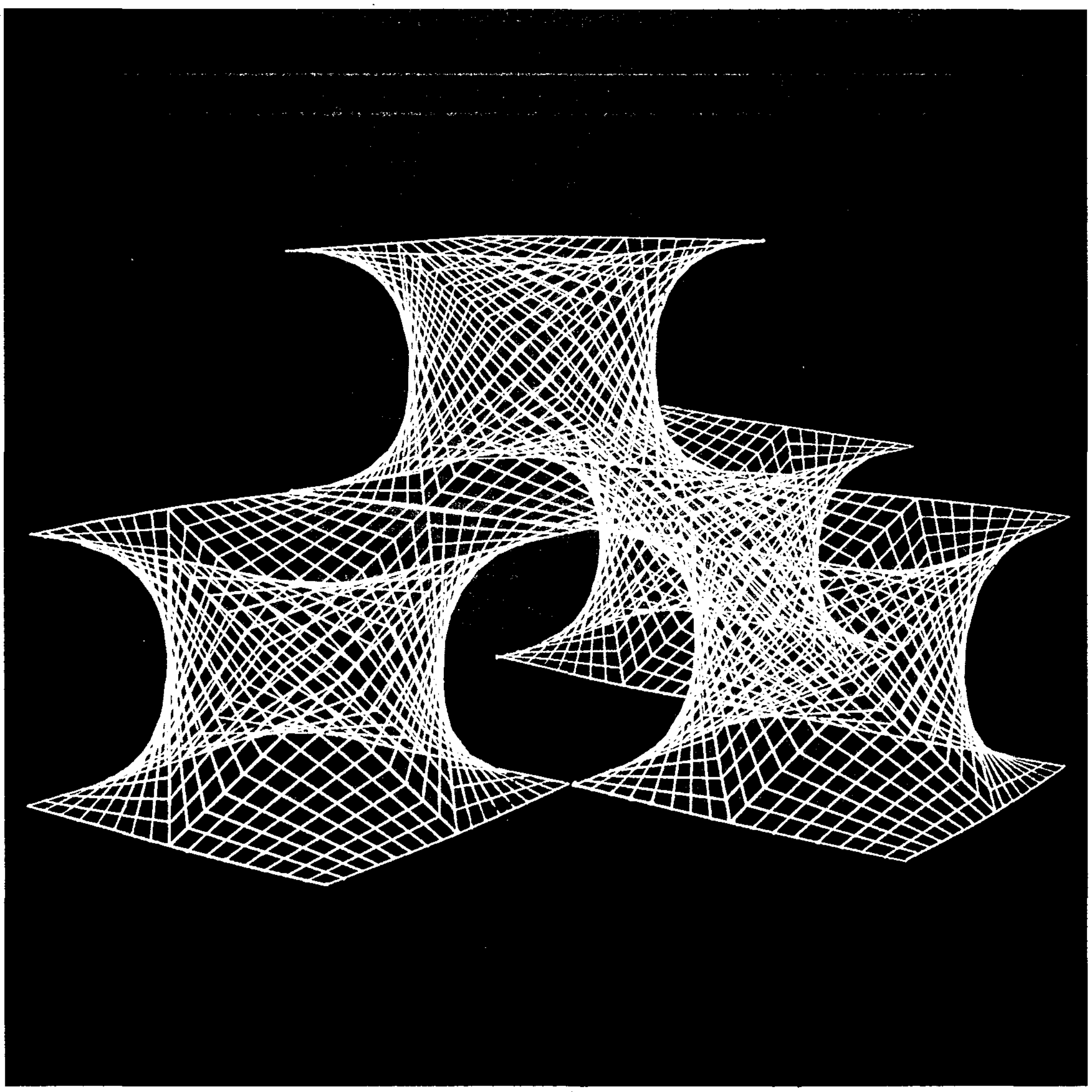
Symmetry: Culture and Science

ORDER / DISORDER
Proceedings, 4th Congress

The Quarterly of the
International Society for the
Interdisciplinary Study of Symmetry
(ISIS-Symmetry)

Volume 9, Numbers 2 - 4, 1998



See the *Advisory Board* inside back cover

Symmetry: Culture and Science

Founding editors: G. Darvas and D. Nagy

The Quarterly of the International Society for the Interdisciplinary Study of Symmetry (ISIS-Symmetry)

Temporarily edited by the *curatory* of the International Symmetry Foundation
Published by the *International Symmetry Foundation*

Volume 9, Numbers 2-4, 113-464, 1998

SPECIAL ISSUE:

ORDER / DISORDER
Organisation and hierarchy in science, technology, art, design, and the humanities

Guest Editors:
G. Darvas, D. Nagy and D. Shechtman

Selected papers, presented at the
Fourth Congress and Exhibition of ISIS-Symmetry,

Haifa, 13-18 September, 1998

SYMMETRY: CULTURE AND SCIENCE is *edited* by the Board of the International Society for the Interdisciplinary Study of Symmetry (ISIS-Symmetry), and *published* by the International Symmetry Foundation. The views expressed are those of individual authors, and not necessarily shared by the Society or the editors.

Any correspondence should be addressed to the
International Symmetry Foundation
c/o Symmetrion
P.O. Box 994, Budapest, H-1245 Hungary
Phone: 36-1-331-8326 Fax: 36-1-331-3161
E-mail: sym@freemail.hu, bercziszani@ludens.elte.hu, h492dar@ella.hu,

Subscription:

Annual membership fee of the Society:

Benefactors,	US$780.00;
Ordinary Members,	US$78.00 (includes the subscription to the quarterly),
Student Members,	US$63.00 (includes the subscription to the quarterly);

Institutional Members, please contact the Secretary General.

Annual subscription rate for *non-members*: US$100.00.

Make checks payable to *ISIS-Symmetry* and mail to G. Darvas, Secretary General, or transfer to the following account number: ISIS-Symmetry, International Symmetry Foundation, *10300002-50400004-21004015*, Hungarian Foreign Trade Bank, Budapest, Szt. István tér 11, H-1821 Hungary; Swift MKKB HU HB. Checks and transfers in Euro (€90.00 for Ordinary Members) are also accepted.

ISSN 0865-4824

Cover layout: Günter Schmitz;
Images on the cover:
from the book *The Periodic Table of the Polyhedral Universe* by Michael Burt,
Graphic design: Ami Korren
Front cover: Periodic polyhyparic surface of genus 3,
subdividing space between two cubic lattices into two identical subspaces
Back cover, top: Periodic table of the polyhedral universe, polyhedral mechanisms in the $g = 0$ domain
Back cover, bottom: Location pattern of proper polyhedral mechanisms
within the $g = 0$ domain and isotopes sharing V = 6
Ambigram on the back cover: Douglas R. Hofstadter
25368 Akaprint Kft. F. v.: Freier László

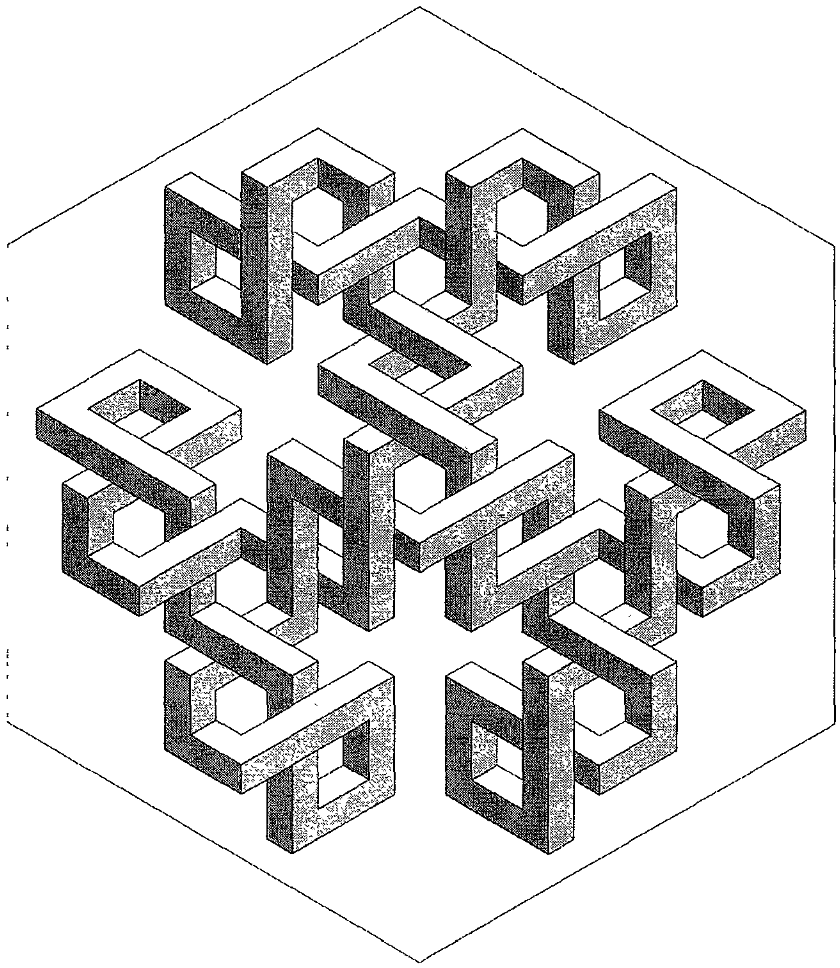
CONTENTS

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 117, 1998

EDITORIAL

This triple issue contains a selection of papers presented at the 4th International Congress and Exhibition of ISIS-Symmetry, *ORDER / DISORDER, Organisation and Hierarchy in Science, Technology, Art, Design and the Humanities*, held 13-18 September 1998 at the Technion - Israel Institute of Technology, Haifa. The Congress was presided by *Dan Shechtman*, professor of material science, Technion, Haifa; the honorary president was *Yuval Ne'eman*, professor of physics, Tel-Aviv University. The papers are arranged in alphabetical order by the name of the first author. Further papers presented at the Congress will be published in later issues.

The present editors apologise again for the long delay in the publication of these proceedings. We should remember the reader that one of the former editors resigned from the editorship in 1997, just for the unjustified delay in the publication of the journal. The next years provided an evidence for the reasons of the delay; under the control of the other editor, who took on the task to continue, no single issue was published in 4 years. When the Board of ISIS-Symmetry appointed the present editorial team in 2001, we should remedy this delay. We would like to express our thanks to all subscribers for their patience, and to all authors for their assistance and patience in preparing this issue.



Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 1998, 119-129

ORDERING PRINCIPLES IN 20TH CENTURY URBAN DESIGN APPROACHES

Iris Aravot

Name: Iris Aravot, B.Arch., M Sc., D.Sc. (b Romania., 1954)
Address: Faculty of Architecture and Town Plannmg, Technion I I T., Haifa 32000, Israel. *E-mail* aria@tx technion ac il.
Fields of interest: urban design, architectural philosophy, architectural education.

Abstract: *The paper presents a comparison of ORDER in twentieth century approaches to urban design. ORDER is discussed from a meta-content perspective, according to two counterparts: 1. order as arrangement; 2. order as command. Fifteen prevailing urban design approaches are examined in relation to the type and degree of physical order they propose. Then, their sources and power of authority are exposed and classified. Historical evidence supports the papers conclusion, namely that the actual "ordering capacity" of any approach may be directly related to the arrangement / command counterparts.*

1. PREFACE

Principles of physical order, usually intended as form generators of actual urban surroundings, are the ideological core of any U.D. (urban design) approach, and have been lengthy discussed (e.g. Gosling and Maitland, 1984; Broadbent, 1990). Thus, orthogonal forms, parallel housing arrangements and extensive open spaces are usually associated with Modernism and the first half of the century. Neoclassical free forms and spatial definition of continual public spaces are the characteristics of Post-Modernism and the second half of this century. Most Modernists' approaches regard urban physical order as a means towards universal solution of social problems, while (some) post-modernists seek the local cultures as bases for humanized urban places.

This paper does not intend to offer a similar clustering around the opposed poles of Modernism and (rediscovered) Classicism. Rather, it is concerned with the parameters of ORDER in U.D. approaches, regardless of their specifics. It raises questions such as: What is the locus of ORDER in various urban design approaches? How could ORDER be compared despite changing formal preferences? What could be the characteristics of ORDER in U.D. principles?

Therefore, our focus moves from urban form to the meanings of ORDER. Fifteen of the leading approaches to twentieth century U.D. are discussed as the empirical sphere for the above enquiry (The list of approaches is presented under "references").

2. TWO MEANINGS OF ORDER

ORDER as noun and verb has two major meanings: (1) arrangement and (2) command.

Order as arrangement is a state in which components or elements are arranged logically, comprehensibly or naturally, according to formulae, rules or laws. Order however, implies more than the state of affairs. It is value saturated in that it is the desired condition of society, peaceful and harmonious, as against chaos, disorder, confusion, mess, anarchism, jungle.

Order as command is an instruction that must be obeyed, a commission or request to produce or supply something. It sets a direction for action, and relays on the power of authority, convention, tradition or force.

The question raises if the two meanings are mutually related, if they complement each other or are mutually dependent. Language suggests it to be a cultural matter. In English, French and German ORDER contains the two counterparts, unlike in Hebrew, Arabic and Russian, where "arrangement" and "command" are referred to by two different words. (In Hebrew: *seder* vs. *tzav*; in Arabic: *nizam* or *tartib* vs. *amir*; in Russian *poaradoc* vs. *prikaz*).

Nevertheless, one is tempted to speculate that the imposition or even the reading of a certain order is always dependent on a degree of ordering as command. The exercise of authority is necessary simply in order to exclude other arrangements or alternative readings.

The speculation of mutual dependence has special importance for the discussion of ORDER in U.D. approaches, since epistemologically these are ideologies, and by definition must have their *sayings about what "is" and what "ought to be".* Accordingly, they must have two counterparts: one which enables description and analysis, and one which channels change according to a set of values. The first is connected to ORDER as arrangement. The second - to ORDER as command.

3. DEGREES OF ORDER AS ARRANGEMENT

Order as arrangement has itself two counterparts: *Distinction of objects* and *definition of relationships* between the objects.

Distinction of objects is the activity of clustering and differentiating within the Heraclitic flux of urban culture. It enables the reading of urban objects for further discussion, evaluation and actual intervention. These objects may be common, such as squares, public gardens and blocks, or unprecedented, such as "monuments" in Rossi's sense (1982) or "decks" *in Smithson's sense (1968).*

These urban entities fix the relevant *perspectives or aspects of observation* and the *suitable scales of reference.* The Athens' Charter, for example, took a functional perspective which required only general differentiation between "dwellings", "recreation", "work", "transportation" and "historic buildings" (Le Corbusier, 1943). Lynch (1960) presented five completely different urban entities based on human perception ("paths", "edges", "districts", "nodes" and "land-marks"). Alexander (1977) introduced 253 patterns, which clustered and differentiated every urban entity, from the general scale of cities in their natural regions to the details of furniture on a veranda.

The urban entities of each approach define also the width and *continuity of the field* under examination. Alexander (ibid.) covers the entire city and surroundings. Rob Krier (1979) - only the public realm. Bill Hillier (1984) relates only to streets.

Thus each approach has its set of urban "bricks of game". There is little sense to ask Hillier about the urban "locus" (Rossi, ibid.), just as it is unthinkable to consult Rem Koolhaas (1977) about "place making" (Norberg Schulz, 1980). Urban entities are not neutral or "objective". They embody value saturated presuppositions, such as application of socialist ideas (Howard, 1961), the salvation of European urbanism (Le Corbusier, ibid.), or a return to neoclassicist urban forms (Krier, ibid.).

detailed scale	detailed-large scale	large scale		
- "2" -	- "1" -	- "3" -		
Cullen	CIAM	Lynch	**extensive**	
	Rossi	O.M.A.	**field**	
	Alexander			
	New Urbanism			
	Critic. Regional			
	Howard	Team X		
				BREADTH
				OF
- "3" -				**FIELD**
Venturi			**a**	
Krier	ØØØ	ØØØ	**fragment**	
Rowe				
Hillier				

Table 1: Classification of urban design approaches according to ENTITIES.

This article, however, focuses on the discussion of the parameters of order in U.D. approaches, regardless of the specific contents of each approach, they presuppositions, etc. These are discussed only as examples.

Table 1 presents the categories for differentiation between entities of U.D., and a classification of the fifteen approaches under examination. The two basic categories are derived from the primary characteristics of any approach: (a) *the scale* of entities it offers; (b) the extension of area they may cover or "the *breadth of field*".

The scale of entities ranges between *detailed scale*, such as "The points at the top of the two domes of the Piazza del Popolo" (Bacon, 1967, pp.25) and *large scale*, such as "Metropolitan regions" (Alexander, 19977, pp.11). Some approaches cover the whole range of scales, others - only a part. The first are rich in quantity and hierarchy. The later may offer few or many entities, but never a rich hierarchy.

The relevant area for analysis or intervention may be a street or a square - i.e. *a fragment* of the city (Krier, ibid.), or an *extensive field* such as the endless urban surroundings of American sprawl (Katz, 1994).

Table 1 clarifies that most approaches which look at extensive urban areas use a whole hierarchy of entities, from detailed to large scale. Conversely, approaches relating to city fragments use only detailed scale entities. It goes without saying that the first approaches imply higher degrees of order than the later. There is an intermediate level of potential order, connected to approaches which cover extensive areas with a restricted hierarchy of urban entities. The three have been ranked accordingly as "**1**", "**3**" and "**2**". Approaches which offer only large scale entities have also been ranked "**3**".

Reference to city fragments by large scale entities is logically impossible, hence Ø.

3.1. Relationships among urban entities

The second counterpart of ORDER as arrangement is the *definition of relationships*, in our case among U.D. entities. Basically, such relationships may be strictly ordered or loosely suggested. Theoretically, they range between (I) one mathematical (geometric) rule which defines all relationships and (II) no definition of relationships among urban entities. Between these poles of extreme *unity* and *plurality*, there are theoretical possibilities of *geometrical, typological* and *hierarchical* unity and plurality. All these categories and *their joint combinations* form the framework for the proposed classification.

However, the empirical examination of U.D. approaches shows, that geometrical unity has never been proposed in itself. If a specific geometrical order is promoted - it is orthogonal, and it always implies a topological principle of separation by means of movement systems. Curiously enough - it is anti-hierarchical. C.I.A.M. wrapped the major transportation system in green spaces, intending to separate "dwelling", "recreation" and "work". (Le Corbusier, ibid.). More than fifty years later O.M.A. described an ideal introverted urbanism by means of the abstracted and functionally impoverished transportation grid of Manhattan. (Koolhaas, ibid.)

Therefore, in Table 2, the approaches of C.I.A.M. and O.M.A. are ascribed both geometrical and topological types of definitions. Regarding the degree of order - they are ranked only second ("**2**"), after Alexander ("**1**"). The "Pattern Language" of Alexander, although consisting of a plurality of rules, is nevertheless the most articulate set of relationships among U.D. entities, and it is extremely hierarchical.

unity **one principle/rule**	**plurality** **several principles/rules**	
- “2“ -		
C.I.A.M.		
O.M.A. (Koolhaas)		**a. geometry**
	Alexander (a+b+c)	
- “3“ -		**b. typology**
Team X, Hillier, Cullen,	- “1“ -	
Krier, Howard , Venturi		
Norberg-Schulz, New Urbanism		**c. hierarchy**
- “4“ -		
Rowe, Lynch, Critical Regionalism, Rossi, Coop-Himmelblau		
		d. none

Table 2: Classification of urban design approaches according to RELATIONSHIPS

Most approaches define a topological principle of relationship among their entities, underlaid by clear hierarchical order. For example: Howard's 1898 Garden City model was never meant to be built in the round or symmetrically, but the Crystal Palace had to be in the center, and the workshops - on the outskirts. The "web" and "stem" of Team X were recommended for neighborhoods, city centers, universities, etc. (Smithson, ibid.). In Table 2, this type of relationships has been ranked "**3**", just in front of the last and extreme category, which does not define any relationships (ranked "**4**"). The best example for the latter is Rowe's "Collage City" (1978). Actually, the designer's free combination of ideal urban entities, with Villa Adriana as the ultimate model, is the central principle of this approach. Critical regionalism might equally refrain from definition of universal principles, in favor of particular rules, namely local traditions of urban compositions. For example: New Gourna by Hassan Fathi (1973).

