1 + m_1(I_1-I_3)k_1k_3$$

$$\dot{I}_3\omega_3 = 0$$

$$\dot{\vec{k}} + \vec{\omega} \wedge \vec{k} = 0$$

siendo $m_0 = mg$, $m_1 = 3g/r$, m = masa del giróstato, $g = GM/r^2$ G = atracción de la gravedad, M = masa del punto fijo P, r = distancia del centro de atracción P (situado en la parte negativa del tercer eje fijo) al punto fijo O. Según se sigue del teorema del momento angular y de la ecuación cinemática de Poisson, admitiendo las integrales primeras:

$$k_1^2 + k_2^2 + k_3^2 = 1$$

$$\vec{k}.(\vec{l} + \vec{l}_r) = c \text{ (cte)}$$

$$\frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2) + mg(x_1^0 k_1 + x_2^0 k_2 + x_3^0 k_3) +$$

$$+ \frac{3}{2}\frac{g}{r}(I_1k_1^2 + I_2k_2^2 + I_3k_3^2) = h \text{ (cte)}$$

$$\omega_3^0 = \omega_3 \text{ (cte)}$$

En particular, la tercera ecuación (integral de Jacobi) y la segunda (integral que expresa el caracter constante de la proyección del momento angular total sobre el tercer eje fijo) se pueden expresar en la forma:

$$\begin{aligned} \omega_1^2 + \omega_2^2 &= A_0 + A_1k_3 + A_2k_3^2 \\ \omega_1k_1 + \omega_2k_2 &= B_0 - B_1k_3 \end{aligned} \qquad (2.1)$$

siendo A_0, A_1, A_2, B_0 y B_1 constantes dadas por las fórmulas:

$$A_o = \left[2h_0 - I_3(\omega_3^0)^2\right] \Big/ I_1 \quad , \quad A_1 = -2m_0 x_3^0 \Big/ I_1$$

$$A_2 = 3g(I_1 - I_3) \Big/ I_1 r$$

$$B_0 = c \Big/ I_1 \quad , \quad B_1 = \left[I_3\omega_3^0 + a_3\right] \Big/ I_1$$

Utilizando las expresiones de las componentes de $\vec{\omega}$ y $\vec{k}$ en variables de Euler, las integrales (1) se reducen a las ecuaciones:

$$\begin{aligned} &\dot{\psi}^2 \operatorname{sen}^2\theta + \dot{\theta}^2 = A_0 + A_1\cos\theta + A_2\cos^2\theta \\ &\dot{\psi}\operatorname{sen}^2\theta = B_0 - B_1\cos\theta \end{aligned} \qquad (2.2)$$

El cambio de variable $u = \cos\theta$, permite separar variables, resultando:

$$\begin{aligned} \dot{u}^2 &= (A_0 + A_1 u + A_2 u^2)(1 - u^2) - (B_0 - B_1 u)^2 = g(u) \\ \dot{\psi} &= (B_0 - B_1 u)/(1 - u^2) \\ \dot{\phi} &= \omega_3^0 - (B_0 - B_1 u)\, u /(1 - u^2) \end{aligned} \qquad (2.3)$$

2.1 Hipótesis adicionales:

Antes de proseguir, supongamos que para $t = t_0$ se tengan las siguientes condiciones iniciales:

$$\begin{aligned} &\psi(t_0) = \psi_0 \qquad \theta(t_0) = \theta_0 \qquad \phi(t_0) = \phi_0 \\ &\omega_1(t_0) = 0 \qquad \omega_2(t_0) = 0 \qquad \omega_3(t_0) = \omega_3^0 \end{aligned} \qquad (2.4)$$

(siendo $u_0 = \cos\theta_0$ distinto de 1 y -1, puesto que dichos valores coinciden con sendas soluciones de equilibrio del problema).
Es decir, vamos a proceder a la integración analítica del problema considerado cuando el giróstato (su parte rígida) gira, en el instante inicial, alrededor de su eje de simetría con velocidad angular constante ω_3^0; siendo la orientación del mismo arbitraria.

Entonces de (2.1) resultan $A_0 = -(A_1 u_0 + A_2 u_0^2)$, $B_0 = B_1 u_0$, con lo cual $\dot{\psi}$ y $\dot{\phi}$ quedan factorizadas por B_1.

Si además suponemos que el centro de masas está por encima del punto fijo, $x_3^0 > 0$, y que $I_1 > I_3$ así como que $B_1 = 0$, es decir, se anula la 3ª componente del momento angular total del giróstato (lo cual podría lograrse por una adecuada elección de los movimientos internos) y el centro de atracción P no es coincidente con el punto fijo 0. Las ecuaciones del movimiento vendrían expresadas por:

$$\begin{aligned} \dot{u}^2 &= (u - u_0)(1 - u^2)(A_1 + A_2 u_0 + A_2 u) = g(u) \\ \dot{\psi} &= 0 \qquad \Rightarrow \psi(t) = \psi_0 \\ \dot{\phi} &= \omega_3^0 \qquad \Rightarrow \phi(t) = \phi_0 + \omega_3^0 (t - t_0) \end{aligned} \tag{2.5}$$

Con estas últimas condiciones que le hemos impuesto, vemos que $A_2 > 0$ y $A_1 < 0$, e introduciendo $u_1 = -u_0 - \dfrac{A_1}{A_2}$ tenemos que :

$$u_1 = -u_0 + \frac{2mx_3^0 r}{3(I_1 - I_3)} = -u_0 + \frac{r}{a_0} > -u_0 \tag{2.6}$$

donde $$a_0 = \frac{3(I_1 - I_3)}{2mx_3^0} > 0 \tag{2.7}$$

es un parámetro que depende de la distribución de masas del cuerpo, pudiendo entonces factorizar

$$g(u) = A_2(1 - u^2)(u - u_0)(u - u_1) \tag{2.8}$$

3. RESOLUCION ANALITICA: CASOS POSIBLES.

Nos queda, pues, expresar la nutación θ en función del tiempo a partir de la primera de las expresiones (2.5), para tener el problema completamente resuelto.

Los distintos casos que se nos pueden presentar vendrán dados por la posición relativa de la raiz u_1 de $g(u) = 0$, estos son:

3.1 Caso: $u_1 < 1 \quad \Leftrightarrow \quad r < (1 + u_0)a_0$

Estamos suponiendo que $-u_0 < u_1 < 1$, con lo cual $\dot{u}^2 = g(u) = A_2(1 - u^2)(u - u_0)(u - u_1)$ caben, entonces, tres posibilidades: **a)** Que $u_1 < u_0$, **b)** Que $u_1 > u_0$, **c)** Que $u_1 = u_0$ que se corresponden en terminos de r con: **a')** $0 < r < 2a_0\cos\theta_0$, **b')** $2a_0\cos\theta_0 < r < (1 + \cos\theta_0)a_0$, **c')** $r = 2a_0\cos\theta_0$

En efecto, si se nos presenta el subcaso **a)**, nos encontramos en el supuesto siguiente:

$$-1 < u_1 < u_0 \leq u \leq 1$$

o sea

$$-u_0 < -u_0 + r/a_0 < u \leq u \leq 1$$

que equivale a $0 < r < 2a_0\cos\theta_0$ donde a_0 es mayor que cero y depende sólo de la geometría de masas del cuerpo.

En tal caso, cuando el punto P esté situado a una distancia inferior a $2a_0\cos\theta_0$ de nuestro punto fijo O, el movimiento tiene lugar para valores de u comprendidos entre el valor inicial u_0 y el valor $u = 1$, es decir, para θ variando entre el valor inicial θ_0 y $\theta = 0$. De la primera de las ecuaciones del movimiento se deduce que (ver Byrd & Friedman (1971)):

$$\sqrt{A_2}\,(t - t_0) = \int_{u_0}^{u} \frac{dx}{\sqrt{(1 - x)(x - u_0)(x - u_1)(x + 1)}} =$$

$$= \alpha\, F(\xi, k) \qquad (3.1)$$

siendo $$\alpha = \frac{2}{\sqrt{(1 - u_1)(u_0 + 1)}}.$$

$$\text{sen}^2\xi = \frac{(1 - u_1)(u - u_0)}{(1 - u_0)(u - u_1)} \quad ; \quad k^2 = \frac{(1 - u_0)(u_1 + 1)}{(1 - u_1)(u_0 + 1)} \tag{3.2}$$

y $F(\xi,k)$ la integral elíptica incompleta de primera clase con módulo k. Invirtiendo dicha integral se obtiene

$$\text{sen}^2\xi = \frac{(1 - u_1)(u - u_0)}{(1 - u_0)(u - u_1)} = \text{sn}^2\omega_0(t - t_0) \tag{3.3}$$

donde sn(-) es la función seno amplitud de la teoría de funciones elípticas de Jacobi con el mismo módulo k. Despejando u, obtenemos la nutación en la forma

$$u = \cos\theta = \frac{(1 - u_1)u_0 - (1 - u_0)u_1\text{sn}^2\omega_0(t - t_0)}{(1 - u_1) - (1 - u_0)\text{sn}^2\omega_0(t - t_0)} \tag{3.4}$$

en las que

$$\omega_0 = \frac{\sqrt{A_2}}{\alpha} = \frac{1}{2}\sqrt{m_1\frac{I_1 - I_3}{I_1}(1 + u_0)(1 - u_1)} \tag{3.5}$$

Por las propiedades de periodicidad de las funciones elípticas, se sigue que el ángulo de nutación es una función periodica del tiempo con periodo $T = \frac{2K}{\omega_0}$, que depende de las condiciones iniciales y de la geometría del cuerpo, donde K es la integral elíptica completa de primera clase con módulo k.

En el subcaso **b)** tenemos

$$-1 \le u \le u_0 < u_1 < 1$$

o bien

$$-1 \le u \le u_0 < -u_0 + r/a_0 < 1$$

entonces

$$2a_0\cos\theta_0 < r < (1 + \cos\theta_0)a_0$$

Concluiríamos afirmando que si el punto P está situado a una distancia comprendida entre $2a_0\cos\theta_0$ y $(1 + \cos\theta_0)a_0$ de O, el recorrido de u está comprendido entre u = -1 y el valor inicial u_0, que equivale a un recorrido en θ comprendido entre $\theta = \pi$ y el valor inicial $\theta = \theta_0$.

La primera ecuación del movimiento podemos expresarla en los términos

$$\sqrt{A_2}\,(t - t_0) = -\int_{u_0}^{u} \frac{dx}{\sqrt{(1 - x)(u_1 - x)(u_0 - x)(x + 1)}} = \alpha\, F(\xi, k) \qquad (3.6)$$

donde $$\alpha = \frac{2}{\sqrt{(1 - u_0)(u_1 + 1)}}$$

$$\operatorname{sen}^2\xi = \frac{(u_1 + 1)(u_0 - u)}{(u_0 + 1)(u_1 - u)} \; ; \; k^2 = \frac{(1 - u_1)(u_0 + 1)}{(1 - u_0)(u_1 + 1)} \qquad (3.7)$$

invirtiendo dicha integral tenemos

$$\operatorname{sen}^2\xi = \frac{(u_1 + 1)(u_0 - u)}{(u_0 + 1)(u_1 - u)} = \operatorname{sn}^2\omega_0(t - t_0) \qquad (3.8)$$

entonces

$$u = \frac{(u_1 + 1)u_0 - (u_0 + 1)u_1\,\operatorname{sn}^2\omega_0(t - t_0)}{(u_1 + 1) - (u_0 + 1)\operatorname{sn}^2\omega_0(t - t_0)} \qquad (3.9)$$

donde

$$\omega_0 = \frac{\sqrt{A_2}}{\alpha} = \frac{1}{2}\sqrt{m_1\frac{I_1 - I_3}{I_1}(1 - u_0)(1 + u_1)} \qquad (3.10)$$

En el subcaso **c)** u_0 es una raiz doble de g(u), pudiendo expresar $g(u) = A_2(1 - u^2)(u - u_0)^2$.

el movimiento tendrá lugar para valores de u comprendidos entre $u = -1$ y $u = +1$, es decir, para θ variando entre el valor $\theta = \pi$ y $\theta = 0$. El punto P estará situado a una distancia igual a $2a_0\cos\theta_0$ de nuestro origen O.

De la primera de las ecuaciones del movimiento deducimos que

$$\sqrt{A_2}\,(t - t_0) = \int_{-1}^{u} \frac{dx}{|x - x_0|\sqrt{1 - x^2}} - \int_{-1}^{u_0} \frac{dx}{|x - x_0|\sqrt{1 - x^2}} =$$

$$= \int_{-1}^{u} \frac{dx}{|x - x_0|\sqrt{1 - x^2}} - \text{cte}, \qquad (3.11)$$

de fácil resolución mediante cambios trigonométricos.

3.2 Caso: $u_1 > 1 \Leftrightarrow r > (1 + u_0)a_0$

como

$$\dot{u}^2 = g(u) = A_2(1 - u^2)(u - u_0)(u - u_1)$$

entonces el movimiento sólo puede tener lugar para valores de $u \leq u_0$ con lo cual

$$-1 \leq u \leq u_0 < 1 < u_1$$

o bien

$$-1 \leq u \leq u_0 < 1 < -u_0 + r/a_0$$

lo que equivale a

$$r > (1 + \cos\theta_0)a_0$$

Luego, si el punto P está situado a una distancia de O superior a $(1 + \cos\theta_0)a_0$, los valores de u estan comprendidos entre $u = -1$ y el valor inicial u_0, en términos de θ la variación tiene lugar entre $\theta = \pi$ y el valor inicial θ_0.

De la primera de las ecuaciones del movimiento se deduce que:

$$\sqrt{A_2}\,(t - t_0) = -\int_{u_0}^{u} \frac{dx}{\sqrt{(u_1 - x)(1 - x)(u_0 - x)(x + 1)}} =$$

$$= \alpha\, F(\xi, k) \qquad (3.12)$$

con

$$\operatorname{sen}^2\xi = \frac{2(u_0 - u)}{(1 + u_0)(1 - u)} \quad ; \quad \alpha^2 = \frac{2}{\sqrt{2(u_1 - u_0)}}$$

$$k^2 = \frac{(u_1 - 1)(u_0 + 1)}{2(u_1 - u_0)} \qquad (3.13)$$

E invirtiendo dicha integral, obtenemos:

$$\operatorname{sen}^2\xi = \frac{2(u_0 - u)}{(1 + u_0)(1 - u)} = \operatorname{sn}^2\omega_0(t - t_0) \qquad (3.14)$$

despejando u, obtenemos la nutación en función del tiempo

$$u = \cos\theta = \frac{2u_0 - (1 + u_0)\operatorname{sn}^2\omega_0(t - t_0)}{2 - (1 + u_0)\operatorname{sn}^2\omega_0(t - t_0)} \qquad (3.15)$$

en la que

$$\omega_0 = \frac{\sqrt{A_2}}{\alpha} = \frac{1}{2}\sqrt{m_1 \frac{I_1 - I_3}{I_1}\, 2(u_1 - u_0)} \qquad (3.16)$$

3.3 Caso: $u_1 = 1 \quad \Leftrightarrow \quad r = (1 + \cos\theta_0)a_0$

El valor u = 1 se nos presenta como raiz doble de la ecuación g(u), pudiendo factorizarse como $g(u) = A_2(u - 1)^2(u_0 - u)(1 + u)$. El movimiento tendrá lugar para valores de $u \leq u_0$, que en términos de θ variará entre el $\theta = \pi$ y el valor inicial $\theta = \theta_0$.

El punto P está situado a una distancia igual a $(1 + \cos\theta_0)a_0$ del origen O de nuestro sistema de referencia.

La primera ecuación del movimiento se expresa por

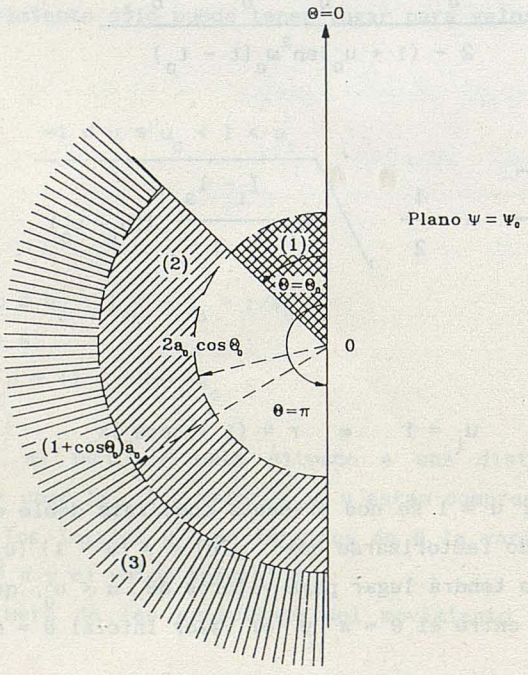
$$\sqrt{A_2}\,(t - t_0) = -\int_{u_0}^{u} \frac{dx}{(1 - x)\sqrt{(1 + x)(u_0 - x)}}\,, \qquad (3.17)$$

de fácil resolución mediante cambios trigonométricos.

4. CONCLUSIONES.

A modo de resumen, las conclusiones obtenidas en nuestro estudio, podemos representarlas gráficamente de la siguiente forma:

Consideremos una esfera de radio suficientemente grande con centro en el punto O, origen de nuestro sistema de referencia, seccionando dicha esfera por el plano $\psi = \psi_0$ tenemos el gráfico siguiente



Con dicha representación gráfica queremos señalar que:

En el sector (1) el ángulo θ puede variar entre su posición inicial $\theta = \theta_0$ y $\theta = 0$, lo cual se verifica cuando el centro de atracción P está a una distancia del punto fijo O inferior a $2a_0\cos\theta_0$; que según vimos se corresponde con $u_1 < u_0$.

En la zona rayada (2) el ángulo θ puede variar entre el valor $\theta = \pi$ y el valor inicial $\theta = \theta_0$, lo que se verifica cuando el centro de atracción P esté situado a una distancia del punto fijo O, comprendida entre $2a_0\cos\theta_0$ y $(1 + \cos\theta_0)a_0$; que según vimos se corresponde con $u_0 < u_1 < 1$.

En la circunferencia de centro el punto O y radio $2a_0\cos\theta_0$ el ángulo θ puede variar entre $\theta = \pi$ y el valor $\theta = 0$, lo cual se verifica cuando el centro de atracción P está a una distancia del punto fijo O exactamente igual a $2a_0\cos\theta_0$; correspondiendose según vimos con el caso $u_1 = u_0$.

En la zona rayada (3) el ángulo θ puede variar entre $\theta = \pi$ y el valor inicial $\theta = \theta_0$, lo cual se verifica cuando el centro de atracción P está a una distancia del punto fijo O superior a $(1 + \cos\theta_0)a_0$; que se corresponde con el caso $u_1 > 1$.

Por ultimo, en la circunferencia de centro O y radio $(1 + \cos\theta_0)a_0$ el ángulo θ puede variar entre $\theta = \pi$ y el valor inicial $\theta = \theta_0$, lo cual se verifica cuando el centro de atracción P está a una distancia del punto fijo O exactamente igual al valor $(1 + \cos\theta_0)a_0$; que según vimos se corresponde con $u_1 = 1$.

En tanto que, las expresiones explícitas de θ en función del tiempo están dadas por las fórmulas (3.4), (3.9), (3.11), (3.15) y (3.17) respectivamente.

BIBLIOGRAFIA.

Arkhangelskii, Iu.A. : 1962, **J. Appl. Math. Mech. 26, 1693.**

Arkhangelskii, Iu.A. : 1963, **J. Appl. Math. Mech. 27, 1059.**

Byrd, F.P. & Friedman, M.D. : 1971, **"Handbook of Elliptic Integrals for Engineers and Scientists"**. Springer-Velag, New York-Heidelberg-Berlín.

Cid, R. & Vigueras, A. :1985, **Celest. Mech. 36, 155.**

Cid, R. & Vigueras, A. :1990, **Rev. Acad. Ciencias. Zaragoza, 45 83, 93.**

Keis, I.A.: 1964, **J. Appl. Math. Mech. 28, 633.**

Leimanis, E. :**"The General Problem of the Motion of Coupled Rigid Bodies about a Fixed Point"**. Springer-Verlag, Berlín.

San Saturio, M.E. & Vigueras, A. :1988, **Celest. Mech. 41, 297.**

Tsopa, M.P.: 1979, **J. Appl. Math. Mech. 43, 189.**

Tsopa, M.P.: 1981, **J. Appl. Math. Mech. 44, 285.**

Vigueras, A. : 1983, **"Movimiento Rotatorio de Giróstatos y Aplicaciones"**. Tesis Doctoral, Universidad de Zaragoza.

Vigueras, A. : 1987, **Actas XII Jorn. Luso - Españolas de Mat., Vol III, 557-563.**

Rev. Academia de Ciencias. Zaragoza 47 (1992)

QUATERNIONS AND NUMERICAL ORBIT COMPUTATION

M. Palacios and C. Calvo

Grupo de Mecánica Espacial. Universidad de Zaragoza, 50009 Zaragoza, Spain.

ABSTRACT

The equations of the motion of a perturbed keplerian orbit are formulated in a non standard way, suggested by Deprit [3], which represents plainly the dynamical behavior of the solution, and avoids singularities in inclination. Quaternions are used for describing rotations between the different frames with a consequent increase of speed and precision. Several test have been performed, and about two significant digits are gained with this formulation.

1. INTRODUCTION

In the last years, requirements in precision in orbit computations have changed drastically. Indeed, at present, centimeter accuracy over one hundred of revolutions of a satellite is required in some missions, mainly in low altitude orbits, where the influence of the atmospheric drag is very strong [2]. Efficiency and speed are requested in numerical methods applied to orbit computation. For obtaining high accuracy in the ephemeris or even for checking analytical theories, we need a good numerical integration.

Recently, from a theoretical point of view, we studied methods for the numerical integration of the equations of the motion in artificial satellite theory. These methods consist of generalizations of the classical Störmer–Cowell methods [4,5]. These generalizations are improvements of the classical ones in several aspects, mainly in the same behavior of the numerical solution that the analytical one, regarding the periodicity properties. Several tests [5] show that with our methods we obtain at least two more digits than with Cowell's method when applied to periodic or quasi–periodic problems. Besides the difficulty in obtaining very high accuracy, there are several constraints that increase the complexity, such as the prohibitive computer time, and the restriction of using arithmetic operations with words of limited length. As it was pointed out in [2], these goals cannot be reached by means of numerical methods as Burlish–Stoer, Adams–Moulton–Cowell or Cowell, whereas the Encke formulation of the equations and a well chosen secularly precessing orbit previous to the numerical integration (e.g. Adams–Moulton–Cowell) bearing into account a reformulation of the gravity vector computation is more efficient.