4. DEGREES OF ORDER AS COMMAND

U.D. approaches, regardless of their ordering principles (as arrangement), may acquire wide acceptance by the professional community, and used in practice, as instructions for application. Conversely, they may also be discarded as unimportant, relegated to the sidelines, forgotten, and eventually rediscovered, or not. For example: the return of Rob Krier to the writings of Camillo Sitte. (Krier, 1979, ch.1). In any case, the impact of an U.D. approach is *context* related. Contextual aspects are the economic situation, historical background, social needs, political ideologies, technological and material possibilities etc. of the potential users (in the broadest sense) of the approach. One major manifestation of context which expresses many of the above aspects, is the *source of legitimation of the approach.* It answers the basic question: why should I (we) accept this approach? Obviously, if the legitimation issue is not resolved convincingly - the approach will not be adopted. Additionally, the more powerful the source of legitimation - the less explicit its questioning.

In contradiction to the discussion of *order as arrangement,* which was by theoretical classification, *order as command* (focussing on legitimation) is characterized according to empirical typification. The categories of sources of legitimation, in a descending order, are presented in Table 3 (together with the approaches which have been supported by each type of the sources).

The highest level of authority (or power of order as command) is manifested when an approach is legitimized by a prevailing social / political ideology ("**1**" degree of authority). When Howard, for example, sowed his garden city ideas, they met the fertile soil of general conviction that a solution was acutely needed counter flight from the land and the overpopulation of towns. Howard clearly addressed "reality", hence his approach had an "obvious" source of legitimation.

Degree of authority	**Type of source of legitimation**	**Supported approach**
- "1" -	Prevalent political / social ideology	Howard, C.I.A.M.
- "2" -	Scientific research	Alexander, Hillier Lynch
- "3" -	Tradition	Cullen, Venturi,
- "4" -	Convention of discipline	New Urbanism, Critical Regionalism Krier, Team X
- "5" -	Philosophy	Norberg-Schulz
- "6" -	Personal preference / view	Rossi, O.M.A. Rowe (collage)

Table 3: Classification of sources of legitimation of urban design approaches

The second powerful source of legitimation is scientific research ("**2**" degree of authority). In Western culture, scientific justification entails natural low and objective knowledge. Everyone must agree about the truth value of the approach, and accept (?) the implied conclusions. For example: ignoring the five variables of urban perception discovered by Lynch (1960) is unthinkable. Their actual application, however, must change according to the changing context.

The third source of legitimation is a living tradition besides the professional community ("**3**" degree of authority). This is a shared outlook of a wider group, including the professionals and the users. Cullen for example, drew on the English tradition of the Picturesque, while Venturi explicitly addressed the American Strip.

Degree "**4**" of authority is attributed to a convention or tradition within the discipline of U.D. This is a shared outlook of a specific community, and is highly dependent on fashion, public relations, etc. Krier, for example, promoted (European) Historicism. This trend was enthusiastically embraced by theoreticians of the eighties, e.g., C. Jencks (1977, pp. 81-90), but had little influence on the built world. A controversial example may be Critical Regionalism, which is sometimes criticized as "architecture for the affluent". (Tailor, 1989 pp.19-35).

The second least power of authority, "**5**", is attributed to approaches which present philosophical texts as their source of legitimation. Norberg-Schulz (1980), for instance, establishes his entire approach on the Phenomenology of Heidegger. This has a rather restricted authority, first and foremost because an acquaintance with Heidegger's writing is a precondition to its acceptance.

Finally, the least powerful source of legitimation, "**6**", is personal preference or individual outlook. Rowe, for example, in his "Collage City", draws on a variety of sources to unfold his personal idea of free composition of historical (and future) precedents. One may share his view, or not. Similarly, O.M.A. "discover" Manhattanism as the prematurely neglected urbanism for the culture of congestion. One may share the preference of the artificial over the natural, or not.

5. U. D. APPROACHES - ORDERING CAPACITIES

The ORDERing capacity of an U.D. approach is a combination of its ordering principles referring to arrangement (U.D. entities and their mutual relationships), and its ordering power as determined by the source of legitimation of each approach. The first counterpart is a meta-content expression of the specific components of the approach, while the second one represents its context. Therefore, the ORDERing capacity of an approach may change according to circumstances (context). The following is a classification representing the current state of the art.

Table 4 simply summarizes the indices annotated to each approach. The results have a ratio significance. The indices serve only as a means to classify the approaches within several *groups of ordinal significance*. The grouping itself is tentative, but the general tendency seems to have some historical support.

The summary in Table 4 clearly produced two groups: group I and all the rest. Then, within groups II-VI, the last one is somewhat outstanding, while the others occupy a continuum of sum total indices.

Here one may ask: what is the meaning or reference of this grouping?

This paper suggests that the *correlative of the ordering capacity* of an approach is its *actual implementation in practice*, both three dimensional building and education. This is reasonably so, because actual implementation of a distinct approach

(a) is recognizable only when some form generating principles are offered (order as arrangement), and

(b) implementation is expectable only when there is a rather powerful source of legitimation (order as command).

Despite possible differences of interpretation, there seems to be no doubt concerning the ordering capacities of the Athens Charter, the Garden City idea and the Pattern Language (group I). They are incomparable to capacities of all other approaches. The Garden City model has been adapted to hundreds of new towns and suburbs all over the world (Stephen, 1991), the C.I.A.M. principles are the basis of post W.W.II Town Planning, and Alexander is probably the most comprehensive attempt to provide an empirical account of integrated urbanism.

approach	arrangement	command	sum total	group
Alexander	1+1	2	4	
Howard	1+2.5	1	4.5	I
C.I.A.M.	3+2	1	6	
Hillier	3+3	2	8	
New Urbanism	1+3	4	8	II
Cullen	2+3	3	8	
Venturi	3+3	3	9	
Norberg-Schulz	1+3	5	9	
Critical Reg.	1+4	4	9	III
Lynch	3+4	2	9	
Krier	3+3	4	10	
Team X	3+3	4	10	IV
Rossi	1+4	6	11	
O.M.A.	2+3	6	11	V
Rowe	3+4	6	13	VI

Table 4: The ORDERing capacity of urban design approaches

Group II includes approaches which are rarely as a group. Still, they have their established influence on practice, though for different reasons. Going on to group III, it is rather difficult to decide whether Venturi has less or more ordering capacity than Hillier (group II). Thus approaches of group III, might eventually change positions with those of group II. The same goes for groups IV and V, or III and IV. However, more extreme changes of location seem improper. The grouping is more like a Wittgenstein arrangement of "family resemblance" then a strict quantitative categorization. For example: the experiment of Aldo Rossi to establish an autonomous theory of U.D., including his particular references by "monument", "locus" etc., had little impact outside the world of academia. Although his position in Table 4 could be exchanged with that of Krier, for instance, it could not be exchanged with the position of Critical Regionalism, which is manifested in numerous urban projects of the Third World (e.g., Curtis, 1988 pp. 144-153, Fathi 1973). The most questionable result of Table 4 is the location of Lynch's approach. His "The Image of the City" is part of most architectural curricula. Retroactively, it is perhaps correct to attribute this approach implicit topological order (marked "**3**" in Table 2), and clearly classify its ordering capacity in group II. Finally, the "Collage" approach is indeed only one of several offered by Collin Rowe during the high time of U.D., in the eighties.

REFERENCES

Alexander, C (1977) *A Pattern Language*, New York. Oxford University Press.

Bacon, E (1967) *Design of Cities*, New York· Viking Press

Broadbent, G (1990) *Emerging Concepts in Urban Design*, New York· Van Nostrand Reinhold International

Cullen, G (1971) *The Concise Townscape*, London: The Architectural Press

Curtis, W (1988) *Balkrishna Doshi. An Architecture for India*, New York Rizzoli

Fathi, H. (1973) *Architecture for the Poor*, Chicago: University of Chicago Press.

Hilier, B (1984) *The Social Logic of Space*, Cambridge, Mass Cambridge University Press.

Howard E. (1961) *Garden Cities of Tomorrow*, London Faber and Faber (first published in 1898).

Jencks, C. (1977) *The Language of Post Modern Architecture*, New York. Rizzoli, pp 39-52

Katz, P.(1994) *The New Urbanism*, Cambridge· Harvard University Press

Koolhaas, R. (1978), *Delirious New York*, New York Oxford University Press.

Krier, R. (1979) *Urban Space*, New York Rizzoli.

Le Corbusier, (1942) *Charte d'Athenes*, Paris (reprinted Paris 1957)

Norberg Schulz, C. (1980) *Genius-Loci*, New York Rizzoli

Plunz, R. (1996) "New Urbanism / Old Urbanism", *AIA*, Vol 85, no 4, Apr. 1996.

Rossi, A (1982) *The Architecture of the City*, Cambridge Mass M.I T. Press.

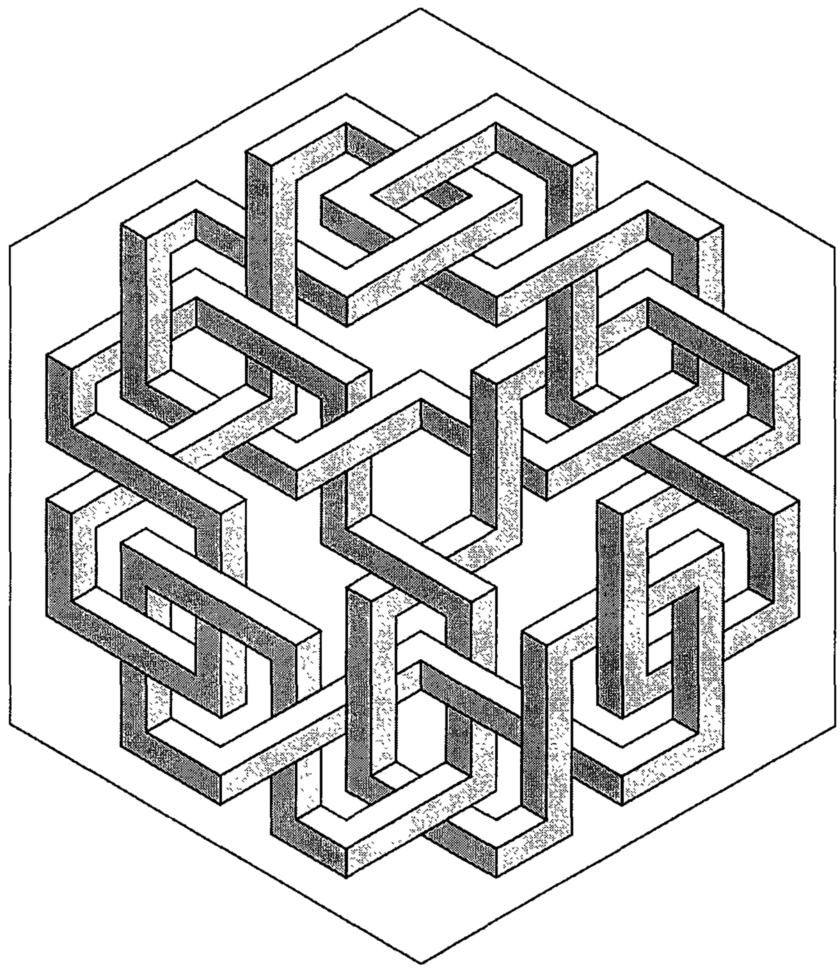
Rowe, C. and Koetter, F (1978) *Collage City*, Cambridge, Mass. MIT Press

Sitte, C. (1889) *The Art of Building Cities*, New York English translation. Stewart, C. T. 1945.

Smithson, A (1968) *Team X Primer*, London: Standard Catalogue.

Venturi, R , Scott Brown, D, and Izenour, S. (1979) *Learning From Las Vegas*, Cambridge, Mass . MIT Press.

Ward Stephen, V (1991) *The Garden City*, Cambridge, Mass



Symmetry: Culture and Science
Vol. 9, Nos.2-4, 131-134, 1998

THE MEASUREMENT OF SYMMETRY: BRIDGING ORDER AND DISORDER

David Avnir[(a)], Omer Katzenelson[(a)], Shahar Keinan[(a)], Mark Pinsky[(a, b)], Yariv Pinto[(a)], Yair Solomon[(a)] and Hagit Zabrodsky Hel-Or[(c)]

Address: The Institute of Chemistry[(a)], The Lise Meitner Minerva Center for Computational Quantum Chemistry[(a)] and The Institute of Earth Sciences[(b)], The Hebrew University of Jerusalem, Jerusalem 91904; and Department of Mathematics and Computer Science[(c)], Bar-Ilan University, Ramat Gan 52900, Israel. Email: david@chem.ch.huji.ac.il

The concept of symmetry has attracted virtually all domains of intellectual activity and has strongly influenced the sciences and the arts. It has functioned as a condensed language for the description and classification of the order within shapes and structures; as an identifier of inherent correlations between ordered structure and physical properties of matter; and as a guideline in artistic and practical aesthetic design. Our study of symmetry has been based on the following thesis: *Symmetry provides a rare realization of Nature, beyond the atomic level.*

The motivation of our studies has been rooted in the stark recognition that much more often then not, natural objects, are *not* symmetric. To realize it one should refine the resolution of observation - spatial or temporal - up to the point where it becomes evident. It appears that symmetry serves in many such instances as an approximate, idealized descriptive language of the physical world, beyond the scale of atoms. While it is true that an imprecise language helps in grasping complex situations and in identifying first order trends, the danger of missing the full picture because of the vague description, is always awaiting the user of the current symmetry language.

This motivation has led us to propose that it is natural to evaluate on a quantitative scale, how much of a given symmetry there is in a structure. Thus, we are treating symmetry as a structural property of continuous behaviour, complementary to the classical discrete point of view. A continuous symmetry scale should be able to express quantitatively how far is a given disordered structure from ideal symmetry, at any temporal resolution, at any spatial resolution, and with respect to any symmetry. Towards this goal, we have designed a general symmetry measurement tool, which is based on a definition which is minimalistic. Our answer to the question "How much of a given symmetry there is in a given structure?" is then:

Find the minimal distances that the points of a shape have to undergo, in order for it to attain the desired symmetry.

In order to translate this definition into practice, we have developed the Continuous Symmetry Measure methodology and computational tool. Using this measurement procedure it is possible to evaluate quantitatively how much of any symmetry exists in a non-symmetric configuration; what is the nearest symmetry of any given configuration; and what is the actual shape of the nearest symmetric structure (see Figure).

We have demonstrated the feasibility and versatility of our approach on two levels. The first one is purely geometric. Here we developed solutions for specific problems such as:

* evaluation of the degree of bilateral symmetry;

* measurement the symmetry content of distorted classical Platonic polyhedra;

* assessment of the symmetry content of objects which contain an element of randomness in their construction;

* analysis of the concepts of left/right handedness and mirror symmetry; and more.

The second level concentrated on applications of the symmetry measure to real problems of the natural sciences. Examples include symmetry analyses of molecules, of crystals, of dynamically changing structures, of small (3-12 molecules) clusters and of large disordered aggregates; of enzymes and their activities, and more. Three major findings emerged:

I. The symmetry measure describes Nature in a well behaved way: Its trends of change agree with intuition and reflect what is visible to the eye and what has been expected on a qualitative level, before measurement was possible.

II. Hitherto unknown quantitative correlations between symmetry and physical/chemical/biochemical properties have been revealed.

III. Far more than before, the importance of a global-shape descriptor for the quantitative observation and analysis of Nature, in distinction from the classically specific geometry descriptors, has been revealed.

In the context of this Symposium, it is important to emphasize that the Continuous Symmetry Measure methodology is applicable, beyond the molecular level analysis, to most other domains of the natural sciences, of the social sciences, and of the arts, where symmetry is an issue, either as a real feature or as an abstract one. Indeed, in separate papers in this Symposium we present the first quantitative assessment in archeology of the degree of symmetry of hand-axes of early man and its correlation to other parameters (Saragusti et al.); the assessment of the orientation of symmetric objects from their images - of relevance to design problems - and the evaluation of facial symmetry (Hel-or et al.,); and in yet another on-going project we explore the possible correlation between evolutionary selection, perception of beauty and attractiveness, and the quantitative degree of bilateral symmetry.

Readers are encouraged to contact us on issues of symmetry measurement in all domains of intellectual activities. We shall provide, free of charge, access to our computer programs. Readers with background in the natural sciences are directed for more technical details to references listed in http://chem.ch.huji.ac.il/~david/index.html.

Review: D. Avnir, O. Katzenelson, S. Keinan, M. Pinsky, Y. Pinto, Y. Salomon and H. Zabrodsky Hel-Or, "The Measurement of Symmetry and Chirality", in "Concepts in Chemistry", D. H. Rouvray, Ed., Research Studies Press, Somerset, 1996, Chapter 9, p.283.

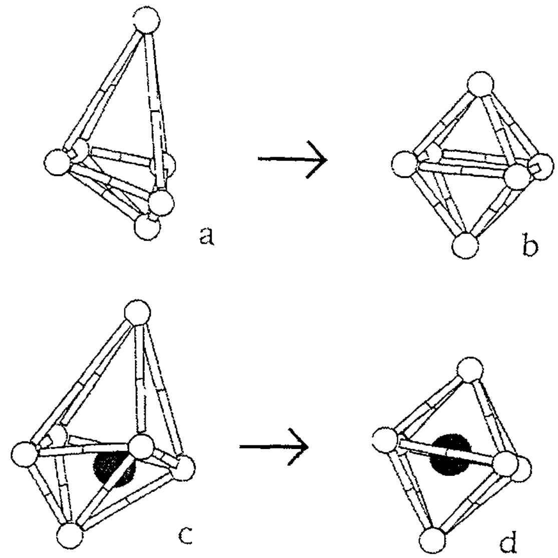


Figure: How much tetrahedricity is there in the distorted tetrahedron **a**?; and how much octahedricity in the distorted octahedron **c**? On a scale from zero (perfectly symmetnc) to 100, both have a symmetry value of 15. The polyhedra **b** and **d** are the nearest perfectly symmetnc objects to **a** and **c**, respectively. Having equal symmetry values, **a** and **c** are equally distant from being fully symmetnc - they are *iso-symmetric* objects.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 135-142, 1998

PHYSICAL ASPECTS OF ORDER IN VISUAL COMMUNICATION

Giuseppe Caglioti

Address: Dipartimento Ingegneria Nucleare, Politecnico di Milano and INFM - Politecnico di Milano. E-mail: caglioti@axp7000 cdc.polimi it, giuseppe caglioti@polimi it.

Keywords: *order, reduction of symmetry, symbols, perception, quantum mechanics.*

Abstract: *The forms of order encoded in some visual messages falling everyday under our eyes are analyzed and their impact on our perception is discussed. The fact that well structured visual symbols or messages are immediately and unambiguously interpreted irrespective of language or cultural biases is ascribed to the universality of natural laws: both the designer (emitter) and the perceiver obey natural laws; the professional designer engaged in the creation of a visual message finalized to promote a specific reaction instinctively finds the way to inject in it the specific forms of structural order enabling the image to fulfill its mission; since these forms of order are inspired by natural laws, the receiver can decode them easily and promptly while unraveling the meaning of the message during the process of perception.*

1. INTRODUCTION

Non verbal communication, based, e.g., on visual (or acoustic) images, is becoming more and more important in our global society. Nowadays people travel more than in the past. More often than in the past they change their country of residence or even their nationality. Society is becoming multiethnical and increasingly complex. Good social relations are necessary for its survival. And yet, as in the past, language barriers continue to hamper mutual understanding.

More and more frequently recourse is made to visual communication. Think for instance to traffic signals and to computer icons: here information must be transferred quickly, irrespective of the mother tongue of the perceiver. Fortunately, in practice, in most cases the visual message, even if it is structured in abstract form, proves to be clearer and more effective and immediate than the verbal message conveying the same meaning.

Visual messages are indeed universal. Like music, they overcome language barriers, they don't need to be translated in order to be grasped. We are tuned on them, we resonate with them, while in order to interpret their corresponding verbal message - whose understanding, in any case, requires some knowledge of the specific language in which it is expressed - we need much longer time.

What are the reasons for that *universality*?

The fact that well structured visual images strike directly and promptly our perception, the fact that they lend themselves to an unambiguous unraveling of their meaning implies that the *forms of order* embedded in their structures are archetypal and in most cases independent of the local culture of the designer.

Somehow the designer creating an icon or a traffic signal, finalized to promote a specific reaction, instinctively finds the way to inject in it the forms of structural order enabling the image to fulfill its mission.

The objective of this contribution is to show that *ultimately the designer's capability to create universally understandable images can be ascribed to the universality of the natural laws.*

Irrespective of the medium (visual, acoustic, ...) we use to communicate, we obey the laws of nature and instinctively apply them. Consequently the designer who is more faithfully capable to interpret the natural laws shapes the structure of his message in such a way as to orient the reaction of the receiver - who also promptly obeys natural laws instinctively - according to the function inspiring the message itself. While doing so he acts biologically, stimulating directly our sensory organs, without hindrances and slowdowns created by cultural biases such as those prevailing in the formulation of a verbal message in a specific language.

In this perspective, the intimate reason for the prompt, mutual understanding between the designer (emitter) and the perceiver of a visual message could thus be identify in the fact that they both operate on the common ground of the natural laws: the designer - most frequently by instinct - encodes such laws in his message and the receiver decodificates them while unraveling the meaning of the message during the process of perception.

2. THE ZEBRA CROSSING: AN EXAMPLE OF DISCRETE TRANSLATIONAL SYMMETRY IMPLYING *AND REQUIRING TO THE PEDESTRIAN* CONSTANT MOMENTUM (OR VELOCITY)

Recollecting our childhood, most of us remember the warning of our parents: *don't start running while crossing the road!*

The visual counterpart of this recommendation is faithfully codified in the form of structural order characteristic of the zebra crossing. (Fig. 1)

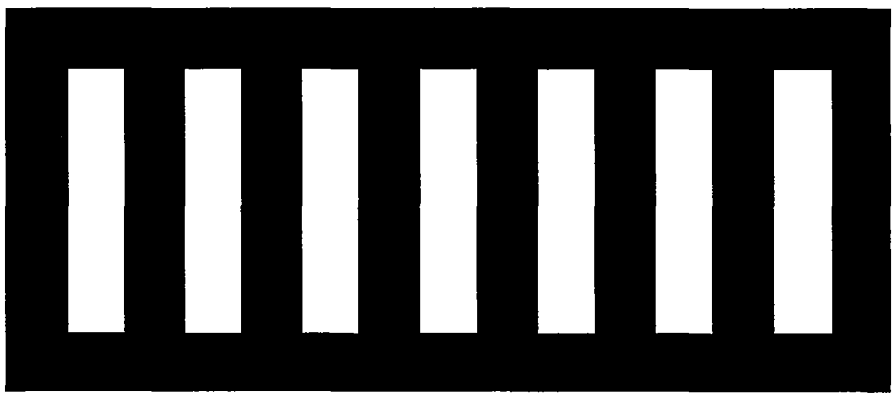


Figure 1: The zebra crossing

In fact the structure of this visual message is obtained by *reducing* the continuous *translational symmetry*, typical of the homogeneous dark gray coloration of the asphalt, to discrete translational symmetry: on the homogeneous coloration of the asphalt *any* translation - irrespective of its size - is a symmetry translation (so that once we have made the translation we are not aware of having made it); on the zebra crossing instead, *only* the discrete translation whose lattice spacing is the width of the white &dark-gray pair of alternating slabs is a symmetry translation. It is indeed this continuous → discrete translational *symmetry reduction* which is responsible for the *order* imprinted in the zebra crossing. But symmetry - a no change as the outcome of a change - is synonym of invariance. At any symmetry element in a structure corresponds a dynamical observable generating the symmetry element itself (we allude here to the so called constants of motion), an observable which keeps constant its value. A discrete translational symmetry is associated to a discrete set of possible constant momenta or, in practice, to a discrete set of constant velocities. Most naturally then, the order imprinted in the zebra crossing is perceived by the pedestrian as an order - or at least a strong recommendation - to maintain a constant pace without abrupt acceleration.

It might be appropriate at this point to recall that analogous arguments apply also to other media and to messages of different nature.

For instance, the repetition of the musical theme in a baroque fugue can be envisaged as a discrete translational symmetry of the structure of the score. And the word *fugue* evokes indeed a constant velocity.

Furthermore in his poems, in order to evoke constant velocity, Virgil systematically adopts pentadactylic hexameters. Try, e.g., to follow the rules of the metric (ictus on the bold type vowels at the beginning of any meter) while reading loudly

D**U**cite ab | **U**rbe do | m**U**m mea | c**A**rmina | d**U**cite | D**A**phnin

or

Qu**A**drupe | d**A**nte pu | tr**E**m soni | t**U** quatit | **U**ngula | c**A**mpum,

With easy gait, your mind will accompany Daphnia running away from the city back to his lover or, assisted by onomatopoeia (namely, here again, by the order associated to the discrete translational symmetry typical of the rhythm), will build the vivid mental image of a horse at a gallop.

The above examples indicate how order arises when translational symmetry is reduced from continuous to discrete. In general, *reduction of symmetry generates order* This aphorism applies to the whole range of conceivable scales: from the macroscopic scale of the objects perceived by our sensory organs all the way down to the atomic scale of quantum physics and subnuclear particles. However symmetry arguments are qualitative in nature and apply not only to all scales - allow me to insist on this crucial point - but also to all conceivable processes: to processes occurring around us in nature as well as to mental processes, such as perception, occurring in our brain (a brain which too is part of the nature.)

3. ANALOGIES BETWEEN THE MENTAL PROCESS OF PERCEPTION AND THE PROCESS OF MEASUREMENT OF QUANTUM STRUCTURES

In a previous paper presented at the ISIS-Symmetry Symposium at Tsukuba (Caglioti 1995), I have proposed that perception can be envisaged as an irreversible ordering process developing in our mind. During perception, the mind, driven by curiosity and attention, interiorizes the proposed image and controls it. The interiorized image evolves from a balanced, statistically symmetric and disordered state, where the stimuli suggested by the proposed figure act initially incoherently, to an unbalanced, more and more ordered state. As a result of this evolution, through a nonequilibrium dynamic instability our mind undergoes a disorder-order transformation, from a meaningless ensemble of uncorrelated signs, to *ordered thought* or *visual thinking* (Arnheim 1969). Incidentally, nothing, perhaps, is more ordered than the thought, notwithstanding the chaotic though statistically selfsimilar pattern of our encephalogram. At the critical point of this dynamic instability the interiorized image eventually comes into coincidence with the archetype of the proposed image, genetically or culturally impressed in our mind. Since the moment this critical state has been attained, the interiorized image behaves as a single entity reminding the quantum mechanics collective, macroscopic wavefunction describing e.g. the laser action resulting from the Bose condensation of the photons of an electromagnetic field optically pumped sufficiently far from thermodynamic equilibrium.

In the Tsukuba contribution quoted above a detailed analysis has been produced of the process of perception of a bistable ambiguous figure obtained by a *graphic condensation* of two cubic moduli. It was concluded that an analogy can be proposed between that perception process and the spectroscopic measurement of the charge transfer spectrum of the hydrogen molecular ion according to quantum mechanics: the analogy turns out to be so stringent that one feels confident to infer from it that the logic underlying the process of perception is the same as the logic of quantum mechanics. In quantum mechanics (as well as in the synergetic behavior of collective nonlinear open systems (Haken 1983)) a central role is assigned to the onset of order produced by symmetry breaking or symmetry reduction (as well as to ambiguity, i.e., to the confluence and coexistence of two incompatible aspects of a same reality - think, in particular, to the coexistence of order and disorder at the critical state of an equilibrium order-disorder transformation or a nonequilibrium dynamic instability.)

In this perspective, extrapolating a bit the above arguments, in what follows we propose that the mental process of perception and the process of measurement in quantum mechanics are governed qualitatively by the same formal rules.

4. THE PRINCIPLE OF INDETERMINACY AND THE TRAFFIC SIGNAL JUST PRECEDING A ZEBRA CROSSING: A REDUCTION OF THE DISCRETE TRANSLATIONAL SYMMETRY IMPLYING *AND REQUIRING TO THE DRIVER* A DECELERATION WHILE APPROACHING A PEDESTRIAN CROSSING

One of the most important characteristic of the process of measurement in quantum mechanics is the existence of an insuperable limit for the accuracy in the simultaneous measurement of pairs of observables whose product identifies an action: for instance it is impossible to measure simultaneously and exactly the position in space of a point-like material particle and its momentum (momentum is the product of mass by velocity) or to determine with unlimited accuracy its energy at an infinitely well known specific time, etc. The *indeterminacy principle* states that the product of the uncertainties in position and momentum of a particle cannot be lower than the Planck constant, the elementary action h.

If, via a measurement, we pretend to localize a particle with great accuracy, we cannot pretend to keep it still. Vice versa, if we want to keep a particle still, we must be ready to accept a large uncertainty in its position.

The above principle could help perhaps to assess the deep reasons underlying the pattern, reproduced in the previous page, of another traffic signal recently appeared along the European road network.

Every driver interprets instinctively this signal as a firm invitation - an *order* - to slow down, so allowing pedestrians to cross the road safely as the car approaches the zebra crossing. Here again, most likely, the order *slow down!* emitted by the visual message above is received correctly and promptly because of the fact that the form of order in the structure of the visual message itself is consistent with the qualitative implications of the natural laws of symmetry and of the indeterminacy principle.

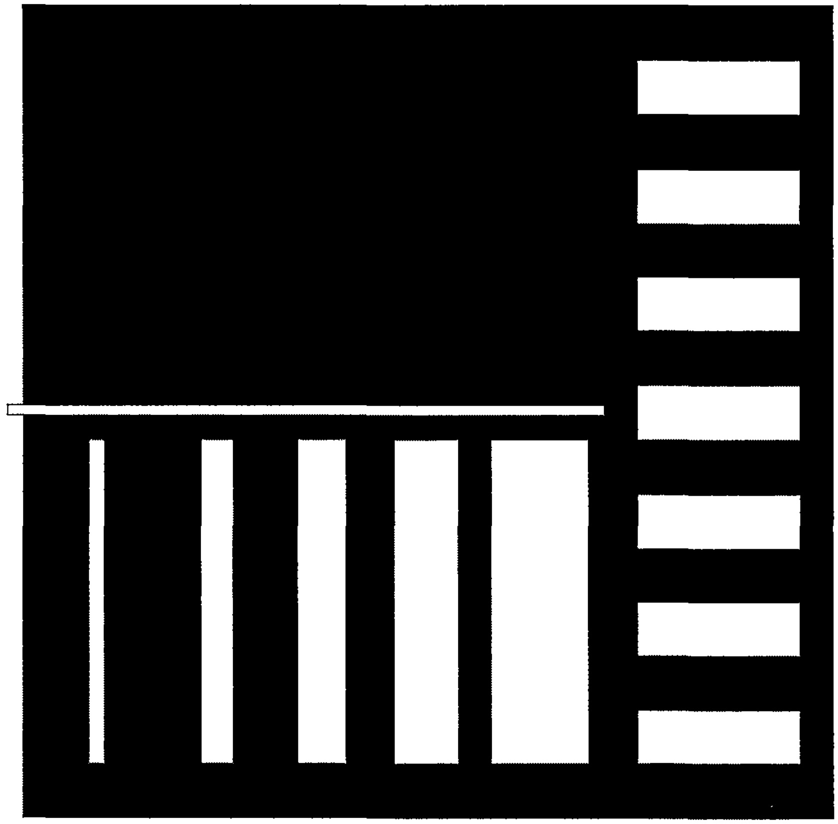


Figure 2

In fact the visual message can be conceived as obtained from the zebra crossing by altering the widths of the alternating white and dark-gray strips. The resultant succession of five white strips of increasing width, could be thus conceived as generated by an additional reduction of the discrete translational symmetry that in the zebra crossing evokes constant momentum and pretends a constant velocity by the pedestrian.

However, any symmetry reduction in a structure produces a form of physical order. Since translational symmetry implies constant momentum and pretends a constant velocity by the pedestrian, a change of lattice spacing implies a change of momentum and pretends a change of velocity by the car driver.

That this driver's velocity change should correspond to a *de*celeration is qualitatively suggested by the indeterminacy principle: according to it, thin strips correspond to large momenta (or velocities) while thick strips, near to the zebra crossing, correspond to small momenta (or velocities).

Similar arguments do apply to the ubiquitous visual message of the arrow:

$$\rightarrow$$

Here again the eye, after having loitered lazily on the body of this symbol (large space corresponds to low velocity) is instinctively brought to fly away as fast as possible from the pointed arrow's tip.

5. CONCLUSION

It is a pleasure to conclude this contribution with the visual message shown after the References.

REFERENCES

Caglioti, G (1992) *The Dynamics of Ambiguity,* New York: Springer-Verlag. (also available in Italian (CittàStudi1994 (2nd edition)), German (Wieveg 1990), Japanese (Kodansha 1997) and Russian (MIR 1998) editions)

Caglioti, G. (1995) From Perception to Thought, in *Katachi $\cup$ Symmetry,* Tokyo: Springer-Verlag.

Haken, H. (1983) *Synergetics. An Introduction.* 2nd edition, Springer-Verlag.

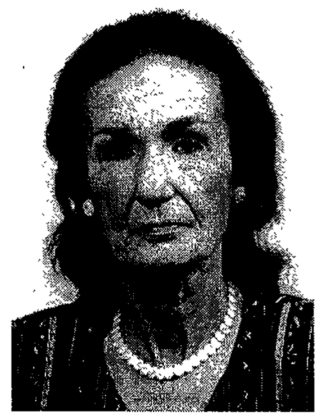
$$\Delta\Delta\Delta\Delta\Delta\Delta\Delta\Delta$$

Symmetry: Culture and Science
Vol. 9, Nos.2-4, 143-164, 1998

ORDER/DISORDER AS A FACTOR IN SHAPING A STYLE SCHEMATA OF REPETITION AND FORM

Dalia Cohen

Musicologist, (b Kibbutz Ein Harod, Israel, 1926).
Address: Department of Musicology, Hebrew Universıty of Jerusalem, Jerusalem, Israel. E-mail: baryosef@hum.huji ac il.
Fields of interest: Universals in music, music perceptıon and cognition, learned and natural musical schemata, style as determined by both the aesthetic ıdeal (of culture, penod, etc.) and cognitive constraints, vocal communication among humans and animals, ERP (Event-Related [Brain] Potentıal) responses to musical stımuli, and symmetry in music.
Awards: Frank Peleg Pnze, 1993, for the book *Comtemplatıon and Experıence ın Music Educatıon* (Jerusalem: Magnes Press, 1991).
Publications: Symmetry in Music, *Symmetry: Culture and Science*, vol. 8., no. 1.; Symmetry in Music: A Historical Perspective, *Symmetry: Culture and Scıence.*



Abstract: *In the present paper I have attempted to define repetition schemata in terms of the order of appearance (along the time axis) of similar and different sections and in terms of some of the structural properties of these sections, including order/disorder (of each section). I assumed that the concept of order is related to the sense of certainty regarding the musical progression; it contributes to "directionality" and may appear on various levels of musical organization. The type of repetition is essentially a system of rules of order on a higher level than in the isolated section. The characterization of repetition schemata is determined in light of their contribution to the types of "directionality" and complexity, which are some of the most important variables in determining the style and the "stylistic ideal." Characterization of the repetition schemata and types of directionality is determined by universal cognitive constraints, but the specific selection is determined in accordance with the stylistic ideal. These conclusions were obtained from theoretical considerations and musical analyses from*

various cultures and periods. As background for the definition of the repetition schemata, I first summarized the relevant concepts— "stylistic ideal" and "directionality and complexity" — and their main variables, in light of which the repetition schemata are examined.

INTRODUCTION

As is well known, most meaning in music derives from—or at least is related to—the manner of organization, i.e., types of order, unlike other arts, in which semantics plays a major role.

In music, the concept of order, which includes symmetry, is reflected most definitively, in our opinion, in the system of expectations and realizations of them (e.g., deviations from expectations) or in the predictability, to varying degrees of certainty, of the continuation of a musical progression on various levels of musical organization.

The concept of expectation and its realization as an important factor in shaping the musical experience occupies a more and more important place in theoretical and empirical musical research. Among the prominent examples are studies by Meyer (1956, 1973), whose ideas served as a basis for other research (particularly that of Narmour, whose formulations [1990, 1991] are being examined empirically), the experiments and summaries by Krumhansl (1990, 1993), and the studies by Granot (1996), who investigated brainwave responses (Event-Related Potential, or ERP) to series of musical stimuli. But the research is still limited to a few types of organization in certain selected parameters on the most immediate level, such as melodic or harmonic progressions in Western tonal music; no comprehensive explanations have yet been provided for the principles proposed; and hardly any attention has been paid to the ideals of the different styles.

Our assumption is that the type of order reflects not only a particular style (i.e., the principles of organization that can be realized in various musical pieces) but also the *stylistic ideal* (calm/tense, clear/unclear, static/dynamic, simple/complex). To our mind, this concept represents meta-principles that guide the specific choice of styles; it may vary among cultures, historical periods, geographical locations, and even individual composers; and it may serve as a reference point for both the composer and the listener. Thus, not only is order/disorder one of the general characteristics of the stylistic ideal, but it interferes with its variables.

In our study we tried to determine types of order/disorder, paying attention both to cognitive constrains and to the stylistic ideal, while looking at a selection of examples from the West and elsewhere. Here we will pay particular attention to the contribution of repetition to the type of order, against the backdrop of the relevant variables of the stylistic ideal, rules of excitement (as one of the variables of the stylistic ideal), and the main variables for determining order/disorder.

The paper will therefore be divided into two main parts: first, a summary of the assumptions regarding the collection of variables, and second, schemata of repetition with attention to all the variables.

ASSUMPTIONS

Most of the following assumptions regarding the relevant variables of the stylistic ideal, rules of excitement, and variables of order/disorder are based on my previous research, but there is, of course, much more to say about them.

A. The Main Variables in the General Classification of Styles by Ideal

1. *Relationship/lack of relationship* with extramusical factors. In the most general terms, we can say that the ideal that guides Western tonal music, in contrast to many non-Western musical cultures, considers the piece a sort of polished diamond that has many internal connections and is detached from its surroundings.