In this point, the idea of Deprit [3] reveals to be very useful. Indeed, as the inclination of the orbit of an artificial satellite varies little, our purpose is to formulate the equations of the motion in such a way that the inclination remains close to zero in order to avoid the loose of significant digits in the numerical calculations and eliminate the singularities related with small

or near 180 degrees inclinations that can arise when the integration is performed in the classical orbital elements. In order to perform this, we proceed in the following way: first, determining the instantaneous plane of motion, and second, to describe the motion in its plane. Remark that the motion of the satellite in its plane (we will define it as the ideal plane) has only two degrees of freedom. We will define the instantaneous orbital plane in terms of quaternions, since this formulation presents no singularities, even for inclination equal to π, and manipulate quantities very close to the unity.

An important aspect to be taken in to account is the numerical check of this new model for the formulation of the problem. We have done several test in a CONVEX machine that shown an improvement in the efficiency of the numerical integration of this formulation versus the classical one in rectangular coordinates. The efficiency obtained, about two more digits, has been mesure in terms of number of digits gained in the determination of radial distance, radial velocity and angular momentum. The cassical along-track, across-track and across-plane errors have been obtained and they show that the more important, along-track error, is of the order of the error in the determination of radial velocity, the across-track error is of lest importance and the across-plane error insignificant.

In section 2., we define the different reference frames to be employed, together with a description of the properties of quaternions which will be very helpful in the integration of the equations of the motion obtained in the ideal frame (subsection 3..1). Several test of the distinct formulations for non-perturbed and perturbed problem are performed in the section 4., which show the efficiency of the integration when formulated in the departure and ideal frames.

2. REFERENCE FRAMES AND QUATERNIONS

Firstable, let us consider the *space frame* $\mathcal{S}$, defined by three orthonormal vectors $(\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3)$. The position $\boldsymbol{x}$ and velocity $\boldsymbol{X}$ (rectangular coordinates) of a particle in a Newtonian potential of constant μ to which is superimposed a perturbing force $\boldsymbol{F}$ is given by the differential equations

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{X}, \\ \dot{\boldsymbol{X}} &= -\left(\frac{\mu}{r^3}\right)\boldsymbol{x} + \boldsymbol{F}. \end{aligned} \tag{1}$$

Following Deprit [3], we shall choose another reference frame and a set of variables (not necessarily minimum) in such a way that the equations show as clear as possible the dynamical properties of the solutions and give us the best efficiency when integrated numerically.

Let us consider now the *orbital frame* $\mathcal{O}$ $(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{n})$, where the orthonormal basis is given by

$$\begin{aligned} \boldsymbol{u} &= \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} = \frac{\boldsymbol{x}}{r} \\ \boldsymbol{n} &= \frac{\boldsymbol{x} \times \boldsymbol{X}}{\|\boldsymbol{x} \times \boldsymbol{X}\|} = \frac{\boldsymbol{G}}{G} \\ \boldsymbol{v} &= \boldsymbol{n} \times \boldsymbol{u} \end{aligned} \tag{2}$$

In particular, in the instant t_0, the orbital frame defines a fixed orthonormal reference system named *departure frame* $\mathcal{D}$ $(\boldsymbol{u}_0, \boldsymbol{v}_0, \boldsymbol{n}_0)$.

Since the departure frame $\mathcal{D}$ may be regarded as the image of the space frame $\mathcal{S}$ by a rotation, the relation between both systems may be given in matricial form orusing the associated quaternion; we shall express this relation by means of quaternions. Quaternions present the advantage of no having inherent geometrical singularities, and from a numerical point of view, we have to make only arithmetic operations among them.

First at all, a quaternion $\bar{q} = (q_0, q_1, q_2, q_3)$ may be considered as a pair $(q_0; \boldsymbol{q})$, composed by an scalar q_0 and a vector $\boldsymbol{q} = (q_1, q_2, q_3)$. The set of all quaternions of norm unity and non negative real part,

$$S^4 = \Big\{ \bar{q} = (q_0, q_1, q_2, q_3) \quad \Big| \quad \|\bar{q}\|^2 = q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1, \quad q_0 \geq 0 \Big\}.$$

may be considered as a group if we introduce the product of two quaternions be defined as

$$\bar{q}\,\bar{q}' = (q_0 q_0' - \boldsymbol{q} \cdot \boldsymbol{q}';\ q_0 \boldsymbol{q}' + q_0' \boldsymbol{q} + \boldsymbol{q} \times \boldsymbol{q}').$$

Let us mention that there exists an isomorphism between the group $O_3^+(\mathbb{R})$ of rotations of order 3 (positive orthogonal matrices) and the group S^4 of quaternions

$$O_3^+(\mathbb{R}) \approx S^4$$

The isomorphism is completely defined when the rotation matrix Q is identified with the following matrix

$$\begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1 q_2 - q_0 q_3) & 2(q_1 q_3 + q_0 q_2) \\ 2(q_1 q_2 + q_0 q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 - q_0 q_2) & 2(q_2 q_3 + q_0 q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix}$$

corresponding to the quaternion $\bar{q}$.

A quaternion given, the obtaining of its associate matrix does not offer any formal difficulty. However, in the inverse process we may have a lost of significant digits, because some small divisisor may appear. In order to avoid this obstacle in our subroutines, we take among the four possible options the greater component q_i of the quaternion as denominator.

Besides, taking in to account that we can define a one-to-one mapping from $\mathbb{R}^3$ to S^4 given by

$$\begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & S^4 \\ \boldsymbol{v} & \longmapsto & \bar{v} = (0; \boldsymbol{v}), \end{array}$$

it is easy to check that if $\boldsymbol{y}$ is the image of the vector $\boldsymbol{x}$ by a rotation, the well known relation $\boldsymbol{y} = Q\boldsymbol{x}$ is equivalent to

$$\bar{y} = \bar{q}\,\bar{x}\,\bar{q}^*, \tag{3}$$

where $\bar{x}$, $\bar{y}$ are the quaternion images of the vectors $\boldsymbol{x}$, $\boldsymbol{y}$ by the previous mapping, $\bar{q}$ is the quaternion corresponding to the rotation, and $\bar{q}^*$ is the conjugate quaternion of the previous one. (For more details on quaternions cfr. [6,1]).

With the aforementioned properties, a vector $\boldsymbol{w}_d$ in the departure frame may be expressed in the space frame $(\boldsymbol{w}_s)$ by means of

$$\boldsymbol{w}_s = \bar{q}\,\boldsymbol{w}_d\,\bar{q}^* = (2q_0^2 - 1)\boldsymbol{w}_d + 2(\boldsymbol{q} \cdot \boldsymbol{w}_d)\boldsymbol{q} + 2q_0(\boldsymbol{q} \times \boldsymbol{w}_d).$$

In analogous way, the inverse transformation is:

$$\boldsymbol{w}_d = \bar{q}^* \, \boldsymbol{w}_s \, \bar{q} = (2q_0^2 - 1)\boldsymbol{w}_s + 2(\boldsymbol{q} \cdot \boldsymbol{w}_s)\boldsymbol{q} - 2q_0(\boldsymbol{q} \times \boldsymbol{w}_s).$$

Other property used below is related with derivative of quaternions. Differentiating equation (3), one obtains

$$d\bar{y} = d\bar{q}\,\bar{x}\,\bar{q}^* + \bar{q}\,\bar{x}\,d\bar{q}^* = d\bar{q}\,\bar{q}^*\,\bar{y} + \bar{y}\,\bar{q}\,d\bar{q}^*$$

and after some arithmetics one obtains

$$\frac{d\bar{y}}{dt} = 2\dot{\bar{q}}\,\bar{q} \times \bar{y}$$

If one name $\bar{\omega}$ the instantaneous quaternion angular velocity, identifying the previous result with $\bar{\omega} \times \bar{y}$, one may obtains $\bar{\omega} = 2\dot{\bar{q}}\,\bar{q}$, or equivalently

$$2\dot{\bar{q}} = \bar{\omega}\,\bar{q}. \tag{4}$$

3. EQUATIONS OF THE MOTION

By definition, the orbital frame $\mathcal{O}\ (\boldsymbol{u}(t), \boldsymbol{v}(t), \boldsymbol{n}(t))$ coincides with the departure frame D in the initial instant t_0.

Having into account the equations (1), and the definition of the angular momentum $\boldsymbol{G} = \boldsymbol{x} \times \boldsymbol{X}$, the derivatives with respect the independent variable t, are:

$$\dot{\boldsymbol{G}} = \boldsymbol{x} \times \boldsymbol{F}, \qquad \dot{G} = r\,\boldsymbol{F} \cdot \boldsymbol{v}. \tag{5}$$

After differenciating the relation $\boldsymbol{G} = G\boldsymbol{n}$, one obtains that

$$\dot{\boldsymbol{n}} = \frac{r}{G}(\boldsymbol{F} \cdot \boldsymbol{n})\boldsymbol{u} \times \boldsymbol{n}.$$

Likewise, by differenciating $\boldsymbol{x} = r\boldsymbol{u}$, we have

$$\dot{\boldsymbol{u}} = \frac{G}{r^2}\boldsymbol{n} \times \boldsymbol{u}.$$

Finally, from $\boldsymbol{v} = \boldsymbol{n} \times \boldsymbol{u}$, and substituting the previous derivatives, we obtain

$$\dot{\boldsymbol{v}} = \left(\frac{G}{r^2}\boldsymbol{n} + \frac{r}{G}(\boldsymbol{F} \cdot \boldsymbol{n})\boldsymbol{u}\right) \times \boldsymbol{v}.$$

Accordingly, the motion of the orbital frame with respect to the fixed frame may be viewed as the rotation

$$\dot{\boldsymbol{u}} = \boldsymbol{\omega} \times \boldsymbol{u}, \qquad \dot{\boldsymbol{v}} = \boldsymbol{\omega} \times \boldsymbol{v}, \qquad \dot{\boldsymbol{n}} = \boldsymbol{\omega} \times \boldsymbol{n}, \tag{6}$$

where the angular velocity is

$$\boldsymbol{\omega} = \frac{G}{r^2}\boldsymbol{n} + \frac{r}{G}(\boldsymbol{F} \cdot \boldsymbol{n})\boldsymbol{u}. \tag{7}$$

3..1 IDEAL FRAME

In order to define the ideal frame, we consider the second term of the angular velocity (7),

$$\omega_I = \frac{r}{G}(\boldsymbol{F}\cdot\boldsymbol{n})\boldsymbol{u}$$

The ideal frame $(\boldsymbol{u}_I, \boldsymbol{v}_I, \boldsymbol{n}_I)$ was defined by Hansen as the departure frame rotated by ω_I. Remark that when the perturbation is small, the vector ω_I is small, that means slow rotation, and, besides, the velocity is the same in both frames (departure and ideal). In this way the motion may be decomposed in a slow rotation of the orbital (or ideal) plane and the motion in this plane.

The ideal frame along the time referred to the departure frame may be viewed as the rotation

$$\dot{\boldsymbol{u}}_I = \omega_I \times \boldsymbol{u}_I,\ \dot{\boldsymbol{v}}_I = \omega_I \times \boldsymbol{v}_I,\ \dot{\boldsymbol{n}}_I = \omega_I \times \boldsymbol{n}_I, \tag{8}$$

with the following initial conditions

$$\begin{aligned}\boldsymbol{u}_I(t_0) &= (1,0,0),\\ \boldsymbol{v}_I(t_0) &= (0,1,0),\\ \boldsymbol{n}_I(t_0) &= (0,0,1).\end{aligned}$$

Considering the aceleration relative to the ideal frame and taking into account the equality of the velocities in both frames, we obtain the following equation

$$\frac{\partial}{\partial t}\dot{\boldsymbol{x}} = -\frac{\mu}{r^3}\boldsymbol{x} + (\boldsymbol{F}\cdot\boldsymbol{u})\boldsymbol{u} + (\boldsymbol{F}\cdot\boldsymbol{v})\boldsymbol{v} \tag{9}$$

where vectors are referred to the ideal frame.

Equations (8) and (9) define the ideal frame and the motion in the ideal plane. These equations (8) may be substituted by the equations for the corresponding quaternion defining the rotation from the departure frame to the ideal frame. That is done by taking into account equation (4) for ω_I and expanding the product, one obtains the following equation for the quaternion:

$$\dot{\bar{q}} = \frac{1}{2}\omega_I\,\bar{q}$$

or equivalently in components, once the value of ω_I has been substituted:

$$\begin{aligned}\dot{q}_0 &= -\frac{1}{2}\frac{r}{G}(\boldsymbol{F}\cdot\boldsymbol{n})(q_1 cos\theta + q_2 sin\theta)\\ \dot{q}_1 &= -\frac{1}{2}\frac{r}{G}(\boldsymbol{F}\cdot\boldsymbol{n})(q_0 cos\theta - q_3 sin\theta)\\ \dot{q}_2 &= -\frac{1}{2}\frac{r}{G}(\boldsymbol{F}\cdot\boldsymbol{n})(q_0 sin\theta + q_3 cos\theta)\\ \dot{q}_3 &= -\frac{1}{2}\frac{r}{G}(\boldsymbol{F}\cdot\boldsymbol{n})(-q_2 sin\theta + q_1 sin\theta)\end{aligned} \tag{10}$$

where θ is the angle between the $\boldsymbol{u}_I$ and the radial direction defined by $\boldsymbol{u}$

The vectorial equation (9) may be substituted by four equivalent scalar equations: the equation for the angular momentum is the second of equations (5); the equation for the radial

distance and radial velocity is obtained by differentiating the expression $r\,\dot{r} = \boldsymbol{x}\cdot\boldsymbol{X}$; the equation for the angle θ is obtained by taking into account that vector $\boldsymbol{u} = (cos\theta, sin\theta, 0)$ with respect to the ideal frame and its instantaneous variation is

$$\frac{\partial \boldsymbol{u}_I}{\partial t} = \boldsymbol{\omega}_I \times \boldsymbol{u}_I = \frac{G}{r^2}\boldsymbol{v}_I = \frac{\partial \boldsymbol{u}_I}{\partial \theta}\frac{\partial \theta}{\partial t}$$

So, one may write these equations in the following form:

$$\begin{aligned} \dot{G} &= r\,\boldsymbol{F}\cdot\boldsymbol{v} \\ \ddot{r} &= \frac{G^2}{r^3} - \frac{\mu}{r^2} + \boldsymbol{F}\cdot\boldsymbol{u} \\ \frac{\partial\theta}{\partial t} &= \frac{G}{r^2} \end{aligned} \qquad (11)$$

Consequently, we have an initial value problem composed by a system of 8 equations ((10) and (11)) with 8 unknowns and the initial conditions

$$q_0(t_0) = 1, \quad q_1(t_0) = 0, \quad q_2(t_0) = 0, \quad q_3(t_0) = 0,$$
$$r(t_0) = r_0, \quad \dot{r}(t_0) = \dot{r}_0, \quad G(t_0) = G_0, \quad \theta(t_0) = \theta_0.$$

However, the system is redundant since the unknowns must satisfy the constraint of normalization for the quaternion, $||\bar{q}|| = 1$, which will be employed to monitor the propagation of round off and truncation errors all over the numerical integration.

4. NUMERICAL TESTS

In order to check these new formulation we have developed several FORTRAN programs for the numerical integration of the initial value problems above mentioned, taking initial conditions and perturbation force corresponding to an artificial satellite with low altitude.

Three different cases of orbits have been tested:

- keplerian circular orbit,
- keplerian elliptical orbit,
- zonal and tesseral harmonics.

We have run our programs in the vectorial computer CONVEX 220 of the University of Zaragoza, but in the begining we have considered the Standard Apple Numeric Environment (SANE) for binary floating point arithmetic. SANE uses the extended data type to perform most computation, which provides the maximum degree of accuracy, which is 19 to 20 decimal digits (with non elemental arithmetic operations the least significant bit of extended may be in error). For this reason all subroutines are prepared in order to vectorize as much as possible and so save time computing

Since we are interested only in checking the different formulations above described and not in the numerical method themselves, we choose a Runge-Kutta of order eight for all cases, with

stepsize equal 60, 90, 180, 270 and 360 seconds (they are in the range of stepsize recommended in other programs). Furthermore, a good reference ephemeris is needed to serve as an absolute mean of comparison. In the first case (circular orbit) we used the exact values, whereas in the other two, we used as orbit of reference one computed in rectangular coordinates in the space frame with the same conditions and stepsize equal to 10 seconds and quadruple precision. We have check that the order of approximation of the orbit of reference is about 10^{-20}.

In a more practical sense and in order to avoid loose in significant digits, we have done a change in the scale of the unities in order to have all quantities of similar magnitude and close to one and, therefore, the significance of the calculations has increased. In particular, we took the following unities: equatorial radius of the Earth (R_0) for distance, time spent in a circular revolution of radius R_0 for time.

To illustrate our numerical test, we present here only some of the graphics corresponding to the test performed; more precisely, the related to a orbit with initial orbital elements $a = 7000$ Kms., $e = 0.1$, $i = 135°$ perturbed with zonal and tesseral harmonics until order two; so each revolution of the satellite spend about 1.3 unities of time. All the figures have been calculated for a final time about twenty revolutions or about 26 unities of time.

In figures 1 to 4 we represent for the mentioned problem the deviation from the reference orbit of the radial distance, radial velocity and angular momentum calculated in rectangular coordinates in the space frame (cf. equations (1)). In the vertical axis we represent minus the decimal logarithm of the differences between the corresponding quantities of the orbit of reference and the calculated ones; this represent approximately the number of exact digits obtained in the numerical integration. We observe that the three quantities are in same range, nevertheless the radial velocity has been calculates with less approximation. This approximation reach from five digits if stepsize equal 360 seconds to 13 digits if stepsize equal 60 seconds (this is the recommended stepsize in most of the usual programmes).

In figures 6 to 9 we represent for the mentioned problem the deviation from the reference orbit of the radial distance, radial velocity and angular momentum calculated in the ideal frame by means of quaternions and making use of the equations (8) and (9). Actually, the angular momentum is better approximated than the radial distance and radial velocity. This is due indoubtly to fact that equations (8) and (9) represent more acdequately the dynamics of the motion, because the angular momentum is near an integral of the motion. One more time the radial velocity has been calculated with less appoximation than radial distance. This time the approximation reach from 9 digits after 20 revolutions if stepsize equals 360 seconds to 15 digits if stepsize equals 60 seconds.

A comparison of both sets of figures mentioned shows that our model gives two more digits of approximation in the worth case, i.e. in the small stepsize. If the stepsize is greater than 60 seconds, more than two digits have been gained. As a consequence, the same approximation as in the rectangular coordinates model may reached using a greater stepsize and so we may saved time computing.

In figures 10 to 15 we represent the along-track, across-track and across-plane errors after

about 20 revolutions of the satellite. We may observe that the maximum of the across- track and across-plane errors grows very slowly and linearly, just as the truncation error of the numerical method used, although the across-plane error is much smaller than the across-track error that is no greater than 10^{-9} if stepsize equal 360 seconds or 10^{-16} if stepsize 60 seconds. The along-track error, much more significant in others model of integration than the other two, is in the same range than the across-track error in our model of integration; its growing is not yet linear but slightly non linear if stepsize 60 seconds is taken. When compared with figures 5. and 5. we may deduce that along-track, across-track and across-plane errors do not reach the radial distance error that always is the greatest of them.

5. CONCLUSIONS

A new formulation for the equations of the motion of an artificial satellite using an ideal frame and quaternions to achieve rotations has been developed. The use of ideal frame enable us the elimination of all type of singularities in inclination as well as the consideration of every orbit as have null inclination. This fact together with the use of quaternions and the mentioned system of unities, that eliminates innecessary loose of significant digits, has enabled us to obtain an improvement of no less than two more digits with respect to the integration in rectangular coordinates with the same numerical method. We must mention that no more time of computations has been required and that the use of quaternions require less and simpler arithmetic operations and consequently an improvement in time computing. All calculations have been performed in the array processor Convex 220 of the University of Zaragoza.

Support came in part from the C.I.C.Y.T. (Proyects # PB90–0921 and # ESP91-0919).

Referencias

[1] Abad, A., Arribas, M. and Elipe, A.:1989, On the attitude of a spacecraft near a Lagrangian point. *Bull. Astron. Inst. Czechosl.* **40**, 302-307

[2] Balmino, G. and Barriot, J.P.:1988, *Revision of numerical techniques*.Work package No. 221 ESA Contract No. 7521/87/F/FL, Final report.

[3] Deprit, A.: 1975, Ideal Elements for Perturbed Keplerian Motions, *Journal of Research of the NBS* **79B**, 1–15

[4] Palacios, M. and Franco, J.M.:1990, Adaptive Cowell methods for orbital motions. *Advances in the Astronautical Sciences.* **69**, 3–22

[5] Franco, J.M. and Palacios, M.:1989, High order P-stable multistep methods. *J. Com. Appl. Math.* **34**, 1–12

[6] Whittaker, E.T.:1905, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Rep. Dover Pub. Inc., New York.

[7] Abad, A., Palacios, M. and Sein Echaluce, M.L.:1991, An Efficient Formulation for Numerical Orbit Computation. *Proc. ESA Symposium Spacecraft Flight Dynamics* , 65–70