2. A *sense of calm versus tension* or emotions of various sorts. This factor, termed *ethos/pathos* by Sachs (1946), is part of what distinguishes different periods in Western music.

3. *Types of directionality and complexity* (Cohen 1994)

   3.1 The *range of "directionality,"* which allows for predictability concerning the continuation of the progression—from "momentary" directionality on the immediate level (as we found in many non-Western cultures and in some contemporary music) to "overall directionality" in a long-term superstructure. In vocal Renaissance music, for example, the primary directionality involves the unit of the musical phrase, which corresponds to the verbal sentence.

3.2 The *degree of clarity or lack of clarity* of directionality on various levels. When the directionality is clear we know where we are going, when we will arrive there, and how we will get there. In the most general terms, we can say that most of Mozart's works have clearer directionality than Bach's.

3.3 The *degree and range of complexity* of units of varying directionality (for example, in African polyrhythmic music there is a great deal of momentary complexity but a total absence of overall complexity and directionality; Western tonal music since the seventeenth century has been marked by complexity and overall directionality; in India we find complexity and directionality on the immediate level and that above it, as well as simple overall directionality).[1]

These variables may appear in various combinations, but with restrictions resulting from their interdependence.

B. Rules of Excitement (Cohen 1971, 1983)

1. *Intensification of energy* in the different parameters (higher, stronger, faster, etc.)

2. A *sudden change* in each parameter: In pitch, a large "leap" (melodic interval); in intensity, *sforzando*; in density, a sudden switch from quick to slow notes and vice-versa; and so on. There is good reason why, in a style whose ideal is tranquility (as in the rules of Palestrina counterpoint), there is strong limitation on the interval sizes, the "spaces" formed by them must be filled by stepwise motion (seconds) in the opposite direction, phrases must be started and ended with slow notes, and so on.

3. Moving to the two extremes in an *"inverted U function,"* based on the existence of a normative optimum in various parameters. Thus, in exciting music, the range of occurrence (in various parameters) may be very high or very low in the pitch register, density, intensity, ambitus, and so on. For example, the rules of organization in Renaissance music, whose ideal is calm and tranquility, are conspicuous in Western tonal music in that they are in the optimal range of neither too much nor too little; outside the West, Tibetan religious singing, which is supposed to transfer from the materialistic to the spiritual, is very tense, and it is conspicuous in that its register is extremely low and its ambitus is very narrow.

---

[1] The most complex and directional overall schema is probably expressed in the well-known "Schenkerian graphs," which relate to many schemata on various levels of the parameter of pitch only, in Western tonal music. This schema deliberately ignores stylistic differences, i e , differences resulting from deliberate distortions of directionality in accordance with the various stylistic ideals in the different eras.

4. *Deviations from expectations* (in addition to the deviations in point no. 3): breaking of learned schemata or deviation from expected natural schemata such as a convex curve and the rules of Gestalt, which essentially reflect limits to change (Lerdahl and Jackendoff 1983). This is very common in the Romantic period, and it is particularly salient in Schubert's later pieces, in which the deviation appears in the midst of a particularly directionality section.

5. *Uncertainty* of various sorts due to "nonconcurrence" (between simultaneous behaviors of various parameters or between different units), randomness (disorder), extreme equality, strong zigzag, and so on. Note that there is reason why the convex curve, which creates maximum predictability concerning the continuation of the progression, was dominant in the rules of Palestrina counterpoint regarding both pitch and duration, on the immediate level, on the level of the phrase, and even on the level of the piece as a whole (Guletsky 1995). Similarly, most folk tunes are convex (Nettl 1964; Huron 1997). Randomness, however, may in rare cases engender tranquility in the stage preceding the formation of basic measurable parameters such as intervals, durations, and beats, which are the important contributors to the existence of a system of expectations. For example, meditative music, which calls for absolute tranquility, is based on indefinability and randomness in non-salient events. All these tension factors may serve as a distinguishing mark of the styles in accordance with their ideal.

C. The Variables of the Rules of Order That Contribute to Types of Directionality

1. *The quantity of different elements:* In the short term, directionality and complexity are inversely related (increasing the quantity increases complexity and reduces directionality); in the long term, up to a certain point, the quantity contributes both to complexity and to directionality.

2. *Types of schemata.* Today we know that listening is always done in relation to schemata that take shape in our minds. Some of the schemata can be considered "natural" (e.g., the "Gestalt rules" and rules relating to texture [Cohen and Dubnov 1997]); others are "learned," culture-dependent, and not found in nature (e.g., scales, chords, and rhythmic patterns). (In our opinion [Cohen and Granot 1995], even the learned ones are not arbitrary; they are determined in accordance with the stylistic ideal and cognitive constraints.) Here we expand the concept of the schema to refer also to the most basic raw materials; to the rules of composition, such as various operations that can be considered cognitive; to meaningful curves of change such as the convex curve; and to types of texture. The schemata are the main factors in

shaping the order, the system of expectations, and types of directionality and complexity. Different kinds of styles are based on different types of schemata or different ways in which they are realized.

3. The *degree of definability* of the elements (pitches, intervals, chords, rhythm, and units) on different levels
4. *Different hierarchies* in schemata and ways in which they are organized
5. The *breaking of schemata* (deviation from expectations) detracts from directionality; when the breaking is entrenched one can talk about a schema of deviation. In extreme cases the deviation is like a "shock," and it stands out from the background of an "orderly" style.
6. *Concurrence or nonconcurrence* between different parameters or different schemata that are based on the same or different parameters. We can also speak of schemata of concurrence and nonconcurrence (Cohen and Dubnov 1997, pp. 400–401).
7. *Repetition:* Different kinds of repetition may contribute both to enhancement of directionality and to uncertainty.
8. *Form organization*

Each point deserves a special discussion, especially the schemata. In the present paper, we chose to focus only on the last two points: the types of repetition and one important aspect of the form that is defined as a schemata of repetitions with respect to order/disorder, such that the form represents a kind of hyper-order/disorder. In our discussion we refer to the other variables presented here without mentioning them by name.

REPETITION

The idea of repetition has served as the focal point of many studies (for an overview, see Ockelford 1991). Note that Meyer refers to an immediate repetition as a "natural" process: "Once established, a patterning *tends* [my emphasis—D.C.] to be continued until a point of relative tonal-rhythmic stability is reached" (Meyer 1973, p. 130). Boroda (1990) and Voss (1988) refer to some "natural" optimum (part of the "inverted U function") governing how often various units "should" be repeated in a piece of tonal music. (Voss discusses repetition of notes and Boroda discusses repetition of motives.) We should also note the emphasis on the importance of repetition to "order." For example, Ockelford (p. 139) states, "The source of perceived musical order lies in repetition." Taking a more extreme position, Feibleman (1968, pp. 3–5) argues, "Order can be identified with similarity and disorder with differences."

We believe it important to distinguish between different kinds of repetition (in addition to the three types pointed out by Schoenberg [1967, p. 9]: "The repetition may be exact, modified or developed"), and we shall try to characterize them in terms of their contribution to the characteristics of the stylistic ideal—directionality and tension.

Repetition refers to the minimum comparison between units of various kinds—only with regard to whether they are different or similar, which is a basic property of our cognitive activity (Tversky 1977). It is meaningful even aside from the "content" of the repeated unit. This meaning, which stems from similarity or difference, was corroborated in ERP experiments (studies of brainwave responses to various stimuli) in which the reactions to two consecutive patterns of identical or different notes were obtained (Cohen and Erez 1991). The result was a strong reaction to the first difference in the successive series, with the intensity of the reaction growing stronger the later the difference appeared in the series. Biologists explain the sensitivity to difference as a necessary condition for ongoing awareness of the environment for the sake of biological survival. In music, however, the situation is not so simple—not even on the most immediate level. For example, a strong ERP response is obtained (as expected) for every deviation from an expectation (as in a sudden violation of the schema). If a note is repeated, we may expect a change (due to the longing for a directional process or in order to prevent boredom), in which case a lack of change will be unexpected!

In music repetitions are manifested in various ways in different parameters and on different levels, so much so that one can speak of schemata of repetitions. Some types of repetition have even been given names, such as "*binary form*" (AA' or AA), "*ternary*" (ABA), "*rondo*" (with refrain, ABACA). Nevertheless, there is no consensus as to the interpretation and terminology of the types of repetition.

In general, in the most preliminary stage we can speak of six types of repetition.

1. A A' (or 2^n) one repetition
2. A A A A . . . many repetitions
3. A B A
4. A B A C A . .
5. "Mosaic composition"—random combination of units from a specific group (units are repeated with varying degrees of randomness)
6. "Negative repetition" or rarity—the appearance of a significant event *without* repetition
7. Simultaneous combinations as in the polyphonic form

In our opinion, the types of repetition should be further divided according to various criteria pertaining to the content of the units and their interrelations. (In reality there are also various combinations of types of repetition.)

The main criteria for classifying the types of repetition (in accordance with the "content" of the units and their interrelations) are as follows:

1. The "content" of the repeated and different units in terms of structural properties: the level on which they appear (immediate or deeper); their size; their definability with respect to the learned and natural schemata and with respect to their separation from their surroundings; their degree of divisibility; their status as "open" or "closed"; their function as background or as a frontal event; and their "orderliness."

2. What is considered different and what is similar? In what way are units that are considered different similar to each other, and in what way are units that are considered similar different from each other? The answers to these questions may vary widely among different styles from different cultures and periods.

3. The existence of directional links: between two similar units (AA') that are linked by a schema to form overall units (e.g., the harmonic schema V $\rightarrow$ I, in which the open unit A ends on V and the closed unit A' ends on the tonic I); between two different units (ABA), where B is an unstable development of A, such that the A that follows B is, in a sense, a resolution of B, in contrast to an independent B, which is stable and opposed to A; and links between the units in an entire series (e.g., the principle of intensification or the convex curve).

4. Precise or imprecise repetitions (in addition to the aforementioned repetition AA') with operations, with variations based on a single schema or "family resemblance" relationship, and so on.

Below we discuss only type 1 (AA'), type 2 (AAA . . .), and type 5 ("mosaic" composition), while referring minimally to the "content" of the repeated unit.

1. One successive repetition: This appears mainly in two ways:

1.1 A single repetition of a medium-sized or large directional unit, which is part of a movement (in a suite or sonata). This is found only in the West and enhances directionality. Directionality is greatest when the repetition is not precise, and the change represents a schema linking the two—A and A'—to form an overall unit. In the West the schema V → I (where A ends on V and A' on the tonic I) is widespread. It appears in the phrases of the "Classical period" and in the sonata form between the exposition and the recapitulation.

1.2 One imprecise repetition by means of extended doubling (1+1+2+4...)—which may be regarded as a natural schema—up to a certain size enhances directionality. This schema of 2^n, which can also be obtained through successive divisions by two (16, 8, 4, 2, 1, ½), is discussed in the West for the first time only in the eighteenth century (Ratner 1980). Until recently, it was regarded as a the main representative of symmetry phenomena in music. It is typical of many children's songs, of the Classical period in the West, and of various types of organization in the Far East, in both the past and the present. Thus, the repetitive, rhythmic *gongan* pattern (in Indonesian *gamelan* music) has 2^n beats ($n = 3–8$, meaning that the pattern may contain 256 beats!). The existence/nonexistence of 2^n is an important characteristic of the style.

2. Extended additive successive repetitions: These appear in many ways:

2.1 The repeated events are small units that serve as a measurable background (such as meter or accompanying patterns) that plays a role in defining the meanings of frontal events and make it possible to measure and compare the different units. This background (calm or excited) may promote clear directionality.

2.2 As frontal events, multiple exact repetitions in all the parameters produce tension (due to uncertainty regarding the continuation of the process) up to the "boredom threshold" at which the events turn into background. ("Equality" prevents hierarchy and directionality, and it serves as a contrast to "inequality" in many musical systems.) This is in contrast to information theory, which does not take into account the natural expectation of change. Thus, it is no wonder that it was forbidden in the Renaissance but common in the Baroque (Cohen 1971). Interestingly, this type of repetition was not taken into account in the various relevant theories (Meyer 1973; Narmour 1990).

2.3 Repetition that can be thought of as extending the duration of the repeated event, thereby making it more prominent (e.g., repetition of the target note in the main *balungan* melody, whose notes are equal in duration [Benamou 1989]).

2.4 Repetitions of complex, predetermined rhythmic patterns that also pertain to timbre. These can be thought of as learned schemata that combine meter and rhythm, and they are characteristic of non-Western cultures (e.g., the *tala* in Indian music and the *mizan* in Arabic music). The repeated patterns may serve both as background for a frontal musical event such as a melodic line or improvisation of the pattern itself; in the latter case the pattern in its various realizations is the frontal event. The directionality lasts throughout the pattern, and the improvisation may either sketch directionality on a level above the pattern or blur it.

2.5 Multiple repetitions of a series of non-directional events characterized by only one parameter, such as the 12-tone pitch series in the dodecaphonic system (also called "serial composition"). In contrast to the seven-tone diatonic system, the 12-tone equal-interval system (from which the seven-tone system is derived) contains no predetermined hierarchy. Moreover, the system of exact repetition of all pitches in a series (with the possibility of change in the octave state) prevents internal repetitions and makes it difficult to produce directionality. This is in accordance with the idea set forth explicitly by the initiator of the system, Arnold Schoenberg (1975), who recommended avoiding tonal schemata and repetitions of patterns within a row. In fact, however, whether overtly or covertly, composers do create relationships within series by means of various operations, or they limit themselves with respect to the series of intervals between notes. Thus one can speak of types of rows.

Despite theoretical studies on the system's inherent regularity, which has also been expanded to include serialism in parameters other than pitch (e.g., Babbitt 1960; Perl 1962), no research has yet been done on the meaning of repetition from the directional standpoint, taking into consideration the constraints of musical perception and cognitition. Therefore this extreme form of organization has not yet been fully explained.

2.6 Multiple repetitions with a shift in accordance with the elements of a directional schema such as a tonal scale, producing inexact units (the unit, which is like a link in a chain, appears on various degrees of the scale, and the process is known as a "sequence"). These repetitions have two opposing influences: the multiple immediate repetitions enhance the uncertainty as to the rest of progression, but the directional (unequal) schema to which the repetition is attached enhances directionality. The sequence is most often a transitional segment, in which case it may underscore the units before and after it and contribute to more general directionality. This was forbidden in the Renaissance and common in the Baroque. It is prevalent in Arabic music but not in *gamelan* music.

2.7 Repetition of a medium-sized, closed directional unit, with variations only, enhances directionality on the level of the unit but not overall. In most cases the series of variations has no overall schema that unifies the variations into a single unit (sometimes we find the intensification schema or a convex curve), but there is certainty regarding the schema of the next variation (which is uniform for all its variations). This form, which does not require long-range resolutions, is generally relaxing; naturally therefore, it often appears in the second movement of a sonata, balancing out the first, dramatic movement. The variations may also appear in succession, not as closed units, and with the repeated factor blurred. One expansion of the Classical variations—in which a single schema common to all is repeated—is a "family resemblance" relationship between successive sections. In this relationship the repeated factor is not fixed and may change from variation to variation, in which case the directionality is much less clear. In Bach's chaconne for violin solo, the two types of variations appear as complementary contrasts. The chaconne opens with a series of six variations, each of which is similar to the preceding one with respect to a different factor. After this non-directional series, Bach presents a clear series of seven variations based on a single harmonic formula (Figure 6).

2.8 Latent repetitions (not always successive), with various operations (such as expansion of the idea inherent in a motive over a large segment) enhances complexity and overall directionality, which is generally unclear.

A *striking example of contrast between the two types of repetition* that contributes to the overall structure is found in the famous Piano Fantasy in D Minor (Mozart, K. 397). The piece is composed of two diametrically opposed sections. In the first section (d-minor), multiple repetitions of small, non-directional units (repetitions of a single note, of a descending half-tone step, of a motive, etc.) produce tension; the second (D-major) section is characterized by an extremely large number of single precise repetitions of a directional unit.

3. Mosaic composition technique

Figure 1a

According to this technique, the musical structure is a juxtaposition of a limited number of predetermined basic units that recur with some possible variations in the combinations and with various degrees of freedom. The freedom in the arrangement of the elementary units hampers directionality on a level beyond the dimensions of the units themselves. This technique is prevalent in some non-Western music (e.g., Idelsohn 1944, Avenary 1963, Wellesz 1961, Cohen 1973) and is found in contemporary works, too.

In Figure 1 we see an illustration of mosaic composition, which governed the Israeli Arabs' oral tradition of singing neo-Byzantine liturgical hymns. The illustration is based on the cadential motives of the phrases in different performances of the same hymn.

FORM AS DEFINED BY "REPETITION SCHEMATA" WHOSE ELEMENTS ARE DEFINED BY MEANS OF ORDER/DISORDER

A specific form can be defined as a schema that represents a principle of overall organization and may appear in two areas: (1) a natural schema that represents various (meaningful) curves of change in the parameters; and (2) schemata of repetition that mean organization based on the difference and similarity between the units. Of course, each form's schema may be realized in many ways. In most styles the schema of the form is predetermined, as in the Classical period, when most pieces were governed by a single structure, the sonata form. However, we can find one-time overall structures even in tonal music, such as Bach's works. In atonal contemporary music most of the forms are one-time overall schemata. This in itself reduces the clear directionality. Some of the predetermined forms also relate to the content. The addition of content conditions naturally enhances directionality and makes possible a long-range increase in complexity. Interestingly, in the West the content combines with the form mainly through learned schemata (which relate to tonality and leave a great deal of freedom for the composer), whereas in many non-Western musical cultures the form is defined by means of natural schemata as well.

Here we present schemata of forms based on only two types of content: orderly and disorderly (hereinafter: O and D).

SOME EXAMPLES OF SCHEMATA OF FORMS WHOSE ELEMENTS REPRESENT (RELATIVE) ORDER AND DISORDER

"Disorder," which may result from many factors—indefinability of units, nonconcurrence, breaking of schemata, etc.—can be thought of as a tension phenomenon that "demands" resolution, like dissonance that is resolved to consonance. Just as dissonance may appear in order to highlight both consonance (as the theoretician Zerlino noted in the sixteenth century) and dissonance (in styles in which the ideal calls for expressions of tension), D may appear with various functions. Another analogy would be putting a tiny bit of salt on watermelon to bring out its sweetness or using the salt for its saltiness.

A	a	h	b	a	d	f					
B	c	h	b	h	ȷ	c	f				
C	c	h	f̃	c	ã	h	c	a	b	d̃	f
D	b	h	b	d	h	ȷ	a	f			
E	c	h	b	d	h	ȷ	f				
F	f	c	h	c	d̃	b	h	f̃	f		
G	b̃	h	c	b	h	ȷ	a	f			
H	c	h	b	h	ȷ	c	a	f			
I	b	h	b	d̃	f	c	b				
J	b	h	b	b	h	f					
K	b	h̃	b	d	b	h	c	b			
L	c	h	a	d̃	h	ȷ	f				
M	b	h	b	b	d	h	c	a	c		
N	b	h	b	d̃	h	ȷ	a	f			
O	c̃	h	b	d̃	h	ȷ	b	f			
P	c	h	b	e	h	ȷ	c	f			
Q	c	h	b	d̃	h	ȷ	c	a	f		
R	b	h	b	b	ȷ	b	f̃				
S	b	h	ȷ	a	b	h	ȷ	f			

Figure 1b

Figure 1: ["Mosaic composition" in the performance of the "Resurrection Hymn" in the third *lahan* (mode) by 19 Christian Arab singers in Israel. Freedom in selecting the order of the cadential motives in the different performances is limited. The restrictions are particularly marked in the finale and in the second phrase.]