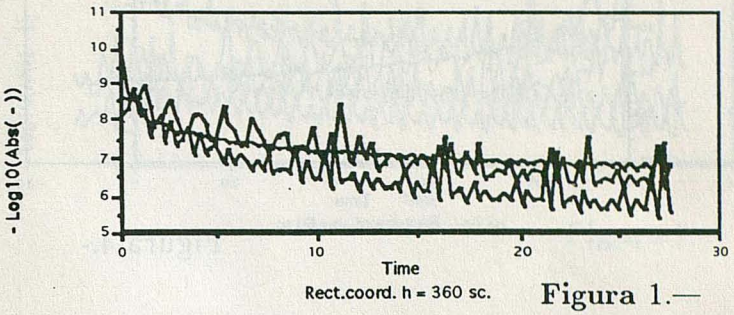


Figura 1.—

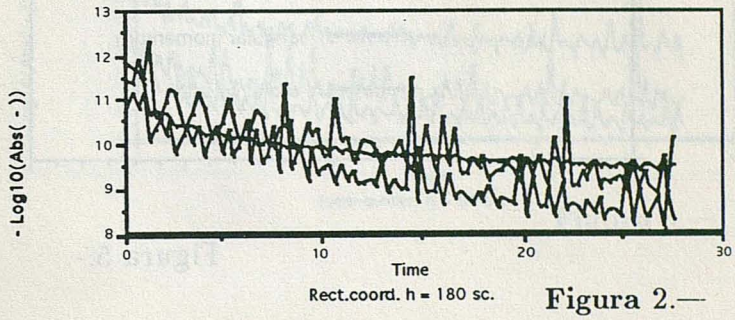


Figura 2.—

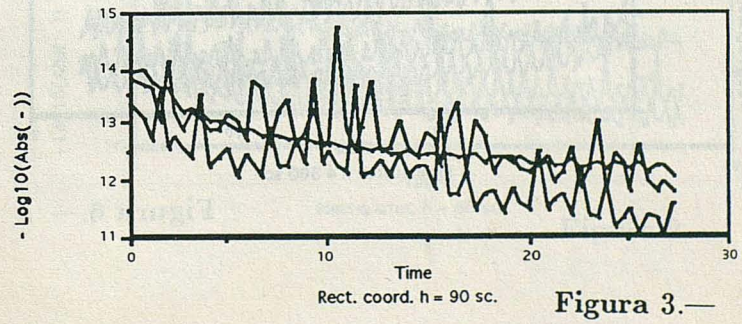


Figura 3.—

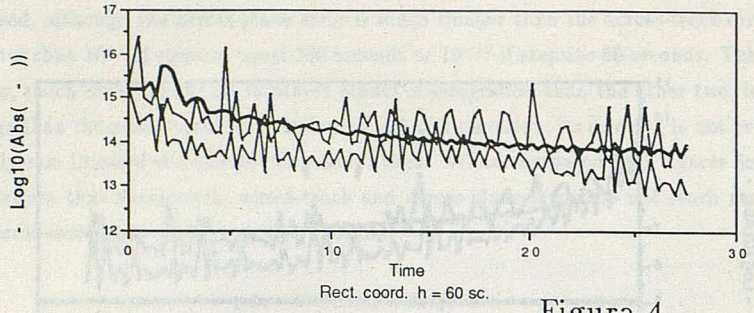


Figura 4.-

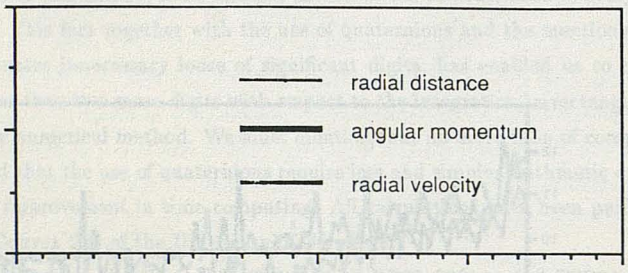


Figura 5.-

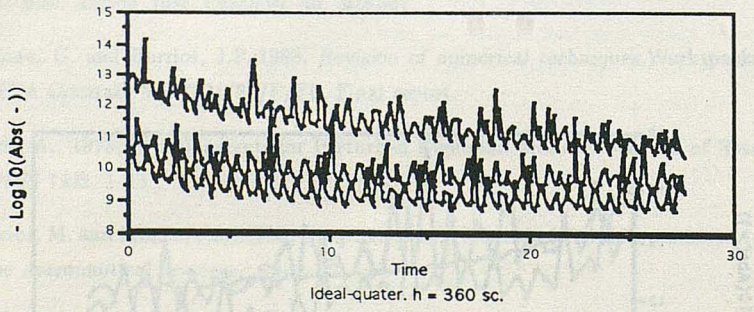


Figura 6.—

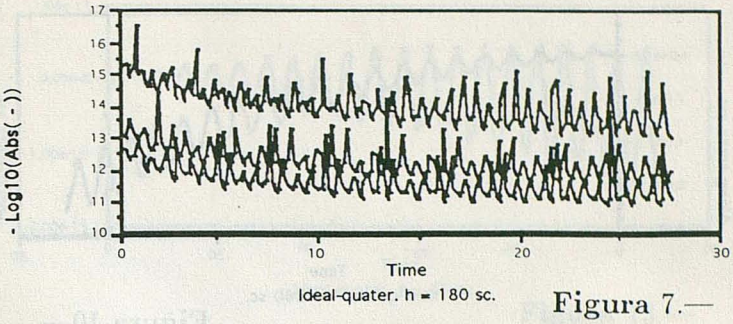


Figura 7.—

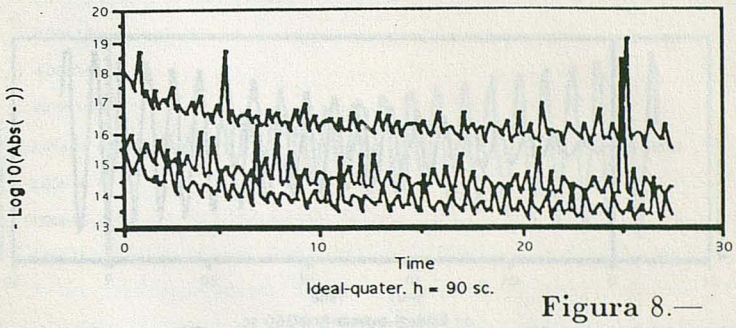


Figura 8.—

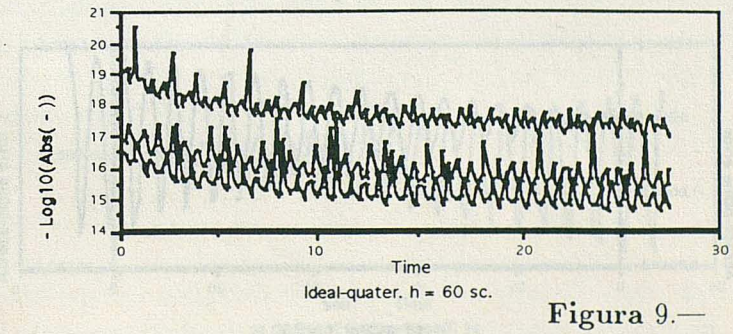


Figura 9.—

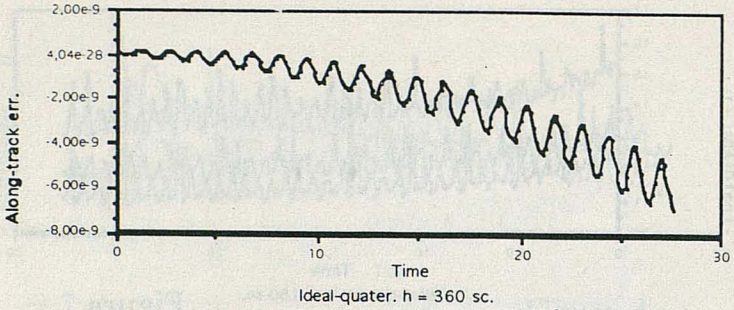


Figura 10.—

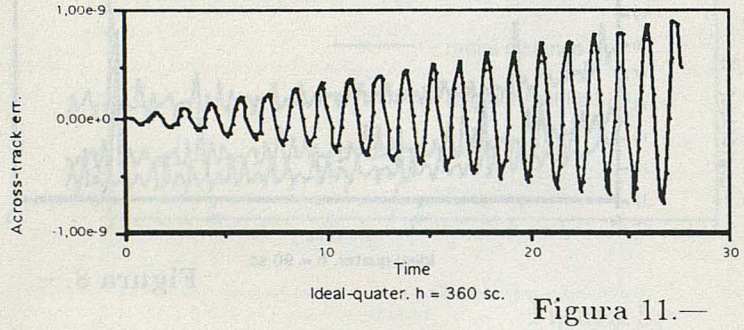


Figura 11.—

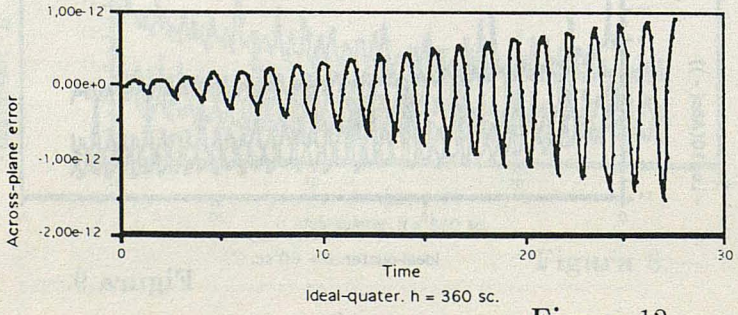


Figura 12.—

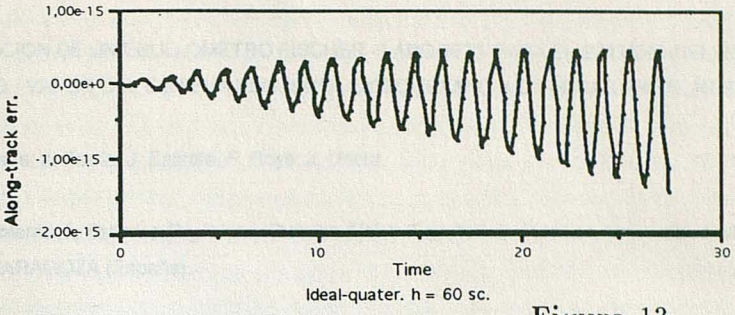


Figura 13.—

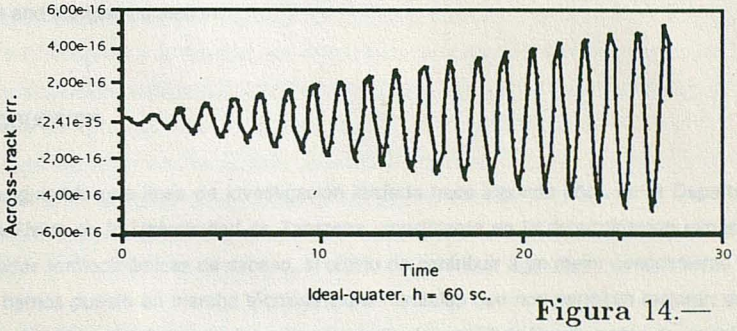


Figura 14.—

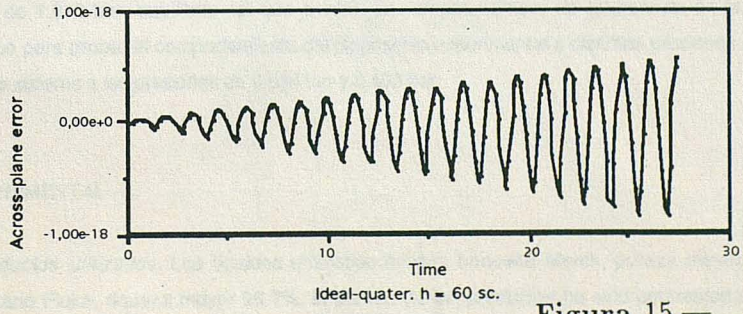


Figura 15.—

Rev. Academia de Ciencias. Zaragoza 47 (1992)

UTILIZACION DE UN EBULLOMETRO FISCHER - LABODEST PARA EL ESTUDIO DEL EQUILIBRIO LIQUIDO - VAPOR DEL SISTEMA BENCENO - CICLOHEXANO A DIVERSAS PRESIONES

C. Lafuente, J. Pardo, J. Santafé, F. Royo, J. Urieta.

Departamento de Química Orgánica y Química-Física. Facultad de Ciencias, Universidad de Zaragoza. 50009 ZARAGOZA (España).

ABSTRACT

The isobaric vapor-liquid equilibrium data of benzene-cyclohexane were determinated at 1.013 bar, 0.666 bar and 0.400 bar. Activity coefficients were evaluated and correlated with Redlich-Kister equation and Margules equation.

1. INTRODUCCION

Prosiguiendo una linea de investigación iniciada hace algunos años en el Departamento de Química-Física de la Universidad de Zaragoza, consistente en la determinación experimental de propiedades termodinámicas de exceso, al objeto de contribuir a un mejor conocimiento del estado líquido, hemos puesto en marcha técnicas experimentales que nos permitan mejorar, en tiempo y precisión, la obtención de resultados para el estudio del equilibrio líquido-vapor de mezclas líquidas binarias no electrolíticas.

Con objeto de comprobar el correcto funcionamiento y rendimiento del ebullómetro Fischer-Labodest se estudió con el mismo el equilibrio líquido-vapor para el sistema benceno-ciclohexano a la presión de 1.013 bar, equilibrio del que existen un amplio número de referencias en la literatura. Asímismo para probar el comportamiento del dispositivo experimental a distintas presiones se estudió el mismo sistema a las presiones de 0.666 bar y 0.400 bar.

2. EXPERIMENTAL

2.1 *Productos utilizados.* Los líquidos utilizados fueron: benceno Merck, pureza mayor 99.7%, y ciclohexano Fluka, riqueza mayor 99.7%; la pureza de los productos ha sido contrastada mediante medidas de densidad y GC. Los valores obtenidos para la densidad a 298.15 K fueron: 0.87327 $g \cdot cm^{-3}$ para el benceno y 0.77344 $g \cdot cm^{-3}$ para el ciclohexano, siendo los valores de la literatura

(Timmermans,1950): 0.87366 $g \cdot cm^{-3}$ y 0.773832 $g \cdot cm^{-3}$ respectivamente.

2.2 *Aparato y procedimiento*. El aparato utilizado para el estudio del equilibrio líquido-vapor fue un ebullómetro de recirculación Fischer-Labodest, en el que introdujimos algunas mejoras para conseguir un funcionamiento óptimo. En la figura1 se muestra un esquema del aparato.

El calentamiento hasta ebullición de la mezcla se consigue mediante un baño de aceite de silicona (1) que permite una calefacción suave y eficaz al mismo tiempo que permite alcanzar, cuando es necesario, elevadas temperaturas. La vasija de equilibrio tiene una capacidad de 200 cm^3. Por medio de un agitador electromagnético (1) y un par de nucleos magnéticos situados uno en el baño y otro en la vasija, se garantiza por un lado una calefacción uniforme y por otro una buena mezcla de las corrientes de recirculación de vapor condensado y de líquido con el contenido de la vasija. El aparato dispone de la correspondiente bomba de Cottrel (2) para una medición correcta de la temperatura de ebullición. La parte superior del ebullómetro, en la que se encuentra confinada la bomba, está rodeada de una camisa de vidrio plateada interiormente que limita un espacio evacuado que proporciona aislamiento y lleva además exteriormente una manta con calefacción eléctrica (3) que previene posibles condensaciones de la fase vapor. Para la medición de la temperatura de ebullición se dispone de una sonda de platino PT100 (5) calibrada previamente con otra patrón y un multímetro Hewlett-Packard 34401A de 61/2 cifras (9). La precisión en la medida de las temperaturas asi determinadas es mejor de $\pm 0.1^{o}$ C. El condensador (4) proporciona la corriente de vapor condensado. El uso de las válvulas electromagnéticas (6,7) facilita la extracción de muestras al mismo tiempo que permiten mantener aisladas las muestras de las corrientes de recirculación. Para el ajuste de la presión se utilizó un manostato (8) altamente sensible, del tipo regulador eléctrico de presión, acoplado a una bomba de vacio y un trasductor de presiones Druck PDCR 110/W con indicador DPI 220 (10) que permite medidas de presión de hasta 1.500 bar con una precisión de ± 0.1 mbar.

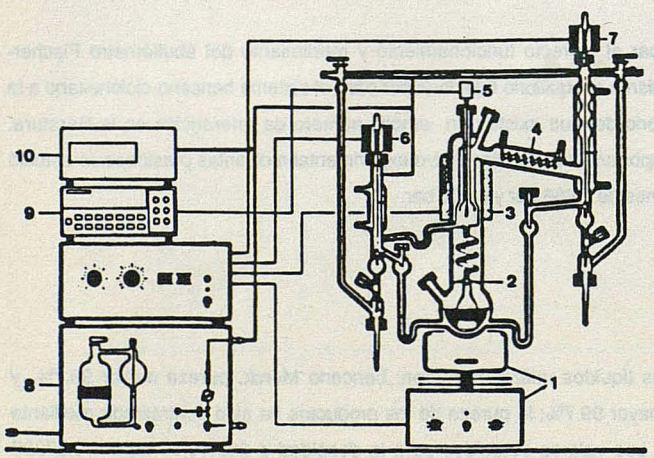


Figura 1. Ebullómetro Fischer-Labodest

La determinación de la composición de las fases líquida y vapor se realizó mediante medidas de densidad a la temperatura de 298.15 K, usando a tal efecto un densitímetro de tubo vibrante Anton Paar DMA-58. La composición de las muestras obtenidas en el ebullómetro se determina por comparación de sus densidades con las densidades de mezclas de composición conocida, medidas y ajustadas previamente.

3. RESULTADOS Y DISCUSION

El estudio del equilibrio líquido-vapor para el sistema benceno-ciclohexano se llevó a cabo a las presiones de 1.013, 0.666 y 0.400 bar. Los resultados experimentales se recogen en las tablas 2-4 y en las figuras 2-7.

El sistema presenta un azeótropo de temperatura de ebullición mínima para cada una de las presiones consideradas. Las temperaturas de ebullición y las correspondientes fracciones molares de los azeótropos son: a la presión de 1.013 bar: $x_1=y_1=0.540$ y t=77.5º C; a la presión de 0.666 bar: $x_1=y_1=0.532$ y t=64.0º C; a la presión de 0.400 bar: $x_1=y_1=0.516$ y t=49.2º C:

Los coeficientes de actividad se calcularon a partir de la ecuación (Van Ness,1964):

$$\ln \gamma_i = \ln (p\, y_i / p_i^o\, x_i) + (B_{ii} - V_i^o) (p - p_i^o) / R\,T + p_i^o (1 - y_i)^2 \delta / R\,T \qquad (1)$$

donde

$$\delta = 2B_{ij} - B_{ii} - B_{jj} \qquad (2)$$

Las presiones de vapor de los componentes puros fueron calculadas usando la ecuación de Antoine

$$\log p_i^o = A - B/(t+C) \qquad (3)$$

Las constantes utilizadas en estos cálculos se muestran en la tabla 1.

TABLA 1. Constantes de la ecuación de Antoine.

Compuesto	A	B	C
benceno	6.90565	1211.03	220.790
ciclohexano	6.83917	1200.31	222.504

Tanto las constantes de la ecuación de Antoine como los volumenes molares de los líquidos y los

segundos coeficientes del virial se obtuvieron o estimaron a partir de datos tomados de tablas de propiedades físico-químicas de compuestos orgánicos (TRC Tables,1974).

Los datos experimentales (temperatura y composición de las fases líquida y vapor) son termodinámicamente consistentes de acuerdo con el criterio de Herignton (Herignton,1951).

La correlación de los datos experimentales se llevó a cabo utilizando las ecuaciones de Redlich-Kister y Margules, los coeficientes de las cuales se obtuvieron por ajuste de la función $Q = G^E_m/RT$. En la tabla 5 aparecen los coeficientes de las ecuaciones para el sistema estudiado a las distintas presiones de trabajo.

Los coeficientes de actividad a dilución infinita se calcularon a partir de los coeficientes de la ecuación de Margules y los resultados obtenidos que se consignan junto con algunos de la literatura aparecen en la tabla 6, observandose para las medidas correspondientes a la serie de 1.013 bar, única de las tres para la que se disponen de suficientes referencias en este sistema, que existe una concordancia satisfactoria entre nuestros resultados y los encontrados en la literatura.

Tabla 2. Datos del equilibrio líquido-vapor para el sistema benceno(1)-ciclohexano(2) a 1.013 bar.

Temp.(º C)	x_1	y_1	γ_1	γ_2
79.7	0.0781	0.1035	1.341	1.002
79.2	0.1258	0.1590	1.298	1.006
78.7	0.1905	0.2295	1.256	1.010
78.4	0.2310	0.2718	1.238	1.013
78.0	0.2888	0.3253	1.199	1.027
77.8	0.3662	0.3949	1.155	1.040
77.6	0.4639	0.4766	1.107	1.070
77.4	0.5045	0.5108	1.098	1.088
77.4	0.5505	0.5486	1.081	1.107
77.5	0.6109	0.5986	1.058	1.132
77.6	0.6553	0.6358	1.046	1.158
77.7	0.7099	0.6834	1.033	1.190
78.0	0.7782	0.7450	1.018	1.243
78.2	0.8239	0.7903	1.014	1.280
78.6	0.8664	0.8347	1.008	1.316
79.2	0.9301	0.9077	1.002	1.380

Tabla 3. Datos del equilibrio líquido vapor para el sistema benceno(1)-ciclohexano(2) a 0.666 bar.

Temp.(° C)	x_1	y_1	γ_1	γ_2
66.0	0.1033	0.1355	1.360	1.009
65.6	0.1516	0.1905	1.323	1.011
65.3	0.2048	0.2460	1.274	1.015
64.9	0.2598	0.2994	1.239	1.026
64.5	0.3210	0.3527	1.195	1.045
64.3	0.4142	0.4351	1.152	1.066
64.1	0.4722	0.4828	1.129	1.090
64.0	0.5164	0.5192	1.114	1.110
64.0	0.6002	0.5881	1.086	1.150
64.1	0.6442	0.6231	1.068	1.179
64.2	0.6992	0.6741	1.059	1.200
64.5	0.7695	0.7348	1.040	1.264
64.9	0.8342	0.7979	1.028	1.322
65.4	0.8899	0.8581	1.020	1.375
65.7	0.9214	0.8930	1.015	1.438

Tabla 4. Datos del equilibrio líquido-vapor para el sistema benceno(1)-ciclohexano(2) a 0.400 bar.

Temp.(° C)	x_1	y_1	γ_1	γ_2
51.2	0.0979	0.1300	1.402	1.017
50.8	0.1487	0.1920	1.383	1.015
50.4	0.2014	0.2464	1.331	1.024
50.1	0.2578	0.3039	1.296	1.029
49.8	0.3032	0.3426	1.257	1.046
49.7	0.3382	0.3713	1.226	1.057
49.6	0.3614	0.3919	1.215	1.064
49.5	0.3999	0.4235	1.191	1.077
49.3	0.4380	0.4542	1.175	1.097
49.2	0.4909	0.4957	1.149	1.123
49.2	0.5337	0.5292	1.128	1.145
49.3	0.5777	0.5668	1.112	1.159
49.4	0.6410	0.6169	1.086	1.201

Tabla 4. Continuación

Temp.(° C)	x_1	y_1	γ_1	γ_2
49.6	0.6953	0.6618	1.066	1.240
49.9	0.7675	0.7244	1.046	1.310
50.3	0.8219	0.7797	1.035	1.347
50.7	0.8802	0.8410	1.028	1.425

Tabla 5. Coeficientes y desviaciones standard para las ecuaciones de Redlich-Kister y Margules.

	Redlich-Kister				Margules		
Presión (bar)	A_0	A_1	A_2	δ	A_{12}	A_{21}	δ
1.013	0.3457	0.0237	0.0187	0.0014	0.3719	0.3250	0.0015
0.666	0.4138	0.0824	0.0834	0.0018	0.5087	0.3437	0.0034
0.400	0.4948	0.0697	0.0723	0.0027	0.5722	0.4353	0.0035

Tabla 6. Coeficientes de actividad a dilución infinita.