The most directional schema in which D highlights O will be O, D, O (the realization formula of A, B, A; this can also be regarded as a convex curve of tension as a function of time. Indeed, there is good reason why this schema underlies the movement in the sonata form, which is the form with the most directional superstructure. In the Romantic period, in contrast, sometimes we find the opposite structure D, O, D, which can be represented by a concave curve, followed by a coda.

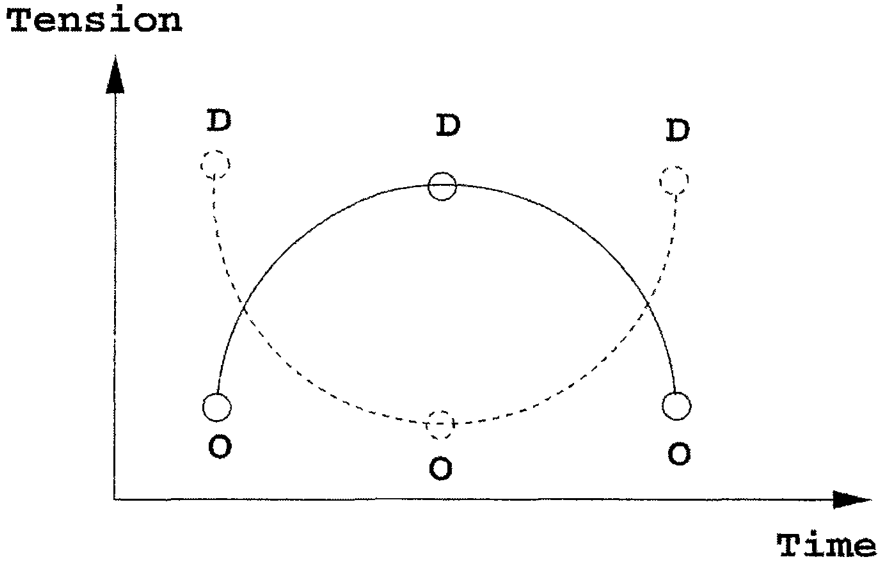


Figure 2: [Two realizations of the repetition schema ABA with O (order) and D (disorder)
— ODO in the Classical sonata form
- - - DOD in Romantic music]

Another common schema is a gradual transition from D to O by means of a transition from undefined to defined. This schema has various aspects. In non-Western cultures it is related to the schema of intensification in various parameters. It has even been given names: *alap-jor-jahala* in Indian music, *jo-ha-que* in Japanese music.

Thus, in some forms of Indian music, a piece opens with a slow, low, quiet section, with no meter or beat, and in extreme cases, without even a sense of duration and interval. Then directionality is extremely momentary and a "perpetuation" of time is achieved. The gradual, unclear beginning (so unclear that we do not know whether it is a tuning of instruments, a "warming up" of the musician's hands and soul, or the beginning of the piece) dismantles barriers between the piece and its surroundings in accordance with the ideal of Indian music and in contrast to the ideal that guides Western tonal music.

In Indian music we find a gradual increase in the definability of notes and intervals in the *raga*, of durations, and of the beat until the predetermined metric pattern—the *tala*—appears. This increase, as stated, entails intensification of all the parameters (including tempo, density, ambitus, intensity, and equality) that promote excitement, as well as uncertainty from the other direction. After the clear appearance of the meter (which is complex and repetitive, with improvisations all the way to the end of the piece) comes the resumption of an intensification process that increases equality and uncertainty. Without going into detail, we shall note that the tension curve, which is determined both by the intensification of the parameters and by the uncertainty, is in part parallel and in part opposite to the curve of definability.

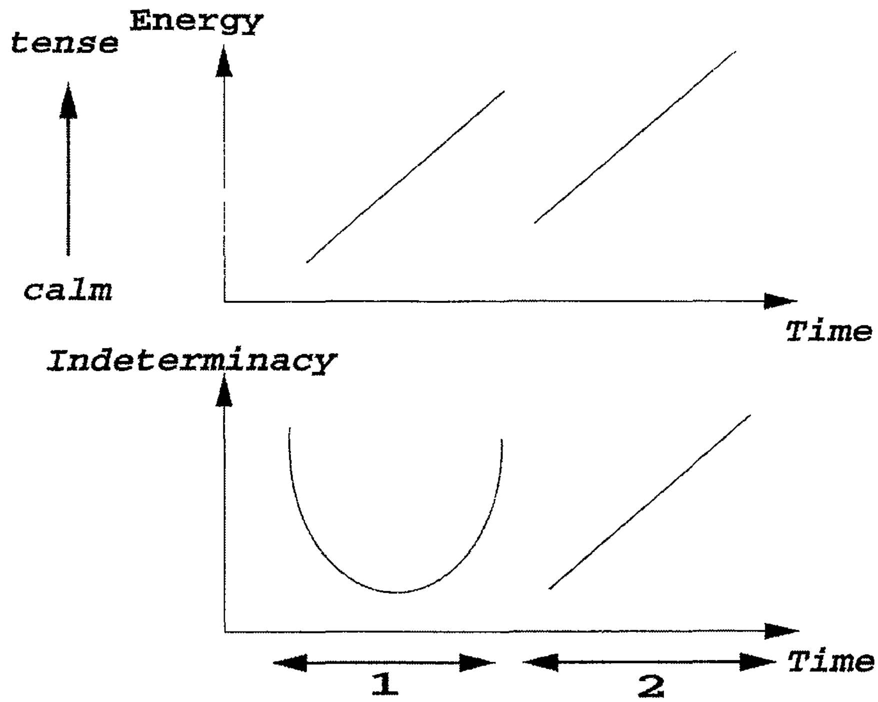


Figure 3· [The ratio of the intensification curve (following the *alap-jor-jahala* principle) regarding energy and tension to curves of uncertainty (disorder) in the performance tradition of the *raga* in northern India.
(a) The non-metric opening
(b) The metric (*tala*) part]

Another manifestation of the transition from D to O (from chaos to order) is typical of the works of Beethoven, who was guided by the principle of transition from indecision to assertiveness. In his music, however, the organization begins on a higher level than in

Indian music, and consequently the indefinability always produces tension. This process has various manifestations on different levels of organization and in different ranges.

First, let us mention the general schema of the sonata form with the two overall repetition schemata AA' (Exp. Rec) and ABA (Ex. Dev. Rec).

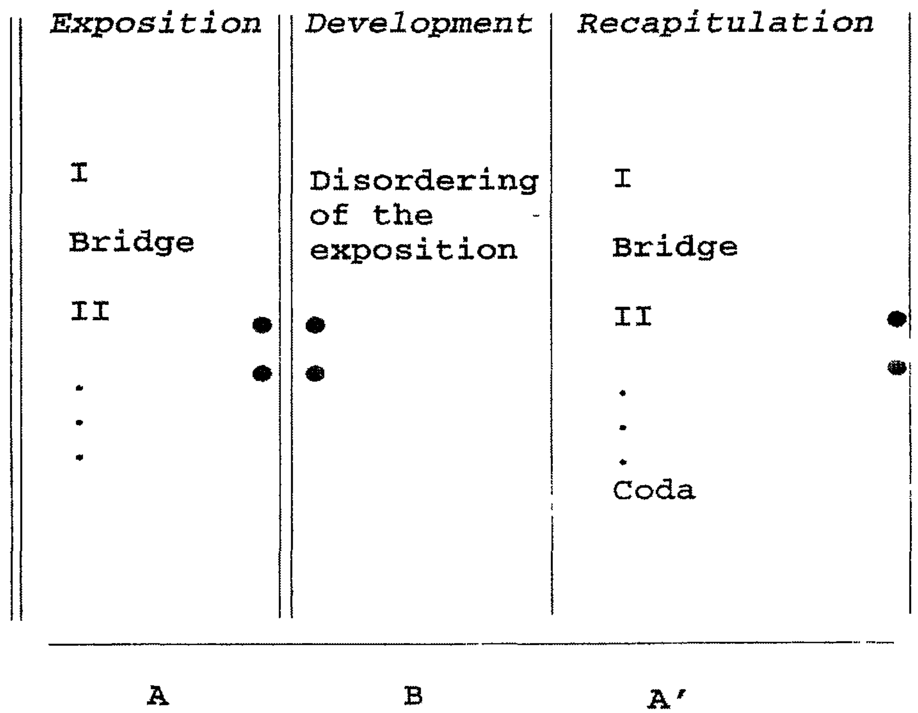


Figure 4: [Two overall repetition schemata—AA' and ABA—in the sonata form]

In Beethoven's work, we find various examples of the schema D → O within the general sonata schema, as in his First Piano Sonata in F Minor (Fig. 5): The first theme begins with no bass support, and later the accompanying chords appear, but not on the first (stressed) beat in each measure. In the recapitulation, the theme appears throughout with accompaniment only on the stressed beat. In the same movement, at the end of the first theme (measures 7–8), there is a descending series of six notes in the upper voice, beginning at melodic degree 5 and ending on the leading tone (as part of the dominant) without resolution to the tonic. This series, which was obtained through expansion of the ornamentation in the second measure, is repeated several times. In the bridge theme (in A-flat major) it appears three times in a row, first hesitantly for six measures, then clearly and concentrated in two measures, and the third time doubled in octaves and loudness. The series also appears (inverted) after the second theme and in the development section. At the end of the movement (in the coda) it is expanded to six measures with clear emphasis of each note and the addition of resolution to the tonic.

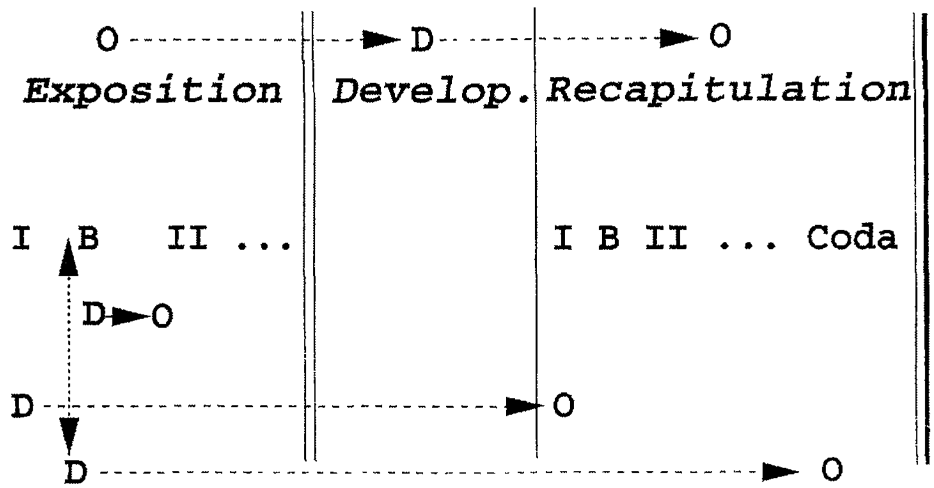


Figure 5b: [Summary of the D → O schemata in Beethoven's sonata (Fig. 5a) within the overall schema O → D → O (in addition to the binary schema AA)]

The peak of the schema D → O in Beethoven's works appears in his Ninth Symphony, where the schema is manifested explicitly on a few levels, within the movements (especially in the final movement) and in the piece as a whole, which opens with extended hints that are open to various interpretations.

Beethoven is an extreme example, Generally, in the Classical period we find the reverse situation: from order to disorder. However, we find D → O on extremely rare occasions in works by Haydn and Mozart, too. Haydn uses it explicitly in the beginning of *The Creation*, whose title proclaims a transition from chaos to order; it is also prominent in Mozart's Dissonant Quartet. In both cases the indefinability is manifested mainly with respect to harmony, which is not crystallizing into a clear schema and contains many violations of the rules.

Beethoven's Piano Sonata op. 2, no. 1

I (1st theme) Exposition

I Recap.

Coda

Bridge theme

Figure 5a: [Musical notation of several themes in the D → O process, from the first movement of Beethoven's First Piano Sonata (Op. 2, no. 1): between the first theme (I) in the Exp. and the Rec ; in the bridge (B); between the end of I and the coda.]

Another interesting manifestation of organization guided by the overall schema D → O is found in Bach's chaconne movement in the Partita in D Minor for Violin Solo. As we saw earlier, one of the manifestations of D → O is a transition from a series of variations that are related to each other through "family resemblance" to "regular classical variations" that are all based on the same harmonic schema.

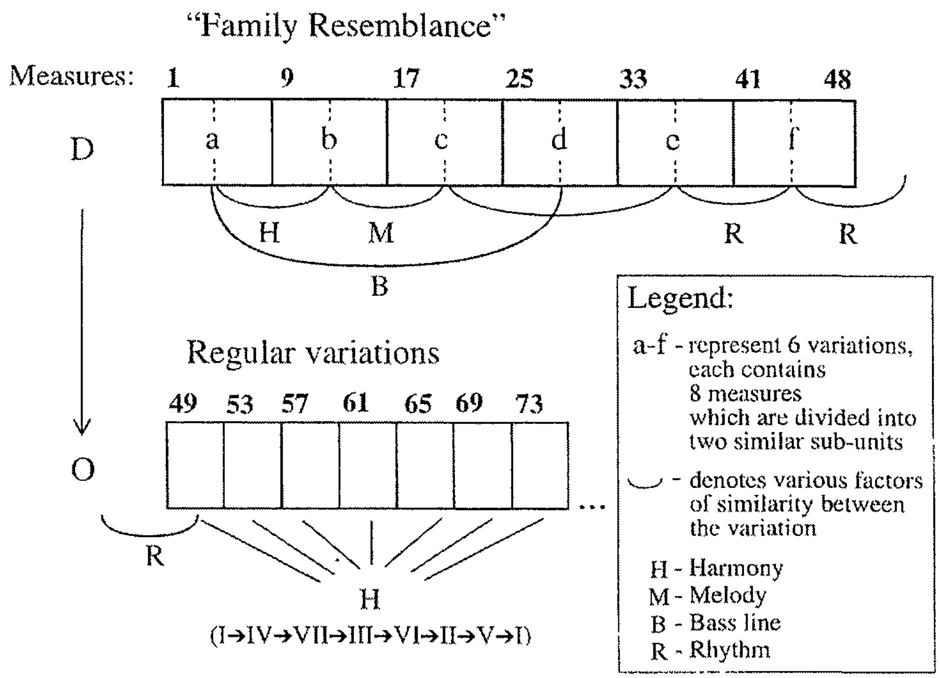


Figure 6: [The D → O ratio in the opening of Bach's chaconne for violin solo of a group of variations with "family resemblance" relationships to a group of four-measure variations that are all based on the same harmonic schema of descending fifths.]

To sum up, based on theoretical assumptions, many musical analyses, and empirical studies, we have tried to summarize in brief the variables of order/disorder (types of schemata, degree of definability, types of repetition, etc.) that are the determining factors in shaping a style. We suggested that in music the concept of order is reflected in the system of expectations and their realization and we paid attention to cognitive constraints and the stylistic ideal (e.g., calm/excitement and types of directionality and complexity). Of these, we focused on the variable of repetition against the backdrop of the other variables, and we examined examples from various styles in the West and elsewhere. We defined repetition schemata that contribute to types of order and proposed a way of determining schemata of forms by means of repetition schemata whose units are characterized by order/disorder (O/D); in this way varying levels of order/disorder are obtained. For example, the repetition schema A B A may be realized in various ways, of which O D O is the most directional (expected) and therefore the most orderly. In the form based on "mosaic composition," there may be order in the small units and disorder in the overall unit.

The types of forms (definition of which is just beginning) and of the repetition schemata were defined on the basis of general structural characteristics and therefore allow for comparisons and analogies with order in other, non-musical fields. Further research is, of course, necessary (as is attention to the psychoacoustic constraints on perception of specific parameters). Already, however, we believe that the results obtained thus far contribute to a deeper understanding of learned and natural schemata and of the concept of order on different levels and to expansion of the possibility of comparing styles.

REFERENCES

Avenary, H. (1963) *Studies in the Hebrew, Syrian and Greek Liturgical Recitative,* Tel Aviv: Israel Music Institute (IMI).

Babbitt, M. (1960) Twelve-tone invariants as compositional determinants. *Musical Quarterly,* 46, 246–259.

Benamou, M. L (1989) *Approaching and Hanging: Metrical and Melodic Organization in Javanese and Western Music,* [Master's Thesis], Ann Arbor: University of Michigan.

Boroda, M. G. (1990) The organization of repetitions in the musical composition Towards a quantitative-systemic approach. *Musikometrika,* 2, 53–105.

Cohen, D. (1971) Palestrina counterpoint: A musical expression of unexcited speech. *Journal of Music Theory,* 15, 68–111.

Cohen, D. (1983) Birdcalls and the rules of Palestrina counterpoint: Towards the discovery of universal qualities in vocal expression *Israel Studies in Musicology,* 3, 96–123

Cohen, D (1994) Directionality and complexity in music. *Musikometrika,* 6, 27–77.

Cohen, D. and Dubnov, S. (1997) Gestalt principles in musical texture, In: Leman, M., ed., *Music, Gestalt and Computing: Studies in Systematic and Cognitive Musicology,* Berlin: Springer, 386–403.

Cohen, D. and Erez, A. (1991) ERP measurements of cognitive components in response to pitch patterns. *Music Perception,* 8, 405–430.

Cohen, D. and Granot, R. (1995) Constant and variable influences on stages of musical activities: Research based on experiments using behavioral and electrophysiological indices. *Journal of New Music Research,* 24, 197–229

Crozier, J. B. (1974) Verbal and exploratory responses to sound sequences varying in uncertainty level, In: Berlyne, E., ed., *Studies in the New Experimental Aesthetics,* 27–87.

Feibleman, J. K (1968) Disorder, In: Kuntz, P. G., ed., *The Concept of Order,* Seattle: University of Washington Press, 3–5.

Granot, R. (1991) *The Concept of Musical Expectancy Examined by Brainwave Recordings (ERP) and Verbal Responses,* [Ph.D. Dissertation], Jerusalem: Hebrew University of Jerusalem.

Guletsky, I. (1995) Proportion in Palestrina's masses, In· *Proceedings of the Third International Conference of Study of "Palestrina e L'Europa"* [Palestrina, October 1994]

Hargreaves, D. J. (1986) *The Developmental Psychology of Music,* Cambridge: Cambridge University Press.

Huron, D. (1997) The melodic arch in Western folksongs. *Computing in Musicology,* 10, 3–23

Idelsohn, A. Z. (1944) *Jewish Music in Its Historical Development,* New York: Tudor.

Knophoff, L. and Hutchinson, W (1983) Entropy as a measure of a style: The influence of sample length *Journal of Music Theory,* 27, 75–97.

Krumhansl, C. L (1990) *Cognitive Foundation of Musical Pitch,* New York. Oxford University Press.

Krumhansl, C. L. (1993) Modeling melodic expectation: Using three "musical forces" to predict melodic continuation, In: *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society.* Hillside, NJ: Lawrence Erlbaum Associates, 629–634.

Lerdahl, F. and Jackendoff, R. S. (1983) *A Generative Theory of Tonal Music,* Cambridge, Mass.: MIT Press.

Meyer, L. B. (1956) *Emotion and Meaning in Music,* Chicago: University of Chicago Press.

Meyer, L B. (1973) *Explaining Music,* Berkeley: University of California Press.

Narmour, E. (1990) *The Analysis and Cognition of Basic Melodic Structure,* Chicago: University of Chicago Press.

Narmour, E. (1991) The top-down and bottom-up systems of musical implication: Building on Meyer's theory of emotional syntax. *Music Perception,* 9, no. 1, 1–26.

Nettl, B. (1964) *Theory and Method in Ethnomusicology,* London: Free Press of Glencoe, Collier-MacMillan.

Ockelford, A (1991) Repetition in perceived musical structures, In: Howell, P., West, R., and Cross, J., eds., *Representing Musical Structure*, Cognitive Science Series, London: Academic Press.

Perl, G. (1977) *Serial Composition and Atonality,* 4th ed., Berkeley: University of California Press.

Sachs, C. (1946) *The Commonwealth of Art,* New York: Norton.

Schoenberg, A. (1967) *Fundamentals of Musical Composition,* London: Faber and Faber.

Schoenberg, A. (1975) Composition with twelve tones, In: *Style and Idea*, 2nd ed., New York: Faber, 214–249.

Tversky, A (1977) Features of similarity *Psychological Review,* 84, no. 4, 327–352.

Voss, R. F (1988) Fractals in nature: From characterisation to simulation, In: Peitgen, H. O. and Saupe, D., eds., *The Science of Fractal Images*, New York: Springer.

Wellesz, E. (1961) *A History of Byzantine Music and Hymnography,* London: Oxford University Press.

Symmetry: Culture and Science
Vol 9, Nos. 2-4, 165-178, 1998

ORDER FROM RHYTHMIC ENTRAINMENT AND THE ORIGIN OF LEVELS THROUGH DISSIPATION

John Collier[1] and Mark Burch[2]

Addresses: [1]Department of Philosophy, University of Newcastle, University Drive, Callaghan, NSW 2308, Australia. Mailing address. Box 209, Carp, Ontario, Canada K0A 1L0 E-mail: ag659@ncf.ca, URL: http://www.newcastle.edu.au.
[2]P.O. Box 1328, Pearl City, HI 96782, USA. E-mail: burch@hawaii.edu.

Abstract: *Rhythmic entrainment is the formation of regular, predictable patterns in time and/or space through interactions within or between systems that manifest potential symmetries. We contend that this process is a major source of symmetries in specific systems, whether passive physical systems or active adaptive and/or voluntary/intentional systems, except that active systems have more control over accepting or avoiding rhythmic entrainment. The result of rhythmic entrainment is a simplification of the entrained system, in the sense that the information required to describe it is reduced. Entrainment can be communicated, passing information from one system to another. The paradigm is a group of jazz percussionists agreeing on a complex musical progression. The process of rhythmic entrainment is complementary to that of symmetry breaking, which produces information. The two processes account for much, if not all, of the complexity and organization in the universe. Rhythmic entrainment can be more or less spontaneous, with the completely spontaneous form being uncontrollable. A balance between the two forms can produce a more robust system, requiring less energy to maintain, whether in physical, biological or social systems. We outline some applications in physics, chemistry, biology, measurement and communication, ending with the especially interesting case of social and economic order. First though, we must introduce some basic principles.*

1. INTRODUCTION

Rhythmic entrainment is the formation of regular, predictable patterns in time and/or space through interactions within or between systems that manifest potential symmetries. We contend that this process is a major source of symmetries in specific systems, whether passive physical systems or active adaptive and/or voluntary/intentional systems, except that active systems have more control over accepting or avoiding rhythmic entrainment. The result of rhythmic entrainment is a simplification of the entrained system, in the sense that the information required to describe it is reduced. Entrainment can be communicated, passing information from one system to another. The paradigm is a group of jazz percussionists agreeing on a complex musical progression.

Rhythmic entrainment can either be forced (driven) or spontaneous (self-organizing). Forced entrainment can be either high power or low power. In high power entrainment, one powerful system drives another through immediate force, e.g., a boat's movements on storm waves at sea. Low power forced entrainment is of interest because it depends more on persistence and careful application of force than on immediate power. An example would be driving a large oscillator (say a swing) with small applications of force just off a node ("pumping" the swing). Forced cases always transfer pre-existing order. Forced resonance can be destructive, as when a singer shatters a glass by driving it at or near its resonant frequency too strongly.

Spontaneous entrainment creates *new* symmetries via the dissipation of energy and/or information. Systems tend towards minimal energy and tend to organize themselves so as to minimize dissipation (and consequently loss of available energy within the system – self-organization tends to increase efficiency). This process increases higher level order, or symmetry, and is mutual among the parts of the system, with excess energy being dissipated externally, unlike many cases of forced resonance. Simple examples can be found in resonances in the solar system resulting from tidal dissipation. Resonance tends to reduce dissipation and lower the energy of the solar system, as in all other cases of self-organization. We argue that similar processes are widespread, and that more complex cases can direct energy more efficiently than similar forced systems, allowing more effect for less effort.

Some symmetry is relic from either earlier undifferentiated conditions and/or deep universal principles, but individual systems are usually individuated through the production of information that distinguishes them from other systems (Collier 1996).

Rhythmic entrainment is a counterpoint and complement to the production of information by symmetry breaking, though similar principles are involved. In particular, both symmetry breaking and rhythmic entrainment, when spontaneous, are the result of dissipative forces (of which friction is a paradigm). The two processes are responsible for much (if not all) of the complexity and organization in the Universe.

We will outline some applications in physics, chemistry, biology, measurement and communication, ending with the especially interesting case of social and economic order. First though, we must introduce some basic principles. This will be rather brief, but necessary to discuss the examples.

2. REVIEW OF BASIC PRINCIPLES

The notion of information places a central role in our treatment. The basic idea of information is that of a distinction between two things. In standard language the notion is restricted to recognized distinctions, or at least ones that are in a position to be recognized, but information theory, as it has developed in abstract mathematical form, does not restrict itself to just meaningful distinctions, but to any distinction. This idea has three roots: i) logic, which can be traced back to Leibniz at least, but reaches its fullest form in the algorithmic complexity theory, which gives a measure of information in terms of the minimal number of distinctions needed to identify something uniquely, ii) physics, going back to Maxwell and his demon, but expressed more clearly by Szilard (1921), Schrödinger (1946) and Brillouin (1962), and finally iii) communications theory, due to Shannon (1949). We will have little to say about the last because of its highly specialized nature. Ideally, the three approaches should be unified, but such a unification is still in the future. One thing that can be said, is that any unification must be able to explain how information can be dynamically, or causally based, with the logical and communications theory forms being abstractions.

We therefore first focus on the connection between information and effort. Producing information requires effort, or work, which in turn requires available energy, sometimes called exergy. Maxwell recognized that the statistical account of the Second Law of Thermodynamics, that the entropy of an isolated system does not decrease, in all probability, would be violated by a sorting demon that could sort fast and slow molecules. Szilard (1921) showed that such a demon was impossible, because to get the information to do the sorting, the demon would have to expend at least as much exergy as would be gained by the sorting. Schrödinger (1946) suggested that information of the sort found in biological and other organized systems was negentropic, and this idea was codified by Brillouin (1962) as the Negentropy Principle of Information (NPI).

NPI implies that in order to do a measurement, work must be done, and exergy dissipated. Not only that, but any formation of order requires the dissipation of an equivalent or greater amount of exergy. More general proofs for computational systems were given by Landauer (1961, 1987) and Bennett (1982), who showed that a sorting demon would have to have an infinite storage place for waste information in order to work; erasure leads to lost information and consequent entropy increase. Collier (1990) gave a proof by reductio that a dynamical demon could not reverse the flow of entropy without some supernatural or very lucky source of information. The Second Law is empirical, but the connection to information through the arguments for the impossibility of a sorting demon establishes that producing information requires work. Conversely, dissipation of energy leads to a loss of information.

Recent work in logic sheds some light on the relation between information and causation. George Spencer Brown (1969) developed a logic of distinctions that has been shown to be equivalent to predicate logic (Banaschewski 1977, Cull and Franck 1984). Following work by Solomonoff (1964) attempting to develop an information based epistemology that encodes knowledge as minimal descriptions, Kolmogorov (1965, 1968) and Chaitin (1975) showed that information can be expressed as the minimal length of a program that can produce a string that isomorphically maps the yes-no answers to a series of questions that uniquely specify some thing. Basically, following Brown's work, the string is a truth table row that distinguishes the object uniquely, and the information content of the table is the length in bits of the minimal program (of a certain specified type) that can produce the table. This measure of information is equivalent (up to an additive constant) to the probabilistic or combinatorial forms that can be derived from Shannon's work. The connections between information, computation and probability allow a rigorous definition of probability in terms of the compressibility of strings.

Given NPI, and the reasonable assumption that all properties supervene on causal properties (that is, there can't be two worlds with the same causal properties that differ in other properties), causation is equivalent to the transfer of the same instance of information (Collier 1999). The only way new information can appear is through work, but information can dissipate spontaneously. This notion of causation guarantees that work requires that entropy not increase, and that obtaining information requires work. This allows us to define a dynamical system in information theoretic terms.

Consider what individuates a system. If it is not just a nominal system, then it is individuated by causal connections within the system that bind it together.

Collier (1988) introduced the notion of cohesion to refer to the closure of the causal connections within a system that unify it and separate it from other systems. Collier and Hooker (1999) have refined the idea to a cohesion profile, which is a multidimensional probabilistic description of the unity dynamical conditions. The basic requirement for dynamic individuality is that the cohesion profile of the system is stronger than any cohesion profile that can be constructed involving other components. Thus cohesion both unifies a system and distinguishes it from other systems, providing the individuation conditions for dynamical systems. The information in the cohesion of a system cannot be completely localized, since any system is spread over space and time. In simple systems, for example a rock crystal, the bonds are local, and the information will be highly redundant. In an ideal gas in a container, all of the information of cohesion of the system is given by the macroscopic thermodynamic variables and the information of the cohesion of the container. Most systems are someplace in between. Highly organized complex systems will show information at a high level of redundancy, that is, it requires large sequences to detect the redundancy. Bennett (1985) has suggested that organization can be measured by the time (number of steps) it takes to compute the surface structure of a string from its compressed form. One of the consequences of this idea is that organization so defined will show high order redundancy. In any case, complex organized systems will not have maximal information (they won't be random), and they won't have minimal information (they won't be highly redundant). We can also expect that they will take time to produce, at least in the initial instance (reproduction from a template can be done more quickly). Also, they require effort to produce the information, which will be relatively high, whether in the initial case or from a template. Quick organization will be inefficient, requiring considerable power, much of which is likely to be dissipated in the process. On the other hand, spontaneous self-organization of complexly organized systems is a slow process, but can be much more efficient from an energetic point of view. The formation of such systems often involves a combination of symmetry breaking to produce complexity and entrainment to produce order. Even in manufacturing processes, raw materials are usually purified and/or cut into pieces and then reassembled. In spontaneous cases, like the formation of Bénard convection cells, symmetry breaking and entrainment can occur together. Generally, however, complexly organized systems will have a long iterative history of such processes, as well as sorting by selection. This is all rather abstract; the details can be found in the references of (Collier 1999). We turn now to the various kinds of entrainment.

3. VARIETIES OF ENTRAINMENT

Cohesion requires entrainment, but entrainment does not imply cohesion: two independent systems can be entrained, but the connection may not be strong enough to create cohesion; connections to other systems may be stronger. In many cases, however, entrainment and cohesion go together, as in a jazz combo playing a specific piece of music. One might imagine that rhythms from external sources are picked up and developed in the piece, but they would not thereby become part of the piece of music. When entrainment does become strong enough to produce cohesion, a new level is formed; we can talk of the emergence of new properties. Without cohesion, we have interacting parts, but no new level.

A taxonomy of rhythmic entrainment starts with the split between forced and spontaneous entrainment mentioned in the introduction. Forced entrainment, sometimes called driven, can be either high or low power. For example, a the movement of boat on a strong sea is driven by high power, and the boat is at the mercy of the sea. A typical low power system is one in which the driving force is applied in small amounts near to nodes of oscillation of the system, as when a child "pumps" a swing to make it move in larger, more energetic arcs. Many processes, like driving a car, combine both high and low power entrainment: the motive force is high power, but it is directed by relatively low power movements of a steering wheel. Control systems in general are low power, but can control large energy flows. To some extent, control is most easily thought of as an informational process, but the distinction is rather arbitrary. Forced entrainment always transfers preexisting information, either through reorganization or through a template. It does not create new information types, but at most new instances of preexisting types. Forced entrainment is especially important for discussion of measurement and perception, but it is also useful as a contrast with spontaneous entrainment. There is no reason, though, why both forced and spontaneous entrainment cannot occur in the same process, as probably happens in the development of organisms and other biological systems (for three quite different accounts, compare Kauffman 1993, with Brooks and Wiley 1988, Brooks et al 1989, and Collier et al in review, and with Weber et al 1989 and Schneider and Kay 1994).

Spontaneous entrainment always involves dissipation of energy. A simple example, is when a bunch of lipids spontaneously form a sphere because one end is polarized; the energy lost in forming this configuration is most likely expelled as heat, but whatever, the entropy of the system and its surroundings will increase. The same thing happens when ice melts, with the difference between the frozen water and its liquid state known as the latent heat of fusion. There are much more complex cases in nature, however, which in many cases have involved both spontaneous self-organization and selection, as in organisms, species and ecologies (for a broader account, see Collier and Siegel-Causey, in press). Notice that when the lipids form a sphere, symmetry is formed, and a new level, that of the lipid sphere is formed. This case can be analyzed almost entirely mechanically, but many case that are not much more complex cannot, such as the onset of convection in Bénard cells. The analysis of this transition *assumes* the convection, and equates the equations of motion for the convecting and conducting cases to determine the conditions at convection onset. A derivation of the convecting state from first principles of molecular motion is almost certainly impossible because the motion is chaotic. We can predict convection because we have observed it before, and the Bénard cell case is carefully controlled to have only one end state or attractor. In unobserved cases prediction is more difficult, and in cases with many attractors prediction is impossible in principle, and this makes it uncontrollable except in general characteristics. This uncontrollability is characteristic of complex self-organizing systems such as ecologies, societies, and economic systems. A major practical problem is what we can do about these circumstances.

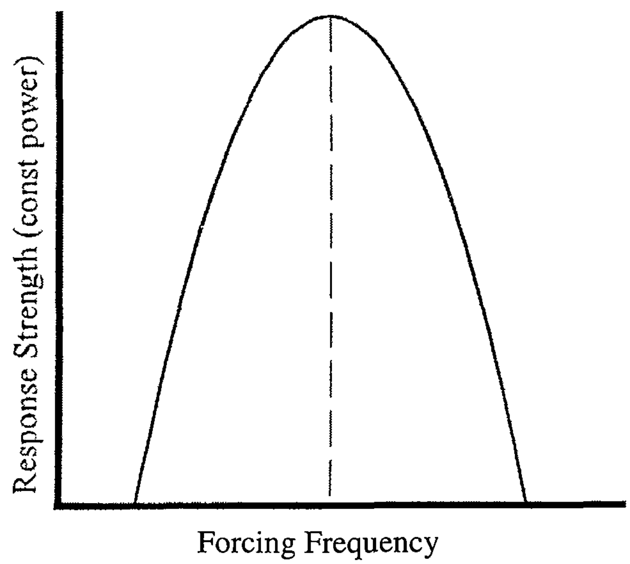


Figure 1: Forced Harmonic Oscillator

4. PHYSICS

4.1 Mechanical systems

Some of the most interesting spontaneous symmetries are in the Solar System. One obvious one is the 1:1 correspondence between the Moon's rotation and revolution times. This is typically attributed to tidal torques lagging behind the direct line between the Moon and the Earth that create dissipation of gravitational energy. The torques are minimized if there is a 1:1 correspondence, since then there is no lag. There are some other much more complex resonances, however. Mercury's rotation period is in a 3:2 relation to its revolution period around the Sun; i.e., Mercury turns three times on its axis for each time it goes around the Sun. This can also be accounted for in terms of tidal torques, since they are less in a 3:2 resonance than in any nearby relation (though greater than they would be for a 1:1 relation, in which there would be no tidal torque). Therefore, Mercury is effectively caught in the 3:2 resonance. How did it arrive in this resonance, rather than some other? Basically, the answer is chance. In the total Sun-Mercury phase space with tidal dissipation there are a number of attractors representing resonances, with fractal boundaries between the resonance basins. With no other information, there is about a 30% of 3:2 resonance, 50% for 1:1, and the other rest cover the other probabilities. Further resonances are a near 5:2 between the rotation period of Venus relative to its passing the Earth, and a resonance between Pluto and Neptune such that although Pluto crosses the orbit of Neptune, it will never hit Neptune (this may be partially a relic of Pluto having been a moon of Neptune at one time).

Physics also provides good examples of forced resonances, as when a force drives an oscillator towards its natural oscillatory frequency (nearby frequencies are damped by dissipation, so not any frequency can be driven unless there is considerable power and lots of energy to waste). Even chaotic oscillators can be put into resonance; the resonance itself is harmonic, but the motion of each oscillator is chaotic. This shows the possibility of forcing a resonance in an otherwise unpredictable system. The expense is dissipated power.

There are many other cases of spontaneous entrainment in physics. One especially simple case is the formation of dissipative structures through the promotion of noise. Bénard convection cells are the simplest of these, since they are so closely controlled. Nonetheless, the cells form spontaneously when the conditions are right. Other forms of entrainment are seen in the formation of eddies, standing waves in streams and waves in the atmosphere. In each case the entrainment creates a macroscopic structure that contains symmetries not present in the original microscopic structure. Although it is possible to create circumstances that will produce a certain resonance, in systems with multiple attractors this can be done only probabilistically, undermining controllability (large applications of power and dissipation of energy are required to overcome this lack of uncontrollability). In systems like the climate and weather, with many attractors, control is virtually impossible, since very small effects can move the system from one attractor to another (the so called "butterfly effect").

Several lessons can be learned from the physical cases. First, for the forced case, there are natural resonances in certain systems that can be driven by forces that contain the relevant frequency. The oscillator will resonate at its natural frequency because other frequencies will be damped by dissipative forces. To drive a system at an unnatural resonance requires a great deal of power, and wastes a lot of energy to overcome dissipation. We believe that the same principles apply across all systems, including social systems for forced resonance. In the case of spontaneous resonance, the properties of the system imply attractors to which the system can be led by dissipation. Systems with multiple attractors are hard to control, and like forced oscillation, require considerable power to drive them to a desired attractor, or else very subtle applications of force in regions near the chaotic zone between attractors. There is an interesting case of Japanese satellite that was supposed to go to the moon, but lacked the power due to other problems. NASA, who launched the satellite, worked out that there was a chaotic region in the earth-moon-sun system, and by applying a small amount of force near that chaotic region, transferred the satellite into a lunar orbit from a terrestrial one. Of course the journey took longer than just blasting the satellite to the moon, but it achieved the purpose.

4.2 Chemical Systems

We have already mentioned the formation of lipid spheres. Many other self-organizing chemical processes are similar, but depend on various thermodynamic parameters. Some specific chemicals have interesting properties from the perspective of symmetry and cohesion. Benzene, for example, is a closed loop of six carbon atoms with double bonds that oscillate. This spreads the cohesion of the molecule over the whole bond structure, and increases cohesion and hence stability.

Another case involves the comparison between ethylene and butadiene (Harris and Bertolucci 1988: 288-297). Ethylene is a double-bonded that 2 carbon unit. Butadiene is a 4 carbon unit with 2 double bonds. To make all things equal, the energy of 2 molecules of ethylene is compared with one molecule of butadiene. The butadiene is more stable by 12kJ. The usual explanation is that the bond energy is delocalized, but it is not clear why delocalizing something should lower its energy. Our explanation, that there is increased cohesion in the form of harmonic entrainment of the bonds explains why the energy of butadiene is lower.