Presión (bar)	γ_1^∞	γ_2^∞	Autores
1.013	1.38	1.45	Este trabajo
	1.37	1.45	Chao,1956
	1.29	1.43	Darmois, Darmois,1964
	1.41	1.50	Donald, Ridgway,1958
	1.42	1.43	Nagata,1962
	1.36	1.45	Nataraj, Rao,1967
	1.40	1.39	Ridgway, Butler, 1967
	1.37	1.46	Sieg,1950
0.666	1.41	1.66	Este trabajo
0.400	1.54	1.77	Este trabajo

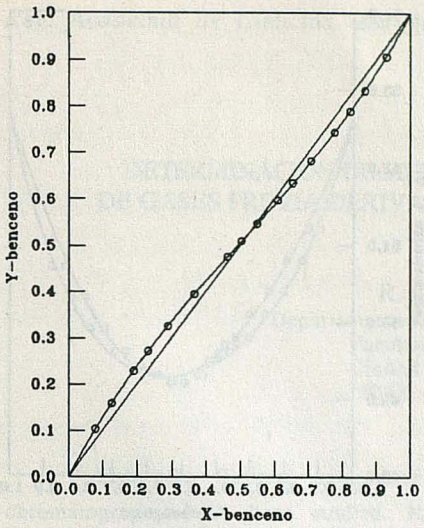


Figura 2. Diagrama X-Y para el sistema benceno-ciclohexano a 1.013 bar.

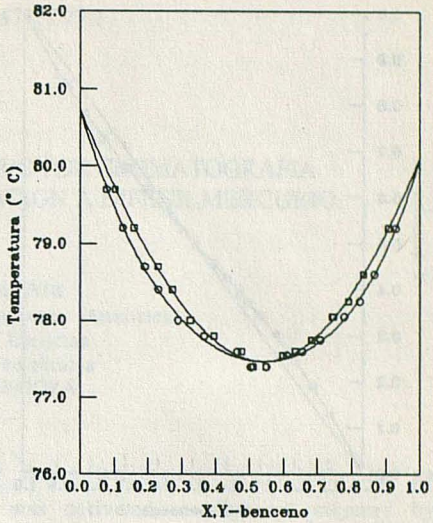


Figura 3. Diagrama T-X-Y para el sistema benceno-ciclohexano a 1.013 bar.

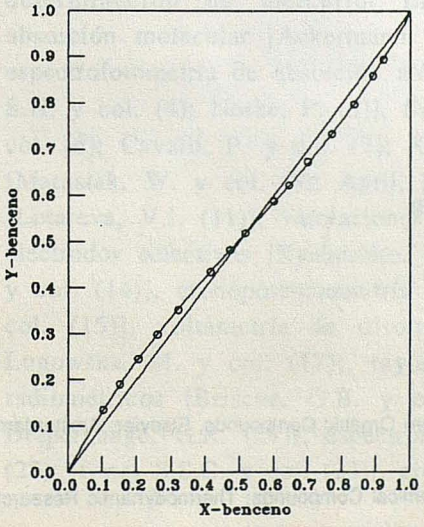


Figura 4. Diagrama X-Y para el sistema benceno-ciclohexano a 0.666 bar.

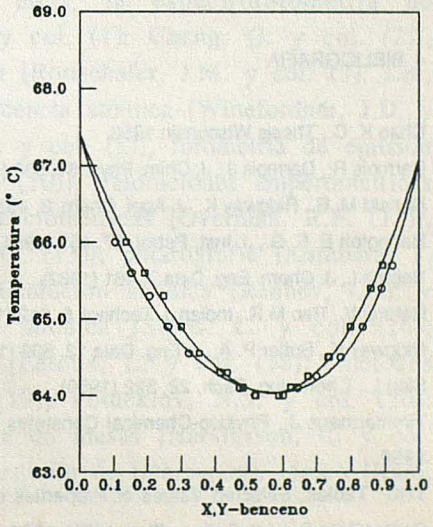


Figura 5. Diagrama T-X-Y para el sistema benceno-ciclohexano a 0.666 bar.

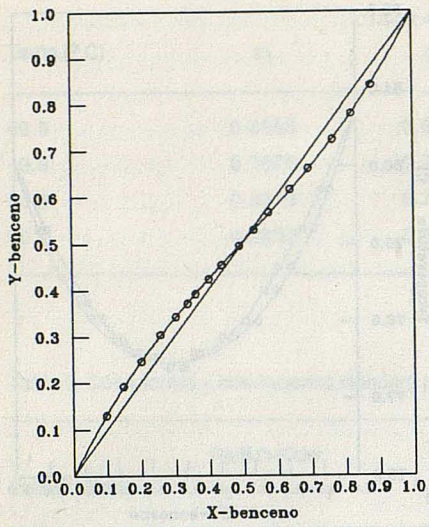


Figura 6. Diagrama X-Y para el sistema benceno-ciclohexano a 0.400 bar.

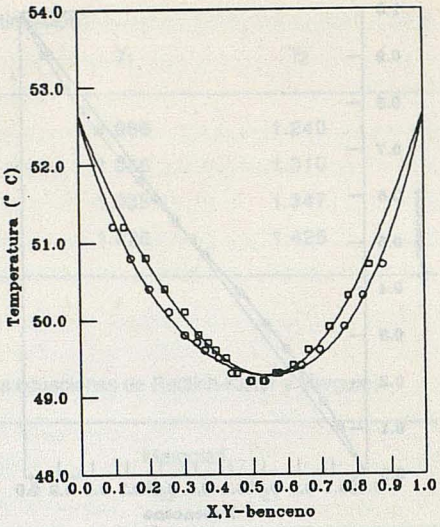


Figura 7. Diagrama T-X-Y para el sistema benceno-ciclohexano a 0.400 bar.

Los autores agradecen a la Diputación General de Aragon la ayuda facilitada para la realización de èste trabajo (Proyecto PCB 5/90). J. Pardo agradece especialmente su apoyo (Beca BCB 15/90).

4. BIBLIOGRAFIA

Chao K. C., Thesis Wisconsin 1956.

Darmois R., Darmois J., J. Chim. Phys. 61, 830 (1964).

Donald M. B., Ridgway K., J. Appl. Chem. 8, 403 (1958).

Herington E. F. G., J. Inst. Petrol. 37, 457 (1951).

Nagata I., J. Chem. Eng. Data 7, 461 (1962).

Nataraj V., Rao M.R., Indian J. Technol. 5, 212 (1967).

Ridgway K., Butler P. A., J. Eng. Data 12, 509 (1967).

Sieg L., Chem. Ing. Tech. 22, 322 (1950).

Timmermans J., Physico-Chemical Constants Of Pure Organic Compounds, Elsevier: Amsterdam, 1950.

TRC Tables; Selected Values of Properties of Chemical Compounds; Thermodynamic Research Center Data Project: College Station, TX, 1974.

Van Ness H.C., Classical Thermodynamics of Non-Electrolite Solutions, Pergamon Press: London, 1964.

Rev. Academia de Ciencias. Zaragoza **47** *(1992)*

DETERMINACION DE MERCURIO POR CROMATOGRAFIA DE GASES PREVIA DERIVATIZACION A DIFENILMERCURIO.

R. C. Diaz Ara
Departamento de Química Analítica
Facultad de Ciencias
Ciudad Universitaria
50009 - ZARAGOZA.

Abstract: Optimal parameteres for the determination of mercury by gas chromatography have been studied. Hg(II) was derivated to diphenyl mercury by addition of sodium tetraphenyl borate and was determined using a graph calibration. The effects of some interferences presents in vegetables have been determined. Chloride interferes only.

1.- Introducción.

Diversas técnicas instrumentales se han aplicado a la micro determinación de mercurio. Entre ellas, la espectrofotometría de absorción molecular [Ackermann, G. y col. (1); Cheng, G. y col. (2)], espectrofotometría de absorción atómica [Rottschafer, J.M. y col. (3); Lee, S.H. y col. (4); Neske, P. (5)], fluorescencia atómica [Winefordner, J.D. y col. (6); Cavalli, P. y col. (7); Xu, J. y col. (8)], fotometría de emisión [Matusiak, W. y col. (9); April, R.W. (10)], valoraciones amperométricas [Lotareva, V.I. (11)], valoraciones potenciométricas [Overman, R.F. (12)], electrodos selectivos [Ryabushko, H.K.O. (13)], polarografía [Kambara, T. y col. (14)], cronopotenciometría de disolución anódica [Rannev, G.G. y col. (15)], voltametría de disolución anódica [Sipos, L. y col. (16); Lugowska, M. y col. (17)], rayos X [Leroux, L. y col. (18)], métodos radiométricos [Briscoe, G.B. y col. (19); Poluektov, N.S. y col. (20); Delpergange, G.R. (21)], espectrometría de masas [Haraldsson, C. y col. (22); Tong, S.S.C. y col. (23)], cromatografía de intercambio iónico [Kato,

K. y col. (24); Livingston, H.D. y col. (25)] y cromatografía de gases [Ohkoshi, S. y col. (26); Nishi, S. y col. (27); Lansens, P. y col. (28)].

La cromatografía de gases es una técnica muy útil no sólo para la identificación de compuestos organomercuriales, sino para la diferenciación de especies de mercurio inorgánico y orgánico, lo cual es de una considerable importancia, ya que la toxicidad de los compuestos de mercurio depende mucho del estado de la combinación química. Por ejemplo, los compuestos de metilmercurio (II) son más tóxicos que sus análogos de fenilmercurio (II).

En el presente trabajo se han optimizado las condiciones cromatográficas para la determinación de mercurio por cromatografía de gases, previa derivatización a difenilmercurio. Así mismo, se ha estudiado la influencia de algunas interferencias en la determinación de mercurio en vegetales.

2.- Parte experimental.

2.1 Aparatos.

Se ha utilizado la siguiente instrumentación:

Cromatógrafo de gases Hewlett-Packard, 5720 A, con detector de ionización de llama. Microjeringa SGE, ARN, de 5 µl. Baño termostático Salvis. Burbujómetro para la medida de flujo de gas portador. Agitador mecánico Kotterman. pH metro Microprocessor Ionalyzer Orion 901.

Las condiciones de trabajo optimizadas para la determinación de Hg(II) por cromatografía de gases se ofrecen en la tabla 1.

TABLA 1

- Columna de acero de 50 cm. de longitud y 3,175 m. de diámetro
  Su designación completa es:
  10% UC W98 80-100 W A.W. DMCS B-20 L-5
- Temperatura de la columna = 180º C

- Temperatura del detector = 350º C
- Gas portador = Nitrógeno
- Flujo de gas portador = 85 ml/min.
- Gases para el detector = Hidrógeno y aire.
- Presión de alimentación de hidrógeno = 35 psi.
- Presión de alimentación de aire = 30 psi
- Sensibilidad = 10 x 4
- Velocidad del papel registrador = 0,38 m/h
- Cantidad inyectada = 5 µl

Tabla 1: Condiciones óptimas de trabajo con el cromatógrafo Hewlett-Packard, Modelo 5720 A.

2.2. Reactivos.

- Solución patrón de cloruro de mercurio (II) de 1000 ppm., a la que se añadieron 5 ml. de ácido nítrico para su conservación.
- Solución de tetrafenil borato sódico al 0,05% (m/v), recién preparada para evitar la descomposición del reactivo en medio ácido.
- Tolueno.
- Sulfato sódico anhidro.
- Soluciones de diversos cationes, para el estudio de interferencias. Se preparan de 1000 µg/ml., a partir de los correspondientes nitratos. Asi mismo se prepara una solución de 1000 µg/ml en cloruros a partir de cloruro de litio

Todos los reactivos son de calidad RA.