Presumably more complex chemical systems as found in organisms increase stability through similar delocalization (although not through double bounds, but networks of pathways, even though the individual molecules are not necessarily especially stable). The thing to note is that it is the whole network that inherits the stability. We could go on about similar issues in development, evolution and ecology, but these have been studied extensively elsewhere (ecological studies by Robert May introduced much of the interest in the general topic. Instead we now turn to measurement as another example of entrainment.

5. MEASUREMENT

Measurement is a special form of entrainment in which the measured property drives a special device into a resonance that correlates in a theoretically predictable way with the driving property. We give a couple of examples from geophysics, and then argue that sensation is the most fundamental form of measurement, working on the same principles.

Micro-gravity surveys are often done with a gravimeter, a device that responds to the local force of gravity by moving a delicate spring holding a weight. Absolute measurements are impossible with this device, but the relative force a gravity at different locations can measured very accurately. Basically, the gravitational force twists the spring, the twisting being proportionate to the change in gravitational field strength. This is an example of directly forced entrainment. Absolute measurements are impossible with this device, but the relative force of gravity at different locations can be determined. Magnetic surveys, on the other hand, are done with a device that has an ability to resonate at various frequencies when driven by a force. The magnetic field, much stronger than the energies required to drive the meter into resonance, drives the meter into one of its many possible resonances, designed to be very close together to permit a high degree of accuracy in measurement. In this case, the forcing is indirect, since the meter has its own resonances, one of which is selected by the force of the magnetic field. The two instruments use different principles for measurement, one directly forced, and the other forced indirectly through oscillations.

Sensing is similar to measurement with geophysical instruments, but has some additional interesting properties, both in operation and origin. In operation, energy is passed from the thing sensed that has a form that stimulates the nervous system so as to entrain the form in a way that we can use to form expectations, predictions, and guide actions (though this process is fallible, partly because of possible failings in the sensory system, but more likely at other points in the process.) Possibly, most sensation is more like the gravity meter, directly forced, but in hearing and smell the sensory system has a number of natural resonances, and the forcing is indirect. It has been suggested that both smell and hearing systems are kept normally in a chaotic state, and that the sensation drives the system into one of the infinity of oscillatory states that make up the chaotic state. Although this mechanism is not universally accepted, it would allow very fast sensation, because all the sensory states are already present in the sensory apparatus. A sensory system near to chaos would be almost as effective in response, but would allow only a predetermined set of responses that might be slightly different than the driving sensory impulse. Possibly more subtle phenomena are grasped in the same way, even across sensory modes. In shape and color sensation it is known that there are specific receptors in the eye and brain, but the way these are combined into perceptual images is unknown. Many philosophers (e.g., Harmon 1973) believe judgment is involved in perception. This is surely true at some stage, and if judgement is interpreted loosely enough, perhaps even at an early stage. It is possible, however, that self-organization plays an important role in forming gestalts. These have been largely ignored by contemporary cognitive scientists. Recognized gestalts, such as the examples in textbooks, are probably forced, but new gestalts appear at a new level of organization, and are good candidates for self-organization; they show new order. This process and subsequent selection may be important in creative thought.

Sensory systems, as we have described them, already contain all the possible responses, but they had to originate in some way. This origin was not designed. The most likely explanation is that the underlying structure was malleable, and gradually responded spontaneously to sensory inputs, whose increasing effectiveness was guided by interaction with the environment and by selection. These guiding processes are a sort of forcing that tune the sensory system as it evolves, but the original formation of sensory attractors must have been spontaneous, with information dissipative in the interactive and reproductive and selection processes. Similar mechanisms might be involved in the formation of learned higher order perceptions as well as learned ideas and practices.

One of the more vexing problems of measurement of our day is what happens in Quantum Mechanical system when they interact with measurement (macroscopic devices). In this case the energies involved are very subtle, and we cannot rely on forcing. We speculate a Bohmian style approach in which the guiding wave entrains the measurement device. Since the energy in the wave can be extracted only a half wave length at a time, there are natural limits on this sort of measurement, relative to the size of the properties to be measured. Conjugate values, like momentum and position, which together make up action, the units of the Planck's constant, cannot be measured entirely independently, since they are entangled in the single wave, and can only be entrained together. This hardly explains the mysteries of Quantum Mechanics, but it does explain the measurement of Quantum systems. The phase information is not available macroscopically (it has no effect on energy), so it cannot be measured. Given that macroscopic values cannot influence phase, the quantum of action places a lower limit on what we can measure. This will be true of any measurement that involves measuring action.

6. MEMES

The communication of memes is one of the most interesting forms of entrainment. Memes can be ideas, practices, ideologies and paradigms, among other things, though they are most often thought of as ideas. In some cases, when the basic primitives are already there, memes are passed by simple resonance, causing an appropriate combination of preexisting memes. In other cases, new primitives must be created, as when an apprentice learns from a master. This involves some simple forcing through recombination, but largely involves the spontaneous generation of memes through the generation of new primitives in the apprentice in the presence of the master's memes, which aid in the entrainment through reward and punishment, but also through copying and practice directed in an appropriate way. This latter form of learning can be carried out independently, and is probably a major factor in the transmission of memes. The basic primitives are already available to the apprentice, and he can reorganize them for new tasks, but mastery comes only when the memes are integrated into the autonomy of the apprentice, so the apprentice can discover new ways to work, and achieve mastery. The passage from childhood to adulthood is not dissimilar. The new organization requires both differentiation and entrainment, and requires much experience and practice.

The original metaphor for entrainment from music is a social case, and we believe that entrainment is common at the social level. This can help to create social order and function, but it can also be wasteful and counterproductive if done poorly. In some cases practices, ideas and ideologies form spontaneously and unpredictably when the right conditions occur, resulting in very rapid change. Fashion and art change, and the fall of the Berlin Wall probably contain large elements of this sort of entrainment.

The population affected need not be prepared for the eventuality, but there must exist a social attractor (perhaps one of many), that random variations allow to become expressed throughout the population. If there are many attractors, the change will be essentially unpredictable, and in that sense random. This suggests that there will never be a fully predictive social science, especially in the case of history.

Authoritarian and totalitarian states put a premium on control. To some extent, they must rely on predispositions in their populations, but largely they rule by terrorist methods and fear. This requires a large concentration of both political, economic and political power, since driving a system artificially a requires a lot of power and waste of energy. This suggests that such states will be unstable. Unfortunately, whenever there are large concentrations of political, economic or physical power, there will be a tendency to use forced entrainment of ideology, however inefficient. A more efficient but less reliable method is propaganda and advertising, which attempt to drive or create resonances through subtle forcing. This method still requires a concentration of power to exclude competitors.

We believe that a stable social system is best founded on spontaneous entrainment. This is both more stable and more efficient than forced coordination and obedience. Its main problem is that it may lead to arbitrary and unproductive entrainments, basically pathological, so some control is mandatory except in the most advanced social systems, in which stability is already well entrained, and mechanisms for the dissipation of concentrations of power already entrenched in the structure of the system. Great variety can be tolerated in such a system, with minimal control, and it allows both the greatest freedom and flexibility.

7. CONCLUSION

The mechanical model persisting since Newton's time suggests the forced model of entrainment. We have offered an alternative self-organizing model that can explain many phenomena, and even has social and economic repercussions. It explains why authoritarian systems need to use a lot of power (making them inherently unstable), and why a self-organizing system, perhaps with gentle control, needs less power and is more stable and self-sustaining. The lessons are from physics and biology, but the extend to systems in general, whether management, social or economic. These higher level systems show an organization that makes them cohere, and follow their own rules ensuring their emergence (Collier and Muller 1999).

REFERENCES

Banaschewski, B (1977) On G Spencer Brown's Laws of Form, *Notre Dame Journal of Formal Logic* 18: 507-509.

Bennett, C.H (1982) The Thermodynamics of Computation: A Review, *International Review of Theoretical Physics* 21: 905-940, reprinted in Lef and Rex, eds. *Maxwell's Demon*.

Bennett, C.H. (1985) Dissipation, Information, Computational Complexity and the Definition of Organization, in D. Pines, ed *Emerging Syntheses in Science. Proceedings of the Founding Workshops of the Santa Fe Institute*: 297-313

Bennett, C.H. (1987) Demons, Engines and the Second Law, *Scientific American* 257, No 5: 108-116.

Brooks, D.R. & Wiley, E.O. (1988) *Evolution as Entropy: Toward a Unified Theory of Biology*, 2nd edition, Chicago University Press.

Chaitin, G.J. (1975) Randomness and Mathematical Proof, *Scientific American* 232, No. 5 (May): 47-52.

Collier, J.D. (1996) Information Originates in Symmetry Breaking, *Symmetry: Culture and Science* 7, 247-256.

Collier, J.D. (1998) Information Increase in Biological Systems: How Does Adaptation Fit?, in Vijver, Gertrudis van der, Salthe, S.N. and Delpos, M., eds. *Evolutionary Systems,* Dordrecht: Kluwer.

Collier, J.D. (1999) Causation is the Transfer of Information, in Sankey, H., ed. *Causation, Natural Laws and Explanation* 279-321, Dordrecht: Kluwer

Collier, J.D. (1999) Autonomy in Anticipatory Systems: Significance for Functionality, Intentionality and Meaning, in Dubois, D.M., ed. *Proceedings of CASYS'98, The Second International Conference on Computing Anticipatory Systems*, New York: Springer-Verlag

Collier, J.D., Banerjee, S. and Dyck, L. (in review) A Non-Equlibrium Perspective Linking Development and Evolution, in Collier, J.D. and Siegel-Causey, D., eds. *Between Order and Chaos,* Dordrecht: Kluwer.

Collier, J.D and Siegel-Causey, D., eds. (in review) *Between Order and Chaos. Studies in Non-Equilibrium Biology,* Dordrecht: Kluwer

Collier, J D. and Hooker, C.A. (1999) Complexly organised dynamical systems, *Open Systems and Information Dynamics*.

Collier, J.D. and Muller, S (1998) The Dynamical Basis of Emergence in Natural Hierarchies, in Farre, G. and Oksala, T., eds. *Emergence, Complexity, Hierarchy and Organization, Selected and Edited Papers from the ECHO III Conference, Acta Polytechnica Scandinavica, MA91,* Espoo: Finish Academy of Technology.

Cull, P. and Frank, W. (1984) Flaws of Form, *International Journal of General Systems* 5: 201-211.

Harmon, G. (1973) *Thought,* Princeton: Princeon University Press.

Kauffman, S.A. (1991) Antichaos and Adaptation, *Scientific American*, August: pp 64-70.

Kauffman, S.A. (1993) *The Origins of Order: Self-organization and Selection in Evolution,* New York: Oxford University Press.

Kolmogorov, A.N. (1965) Three Approaches to the Quantitative Definition of Information, *Problems of Information Transmission* 1: 1-7.

Kolmogorov, A N (1968) Logical Basis for Information Theory and Probability Theory, *IEEE Transactions on Information Theory* 14: 662-664.

Landauer, R. (1961) Irreversibility and Heat Generation in the Computing Process, *IBM J. Res. Dev.* 5: 183-191, reprinted in Lef and Rex, eds. *Maxwell's Demon*.

Landauer, R. (1987) Computation: A Fundamental Physical View, *Phys Scr.* 35: 88-95, reprinted in Lef and Rex, eds. *Maxwell's Demon*.

Lef, H.S. and Rex, A.F. (1990) *Maxwell's Demon: Entropy, Information, Computing*, Princeton: Princeton University Press.

Schneider, E D and Kay, J.J. (1994) Life as a Manifestation of the Second Law of Thermodynamics, *Mathl. Comput. Modeling*, 19, No. 6-8: 25-48.

Solomonoff, R.J. (1964) A Formal Theory of Inductive Inference, *Information and Control* 7: 1-22.

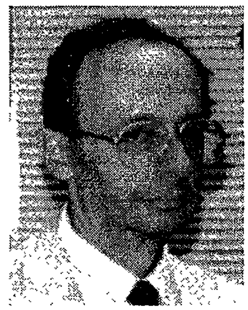
Spencer-Brown, G. (1969) *Laws of Form*, London: Allen & Unwin.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 179-187, 1998

LAWS OF SYMMETRY BREAKING

Closing plenary lecture

György Darvas



Physicist, philosopher (b. Budapest, Hungary, 1948).

Address: Symmetrion and Institute for Research Organisation of the Hungarian Academy of Sciences, 18 Nádor St , Budapest, H-1051 Hungary; Mailing address: P.O. Box 994, Budapest, H-1245 Hungary; E-mail: darvasg@helka iif.hu .

Fields of interest: history and philosophy of science, theoretical physics, symmetry in science and arts.

Publications: (with Agazzi, E., eds.), *Philosophy of Mathematics Today*, Episteme 22 (series ed., M. Bunge), Dordrecht/Boston/London: Kluwer Academic Publishers, 1997, xxix + 361 p.; (with Nagy, D., eds.), *Symmetry of Structure*, Extended Abstracts, Budapest: International Society for the Interdisciplinary Study of Symmetry, Vols. I and II, 1989, 656 p; (with Nagy, D., Pardavi-Horvath, M., eds.), *Symmetry: Natural and Artificial*, Extended Abstracts, Budapest: Symmetrion, 1995, 734 p.; (ed.), *Ars (Dis)Symmetrica '99; Science Reflected in the Arts - Art in the Sciences*, (Satellite art program series accompanying the UNESCO-ICSU World Conference on Science, Science for the Twenty-First Century: A New Commitment, 1999), Institute for Research Organisation, Hungarian Academy of Sciences and Symmetrion, Budapest, 1999, 144 p ; Dichotomies and trichotomies in the eastern and western cultures, pp. 91-99, In: Ogawa, T., Miura, K., Masunari, T., and Nagy, D., eds, *Katachi U Symmetry*, Tokyo: Springer-Verlag, 417 p.

Introduction

We came together at fourth time now. We are representing different fields of study, different kinds of the arts, different cultures. One could raise the question, what brings so many different scientists and artists together again and again? The answer seems very simple: it is the common interest in *symmetry*.

What is that? Well, one can quote definitions of symmetry, asymmetry, dissymmetry, etc. what *denote common concepts* and *common phenomena* in our everyday work, and form a *common platform for our communication* where we understand each other, and where we can *reveal useful information* to each other.

Nevertheless, the case is not so simple. Symmetry does not mean a branch of science or arts. This word denotes an interdisciplinary concept, a class of properties and phenomena. Its essence is in its interdisciplinary character. Symmetry and asymmetry, order and disorder appear in all of our fields of activity. So, the information we can convey, bears *practical value* for each other. Therefore, although there is no such discipline like "symmetrology", but there are common features, and we acknowledge mutually accepted values.

Thus, when we answered the first question, (what brings us together?), we found the *glue* (that was symmetry), but did not answer the question, *what is kept together by this glue*? Semantically, there can be two subjects of this last question. In epistemological sense, one of them are we, ourselves. The other is, in ontological sense, the "common object" of our studies.

Now the next question is, whether there exist common objects of studies, what do not form a discipline? In other words, one can raise the question: are there common principles or common laws what are guiding the symmetry studies conducted in all disciplines? Does this glue keep together (at least symbolically) an *ordered structure*, in which all of us have our quiet space, where we can maintain the privacy of our creative activity, and without which space the whole structure should collapse?

My answer is definite: yes.
Yes, there are general enough *common objects* of our studies (at least in philosophical terms).
Yes, there are *laws* (and not only principles) of symmetry/asymmetry.
Yes, these laws are *common*, irrespective of the discipline, where we are active.
Yes, there are *general* laws, what are glued by the phenomena of symmetry (to be more precise, of symmetry breaking).
Earlier or later any field of learning formulates its laws (this belongs to its identity). So do we now.

These are demonstrated in my paper.

Maybe the question seems too theoretical for some of you, it cannot be quite neutral for any of us after nearly a decade of work together: whether there are general laws guiding our activities; whether there can be formulated what is common in each other's work; how can we determine what is kept together by this glue called symmetry, asymmetry; what is exactly, which brings our achievements close together? I am convinced that the introduction of some general features in the field of symmetry studies makes our co-operation more conscious. The recent decades demonstrated that methods borrowed by the means of symmetry considerations from different disciplines and arts, led to new discoveries (cf., e.g., the quasicrystals, the fullerenes). What has worked on the level of intuition, may become part of the co-ordinated activity of scientists and artists in the future.

Philosophical background

Nature proved to be not always symmetric. Although basic laws could be formulated by the application of symmetry principles, most *new phenomena* appeared by certain distortion of symmetry. Therefore, it was *symmetry breaking* what led scientists to new discoveries. This is why P. Curie stated "dissymmetry makes the phenomenon".

Symmetry plays several different roles in philosophical thought.

In *epistemology* it plays a heuristic role, since the mind often prefers symmetric solutions of problems from among alternatives. Not only prefers, even seeks for such solutions if there are available any such. Thus, symmetry performs a methodological function in the formulation of scientific knowledge. Many examples can be quoted when great discoverers' minds were led by symmetry principles.

The *ontological* basis of the importance of (dis)symmetry is that material reality indeed has both symmetry properties and symmetry breakings. It has been less realised that symmetry breaking plays important role in the *construction of the material world.* There is an order (of symmetry breakings), what can be traced along the evolution. These material properties and order should be reflected in the *laws of the nature.* They are the *laws of symmetry* and *symmetry breaking.*

Science and mathematical description looked for *order* and for the *linear phenomena* for many centuries, because those were so perfect and beautiful. Less attention was paid to *chaos, disordered structures* and *nonlinear phenomena.* The recent two decades turned the attention of scientists to the systematic description (and discovery of laws) of the latter.

Similar is the situation with the *phenomena of symmetry breaking.* We knew and phenomenologically described them. We knew, that they were present in any phenomenon of nature, and most new promising areas for a scientist could be identified by the study of dissymmetric phenomena. However, most works dealing with symmetry itself discussed symmetry and treated laws of symmetry (e.g., Rosen, 1995; van Fraassen, 1989; de Gortari, 1970). Many new transdisciplinary discoveries were made on the basis of the application of symmetry considerations (cf., the discovery of quasicrystals, fullerenes, in the recent two decades). Dissymmetry, and so symmetry breaking was left for the so called "puzzle solving" research (using this term after T. Kuhn, 1963). Making an order in symmetry breaking was a subject only within separated disciplines, like in particle physics and cosmology, as well as in some biological subdisciplines.

Now, we make an attempt to discuss the role of symmetry breaking at a philosophical level. *Laws of symmetry breaking* in a wide context will be formulated first.

The laws of ontological levels and symmetry breaking

(1) The law on the determining role of the lower levels.

(1a) Among two consecutive ($_{\text{a lower}}$ and $^{\text{an upper}}$) levels, the lower level potentially (but only potentially) possesses the characteristic type of interaction of the consecutive upper level; i.e., that the preceding lower level's types of interaction play the determining role in the development and existence of any level's characteristic interaction. However,

(1b) In the interrelation of two different ($^{\text{an upper}}$ and $_{\text{a lower}}$) levels, generally the upper level's structure affects actively the other, since

(1c) Any lower level material structure can reflect its environment only on its own (lower) quality and own level. Within that, the material structure of a lower level can reflect the material structures corresponding to the upper level's forms of material motion also only on its own (lower) level.

(The two statements in (1c) are not-certainly equivalent, because the given levels are determined *per definitionem* by their characteristic interaction and not by the corresponding form of material motion.) For example, any inanimate being can reflect an animal only as a physical object, and cannot reflect its biological properties; no animal can discern the social differences between human beings.