2.3. Procedimientos.

2.3.1 Derivatización de Hg (II) a difenilmercurio (II).

Debido a la relativamente alta volatilidad de los compuestos de mercurio y de que los fenilmercuriales (II) no evidencian descomposición alguna durante el proceso de volatilización, según se ha demostrado por técnicas termogravimétricas [Belcher, R. y col. (29)], se procede a la derivatización de cloruro de mercurio (II) a difenilmercurio, en base a la siguiente reacción [Uthe, J.F. y col. (30)]:

$$NaB(C_6H_5)_4 + 2HgCl_2 + 3H_2O \longrightarrow 2Hg(C_6H_5)_2 + 3HCl + B(OH)_3$$

La reacción se produce a la temperatura de 20º C. El difenilmercurio se determina por cromatografía de gases en las condiciones citadas en la Tabla 1. Se han obtenido resultados similares con nitrato de mercurio (II) como sustrato.

2.3.2. Extracción del difenilmercurio (II) a fase orgánica.

El procedimiento utilizado es el siguiente: En un embudo de separación se coloca 1 ml. de una disolución de mercurio de concentración conocida, 1 ml. de solución tampon HCl - KCl de pH = 2, 1 ml. de tetrafenilborato sódico al 0,05% (m/v) y 1 ml. de tolueno. Se agita durante 5 minutos, dejando separar las dos fases. Se inyecta en el cromatógrafo de gases una alícuota de 5 µl. de la fase orgánica previamente desecada sobre sulfato sódico anhidro.

3.- Resultados y discusión.

3.1. Influencia de la temperatura en la reacción de derivatización a difenilmercurio.

Se ha verificado la reacción entre el Hg(II) y el tetrafenilborato sódico a diferentes temperaturas: 20, 45, 75 y 100º C, utilizando un baño termostático. Solamente en el caso de que la reacción se produzca a 20º C se logra la completa resolución del pico correspondiente al difenilmercurio. A temperaturas superiores se produce la aparición de diversos picos, que se solapan con el del difenilmercurio, indicando la posible descomposición del tetrafenilborato sódico.

3.2. Estudio de la estabilidad del extracto del difenilmercurio.

Tomando una disolución de 10 ppm de Hg(II), se estudia la estabilidad del extracto del difenilmercurio en tolueno. Se inyecta una alicouta de 5 µl cada 15 minutos. Al cabo de 2 horas aparecen variaciones en las alturas correspondientes al difenilmercurio. La explicación fundamental de este hecho radica en la descomposición gradual del difenilmercurio, inestable en medio ácido, lo que lleva

consigo la aparición en el cromatograma de algún pico (B) extraño, tal y como se observa en la figura 1, influyendo además en la mala resolución de la señal correspondiente al difenilmercurio (A), que aparece solapada con la correspondiente al tolueno. Por ello, las determinaciones se realizarán en el acto o, a lo sumo, a las 2 horas de preparado el extracto en tolueno.

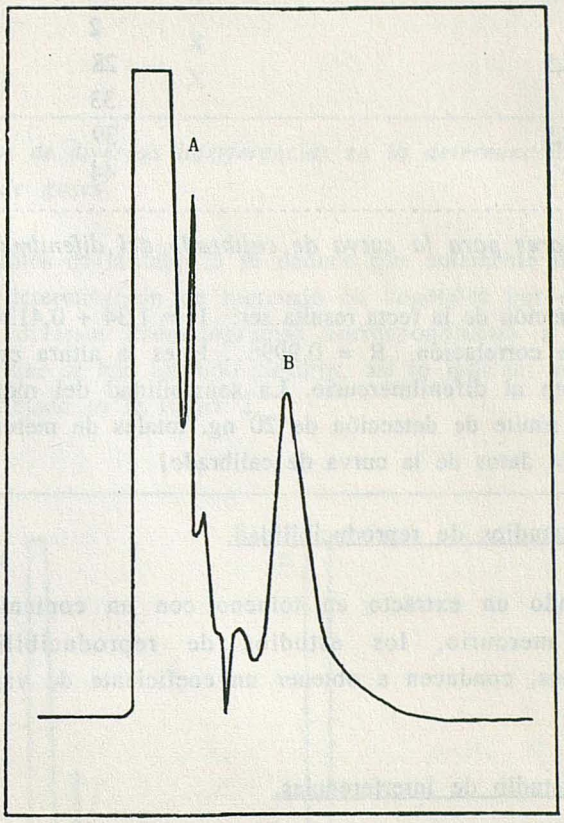


Figura 1: Cromatograma del extracto de difenilmercurio (A) en tolueno, en presencia de algún producto de descomposición (B) del tetrafenilborato sódico.

3.3. Curva de calibrado. Sensibilidad y límite de detección.

Se ha realizado una curva de calibrado entre 5 - 20 µg totales de mercurio. Los valores obtenidos para la altura del pico del difenilmercurio (HgO_2) aparecen en la tabla 2.

TABLA 2

µg totales de Hg	Altura del pico de HgO_2 (mm)
5	7
7,5	16
10	2
12,5	28
15	33
17,5	39
20	44

Tabla 2: Valores para la curva de calibrado del difenilmercurio.

La ecuación de la recta resulta ser: H = 1,34 + 0,4134 [Hg] , con un coeficiente de correlación R = 0,9996 . H es la altura en mm. del pico correspondiente al difenilmercurio. La sensibilidad del método es de 1,4 ng/mm, y el límite de detección de 20 ng. totales de mercurio, calculados a partir de los datos de la curva de calibrado.

3.4. Estudios de reproducibilidad.

Tomando un extracto en tolueno con un contenido de 15 µg totales de mercurio, los estudios de reproducibilidad, para 6 determinaciones, conducen a obtener un coeficiente de variación del 3,2 %.

3.5. Estudio de interferencias.

Se han estudiado los efectos de las posibles interferencias debidas a algunas especies catiónicas y aniónicas presentes en material vegetal, para una relación [Hg]/[X] de 1/1, 1/10 y 1/100, siendo X el ion interferente. Los resultados obtenidos se muestran en la tabla 3, que figura a continuación.

TABLA 3

Relación [Hg]/[X]	Interferencia [X]						
	Cl^-	Mg^{2+}	Ca^{2+}	Mo^{6+}	Fe^{3+}	Na^+	K^+
1/1	X						
1/10	X						
1/100	X						

Tabla 3: Efectos de diversas interferencias en la determinación de Hg por cramatografia de gases.

De los datos de la tabla 3 se deduce que solamente interfieren los cloruros en la determinación de mercurio en vegetales por cromatografía de gases. Los diversos cromatogramas correspondientes a patrones de mercurio, a los que se han añadido cloruros, así lo indican; un ejemplo de ello aparece reflejado en la figura 2.

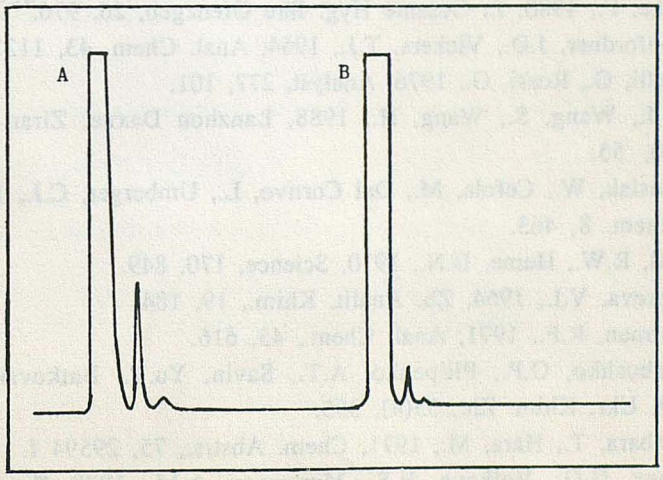


Figura 2: Cromatogramas de difenilmercurio en tolueno, en ausencia de cloruros (A) y en presencia de cloruros (B).

4.- Conclusiones.

Se han optimizado las condiciones para la determinación de mercurio por cromatografía de gases, a partir de la derivatización de Hg(II) a difenilmercurio, por reacción con tetrafenilborato sódico. El límite de detección resulta ser de 20 ng. totales de mercurio. Se ha estudiado la influencia de la presencia de algunas especies catiónicas y aniónicas que pudieran interferir en la determinación de mercurio en material vegetal con el método utilizado. Solamente el cloruro muestra una interferencia apreciable, para cualquier relación de concentraciones con respecto al mercurio total.

5.- Bibliografía.

(1) Ackermann, G., 1970, Z. Anal. Chem., 250, 353.
(2) Cheng, G., XU, S., 1989, Fenxi Huaxue, 17(2), 112.
(3) Rottschaffer, J.M., Jones, J.D., Mark, H.B., 1971, Chem. Abstr., 75, 15945 j.
(4) Lee, S.H., Jung, K.H., Lee, D.S., 1989, Talanta, 368(10), 999.
(5) Neske, P., 1980, Z. Gesante Hyg. Ihre Grenzgeb, 26, 976.
(6) Winefordner, J.D., Vickers, T.J., 1964, Anal. Chem. 43, 1127.
(7) Cavalli, G., Rossi, G., 1976, Analyst, 277, 101.
(8) Xu, J., Wang, S., Wang, H., 1988, Lanzhou Daxue, Ziran Kexueban, 24(3), 55.
(9) Matusiak, W., Cefola, M., Dal Cortivo, L., Umberger, C.J., 1964, Anal. Biochem. 8, 463.
(10) April, R.W., Hume, D.N., 1970, Science, 170, 849.
(11) Lotareva, V.I., 1964, Zh. Analit. Khim., 19, 184.
(12) Overman, R.F., 1971, Anal. Chem., 43, 616.
(13) Ryabushko, O.P., Pilipenko, A.T., Savin, Yu.T., Batkovskaya, L.A., 1989, Ukr. Khim. Zh., 55(4), 395.
(14) Kambara, T., Hara, M., 1971, Chem. Abstrs., 75, 29594 f.
(15) Rannev, G.G., Volkova, V.S., Murtazaev, A.M., 1970, Zavodsk. Lab., 36, 1446.
(16) Sipos, L., Valenta, P., Nurnberg, H.W., Branica, M., 1974, Anal. Abstr. 33(4), 372.
(17) Lugowska, M., Rubel, S., 1987, Chem. Anal., 32, 591.

(18) Leroux, J., Maffett, P.A., Monkman, J.L., 1957, Anal. Chem., 29, 1089
(19) Briscoe, G.B., Cooksey, B.G., Ruzicka, J., Willians, M., 1967, Talanta, 14, 1457.
(20) Poluektov, N.S., Zelyukova, Y.V., 1969, Zavodsk Lab., 35, 186.
(21) Delpergange, G.R., 1964, Chem. Abstr., 60, 16199 c.
(22) Haraldsson, C., Westerlund, S., Dehman, P., 1989, Anal. Chim. Acta, 221(1), 77.
(23) Tong, S.S.C., Gutenmann, W.H., Lisk, D.J.,1969, Anal. Chem., 41, 1872.
(24) Kato, K., Kakihana, H., 1965, Anal. Abstr., 12, 5051.
(25) Livingston, H.D., Smith, H., Stojanovic, N., 1967, Talanta, 14, 505.
(26) Ohkoshi, S., Takahashi, T., Sato, T., 1973, Bunseki Kagaku, 19, 1646.
(27) Nishi, S., Horimoto, Y., 1968, Bunseki Kagaku, 17, 75.
(28) Lansens, P., Meuleman, C., Leermakers, M., 1990, Anal. Chim. Acta, 234(2), 417.
(29) Belcher, J.A., Rodriguez-Vazquez, J.A., Stephen, W.I., Uden, P.C., 1976, Chromatographia, 9, 201.
(30) Uthe, J.F., Solomon, J., Griffitt, B., 1972, J. Assoc. Off. Anal. Chem., 55, 583.