Since the relation of the two (lower and upper) levels are not symmetric, this law does not open the door to any reductionism. A reductionist approach would allow only the following kind of statement, viz., "among two consecutive levels, the lower level possesses the characteristic type of interaction of the consecutive upper level." But, according to our laws, (1a) limits the existence of the upper level's characteristic interaction at the lower levels to potentiality, while (1b) and (1c) together contradict any statement which denies the appearance of new qualities at the upper levels.

(2) The law of correspondence between the ontological levels and their potential symmetry properties.

(*2a*) Each qualitatively higher organisational form in the evolution of matter is marked by the loss of a certain symmetry property, and

(*2b*) Each loss of a potential symmetry property of matter traces a new material quality.

Consequently, the precondition of the development (in its relative totality) of a qualitatively new (material) level is the breaking of a certain symmetry (property), and at the same time, the condition of the continuance (existence) of the new level is to possess (new?) conserved properties. Therefore

(*2c*) Parallel with the appearance of new material qualities and new (higher) ontological levels, there appear also new symmetries.

(*2d*) These new symmetries qualitatively differ from those what existed at the previous (lower) levels and what have been broken at the given level. These new symmetries involve new conserved properties.

As a conclusion, the lower and higher ontological levels can be traced by a sequence of symmetry breakings, thus it can be formulated, that

(3) Each symmetry breaking leads to a higher organisational level of matter.

(4) Each higher organisational level of matter is - in a certain sense - less stable than the former one.

The latter statement needs some further explanation. This will be given in a detailed treatise by encountering examples for all the above four laws. Let's now mention only the decreasing self-reproducibility of the living organisms along phylogeny, or the decreasing forces keeping together the inanimate structures from the subatomic particles to the large molecules.

Some open problems

There are some further problems what arise in the interpretation of the above laws. Two of them are crucial. The first one concerns the so called level theories, what are formulated in the framework of philosophy, and what distinguish fundamental and particular levels. Without going into philosophical details, we mention, that generally there are distinguished three fundamental levels of the material world: inanimate nature, the organic world, and human society and thought, while there are further, particular levels *within* the respective fundamental levels. This is the point, where the next important group of concerns appear: what are the *differentia specifica* of a given level? What is the main concept, according to which one distinguishes the different levels? There are many candidate concepts for this role, e.g., interactions, forms of motion, order of magnitude relations (principle of 'nest of tables'), sequence of genetic evolution, degree of complexity, types of matter, space-time forms, structures. I choose from among them the characteristic *interactions*, because this works, and plays an equivalent role, both at the fundamental and non-fundamental levels. In short, one can speak about two types of level theories: a general one (in philosophy) and particular ones (in inanimate, the organic nature, and in human society). Particular level theories differ from each other *in the* three fundamental ontological spheres, nevertheless in their description and contents. At the same time they may have common features, e.g., all are particular theories concerning their width of validity, and all are based on an arrangement by a common concept, namely the *forms of interaction*. The clarification of these conceptual problems is necessary to understand the laws of symmetry breaking. One can put the question: Is unification of the different types of level theories possible? With certain limits, yes. For this reason, one should accept that all levels can be characterised by a given type of interaction, and they are submitted to the laws (1a-c). Any further detail belongs to the competency of the discipline studying the phenomena of the given level.

The detailed treatise of the mentioned two crucial problems are given in my philosophical works. I do not want to bore too long our interdisciplinary audience with special philosophical problems.

Levels and symmetry breakings

Is there a one-to-one correspondence between the *levels* and *symmetry properties broken* at the given level? One cannot give a definite answer yet, since it has not been studied thoroughly in all disciplines. However, all the available examples affirm the presumption. E.g., the stronger a basic physical interaction is, the more quantities are conserved, and with weakening the type of interaction, the number of symmetry breaking increases. That means also, that the weaker an interaction is, the greater number of material structures (particles) are affected by it, and their interactions are limited by fewer conservation laws. Strong interaction conserves all elementary particle quantities. In electromagnetic interaction Isospin is not conserved, but all the others are; in weak interaction Parity, Charge conjugation and others are not conserved, (however the combination of them with Time reversal (CPT) is conserved). Parity conservation is also violated in the so called united electroweak interaction. The antineutrinos play an important role in the electroweak interaction. These particles exist only in a right-handed chiral form. Antineutrinos are produced during beta decay, where the majority of the electrons produced simultaneously with the antineutrinos have a left-handed chirality (spin) (Ne'eman, 1986). The participants of the electroweak interaction are the electrons of the atom on the one side and the protons and neutrons of the nucleus on the other. From the chirality of the participants follows the chirality of the atoms, and the molecules built of them. This leads to the existence of the enantiomers in the organic molecules (e.g., glucose and fructose), then the L- and D-aminoacids. Proteins are built up (almost) exclusively from L-aminoacids, and therefore it is not by chance that RNA and DNA form only right-handed helixes. Is it surprising that living creations are chiral? All this follows from the electroweak interaction what distinguishes 'left' and 'right' by the charged weak currents and neutral weak currents (or in other words by W and Z forces) (Hegstrom and Kondepudi, 1990).

However, nature is not so simple. Nature reproduces the dominance of left- or right-handedness at any new level by new properties. While left-handed DNA helixes are very rare, we find both left- and right-handed helixes among bacteria, plants, snails, etc. Nature produces again both kinds, although, by a spontaneous symmetry breaking their numbers are different.

The dominance of morphological asymmetry is becoming prevailing in the morphology at the more evolved animals (e.g., circulation system). Another symmetry: irreversibility (e.g., reproducibility of the organs) weakens during the evolution too. Nevertheless, the brain remains symmetric, even at mammals. A new, qualitative change (mutation) takes place, when the lateralisation of the brain starts. This makes possible the real right- and left-handedness, differentiation of the kinetic and the speech centres in the brain, and the separation of the emotional and rational, etc. functions. The loss of the symmetry of the brain is also a typical example of the violation of a symmetry, which did not exist 'always', only since one can speak of 'brain' or 'neural system' as a quality, as an organ of living organisms (2c-d).

Closing remarks

This treatment could not give a detailed philosophical analysis of the full problem. That was not the aim of this paper. Our goal was only to introduce the *laws of symmetry breaking*. We did not want to replace any evolution theory; these laws touch them only tangentially. It is important to note that none of the treated laws take stand on the debate of reductionism. We stressed, that they may be used as arguments by both parties - probably they can bring the parties closer to a decision - but in their presented form they do not fulfil a decisive function. This was also not the aim of this paper. The new features of this treatment were to link the level theories with the laws of symmetry breaking. Finally, what is most important at this forum: we have demonstrated that one can find common ground, based by its own specific laws, for hunters for symmetric/asymmetric phenomena in different disciplines.

REFERENCES

Bérczi, S., Lukács, B., guest eds. (1993) *Symmetry and Topology in Evolution*, Special issues of *Symmetry: Culture and Science*, 4, 1-2, 1-224.

Collier, J., Burch M. (1998) Order from rhythmic entertainment and the origin of levels through dissipation, *Symmetry: Culture and Science*, 9, 2-4 (this issue).

Darvas, G. (1987) Level theories in philosophy and physics - Exploring and interpreting the structure of matter, Vol. 5, Part 3, pp 76-79, In: *8th International Congress of Logic, Methodology, and Philosophy of Science*, Abstracts, Moscow: Nauka.

Darvas, G. (1995) ISIS-Symmetry. Looking for truth and beauty, pp. 9-18, In. Darvas, G., Nagy, D., Pardavi, M., eds., *Symmetry: Natural and Artificial*, Budapest: Symmetrion, 734 p.

Darvas, G. (1997) Mathematical symmetry principles in the scientific world view, pp. 319-334, In: Agazzi, E., Darvas, G., eds. *Philosophy of Mathematics Today*, Dordrecht-Boston-London: Kluwer Academic Publ., 361 p.

Darvas, G. (1998) Ontological levels and symmetry breaking, *Proceedings of the XXth World Congress of Philosophy*, Boston, 10-16 August 1998, 10 pp.

Darvas, G., Nagy, D., eds. (1989) *Symmetry of Structure*, Vols 1-2, Budapest: ISIS-Symmetry, 656 p

van Fraassen, B. C (1989) *Laws and Symmetry*, Oxford: Clarendon Press, 395 p.

de Gortari, E. (1970) Symmetry as a scientific method, pp. 30-34, In: *La symetrie. Comme principe heuristique dans les differentes sciences*, Bruxelles: Office international de librairie, 135 pp.

Hahn, W. (1989) Symmetrie als Entwicklungsprinzip in Natur und Kunst, Königstein: Langewiesche, 320 p.

Hegstrom, R. A., Kondepudi, D. K., (1990) Is the Universe left- or right-handed? *Scientific American*, 3.

Heisenberg, W. (1971) *Physics and Beyond. Encounters and Conversations*, New York: Harper and Row Publ. Inc.

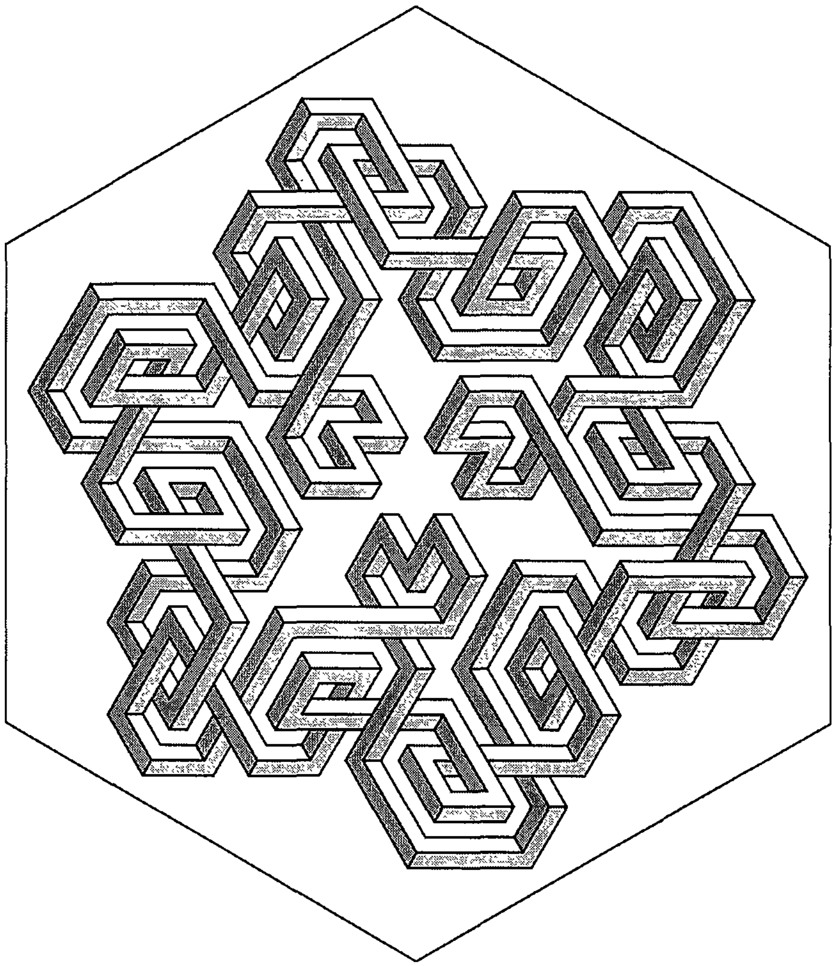
Kuhn, T. (1963) *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

Mainzer, K. (1990) Symmetry in philosophy and history of science, *Symmetry: Culture and Science*, 1, 3, 319-328.

Ne'eman, Y., Kirsh, Y. (1986) *The Particle Hunters*, Cambridge: Cambridge University Press, 272 p.

Rosen, J. (1995) *Symmetry in Science*, New York. Springer-Verlag, 213 pp.

Wigner, E. P. (1964) Symmetry and conservation laws, *Proceedings of the National Academy of Sciences*, 51, 5, 956-965.



Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 189-197, 1998

ON SYMMETRY IN SCHOOL MATHEMATICS

Tommy Dreyfus(1) and Theodore Eisenberg(2)

Address: (1) Center for Technological Education Holon, Israel E-mail: tommy.dreyfus@weizmann.ac.il.
(2) Ben-Gurion University, Beer-Sheva, Israel E-mail: eisen@black.bgu.ac.il.

THE UBIQUITY OF SYMMETRY

Everywhere we turn we can see symmetrical relationships. They are both visual and audio, and they are so pervasive in our daily lives that one is led naturally to wonder if the notion of symmetry is innate in human beings. It is almost as though the notion of symmetry is built into us as a standard against which we measure aesthetic appeal to assess both mental and physical constructs. Hargittai and Hargittai in their text *Symmetry: A Unifying Concept*, illustrate how deeply seated and ubiquitous symmetrical relationships are through hundreds of photographs of man-made objects, from examples in architectural symmetry, to those found in nature, as exemplified by the markings on the wings of a butterfly. The notion of symmetry is itself a mathematician's dream, for point and line symmetries which have been extensively studied in their own right, have been generalized and applied to almost every area of mathematics, even school mathematics. Moreover, entire domains of mathematics, such as group theory have arisen out of the study of symmetry.

Algebra, geometry, trigonometry and calculus are four main areas of school and collegiate mathematics. And in each of these domains students are introduced to generalizations on the notions of symmetry. In algebra for example, we introduce students to the notions of symmetric functions, symmetric determinants, symmetric groups, symmetric systems of equations and to symmetric forms, as in the symmetric form of the equation of a line; in geometry they meet the notions of point and line symmetry, and n-fold symmetry. In trigonometry the students meet basic symmetric relations like $\cos \alpha = \sin (90^\circ - \alpha)$ and at a more advanced level, the notions of symmetric spherical triangles. In elementary calculus they again see that symmetry plays a pivotal role in applying integration techniques and in working with differential forms. The centrality of symmetry as a notion in and of itself, not to mention its use as a heuristic in problem solving, is easily documented in general mathematics, and in school mathematics, too. But whether or not there is a natural, innate, gravitation towards symmetry is an open question, although many giants in mathematics and the physical sciences (Poincare 1913, Einstein 1935, Weyl 1952, Polya 1962 and Penrose 1974) have addressed their own propensities for symmetry and aesthetics, individually saying that they believe it to be one of the driving forces behind their work.

SYMMETRY AS AN AESTHETIC METRIC

It is well known that there seems to be a small set of real numbers which appeal to our psyche more than other numbers. E.g., more than a hundred years ago the psychologist Gustov Fechner made literally thousands of measurements of rectangles commonly seen in everyday life; playing cards, window frames, writing papers, book covers, etc., and he noticed that the ratio of the length to the width seemed to approach the golden ratio ($\tau = \frac{1+\sqrt{5}}{2}$) .He then presented rectangles to hundreds of individuals and asked them to choose one, with no selection criteria specified. He claims that they disproportionately chose rectangles, the ratio of whose sides was close to the golden ratio. He then questioned the innateness of this number in the psyche of man; a number which was already known to be ubiquitous in nature, science and mathematics. According to Huntley (1970), Fechner's experiments were repeated by Witman, by Lalo and again by Thorndike, each a leading psychologist of his time, and each obtained similar results. Huntley's listing of Fechner's data is presented in Figure 1.

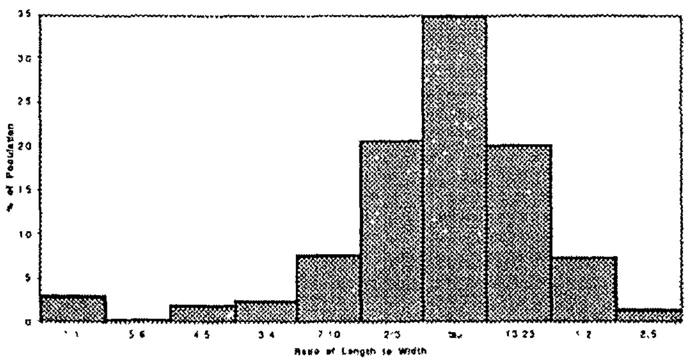


Figure 1

The golden ratio is a number closely tied to symmetry. The ratio is obtained as follows. Given the line segment *AB* there are two special points *C* and *D* on it that divide the line segment into the golden ratio. One of these points *C* is an interior point to the segment *AB*. The other point *D* is an exterior point to the segment; i.e., the segment *AB* must be extended to reach it.

Figure 2

The Golden Ratio τ, is defined as follows:

$$\tau = \frac{AC}{CB} = \frac{AB}{AC}; \qquad \frac{AD}{BD} = \frac{BD}{AB} = \tau$$

$$\tau = \frac{AC}{CB} = \frac{AC+CB}{AC}; \qquad \frac{AB+BD}{BD} = \frac{BD}{AB} = \tau$$

$$\tau = \frac{AC}{CB} = 1 + \frac{CB}{AC}; \qquad \frac{AB}{BD} + 1 = \frac{BD}{AB} = \tau$$

$$\tau = 1 + \frac{1}{\tau} \qquad \frac{1}{\tau} + 1 = \tau$$

$$\tau^2 - \tau - 1 = 0 \qquad \tau^2 - \tau - 1 = 0$$

$$\tau = \frac{1+\sqrt{5}}{2}$$

Constructing the points *C* and *D* for a given line segment *AB* is an instructive exercise which is closely associated with the Appolonian circle for points *A* and *B*; the Appolonian circle is the locus of a point *P* such that $\frac{AP}{PB}$ is a constant $k > 0$, $k \neq 1$. How do we find points *C* and *D*? With *B* as the center, draw a circle of arbitrary radius *r*; now with *A* as the center draw a circle of radius τr. If these two circles do not intersect, choose another value for the radius *r*, and repeat the procedure. Let the circles intersect at *P*. Note that $\frac{AP}{PB} = \tau$. Bisect $\angle APB$ internally and externally. Let the internal bisector meet line segment *AB* at *C;* and let the external bisector meet *AB* extended at *D*. We have now found the two desired points for the golden ratio. If the length of *AB* is one, then $AC = \frac{1}{\tau}$, $CB = \frac{1}{\tau^2}$ and $BD = \tau$.

The aesthetic appeal of the golden ratio and its ties to the Fibonacci sequence, as well as its far reaching connections to nature and science are well documented in the literature (Huntley 1970, Herz-Fischler 1998, Dunlap 1998). But the connection of this number to the human psyche, in the spirit of Fechner's work with investigating our attractions toward it, is an open question. Nevertheless, there are many aspects of symmetry which are embedded in the golden ratio and which are instructive for students to study. (Space limitation allows us to only mention a few.)

The two points C and D are said to divide the line segment AB into its *mean* and *extreme* ratio — and there is no other set of points which does this. I.e., there is no other set of points X and Y on segment AB such that $\frac{AX}{XB} = \frac{AB}{AX} = \tau = \frac{AY}{YB} = \frac{YB}{AB}$. Perhaps the ratio is not connected to our subconsciousness, but it and its reciprocal, $\left(\frac{\sqrt{5}-1}{2}\right)$, are somehow connected to our aesthetic attraction towards simplicity and minimalism.

Hidden symmetries in the golden ratio abound. For example, τ, the golden ratio, satisfies the equation $\tau^2 - \tau - 1 = 0$.

$$\tau^2 - \tau - 1 = 0 \qquad \tau^2 - \tau - 1 = 0$$
$$\tau^2 = \tau + 1 \qquad \tau^2 = 1 + \tau$$
$$\tau = 1 + \frac{1}{\tau} \qquad \tau = \sqrt{1+\tau}$$
$$\tau = 1 + \frac{1}{1+\frac{1}{\tau}} \qquad \tau = \sqrt{1+\sqrt{1+\tau}}$$
$$\tau = 1 + \frac{1}{1+\frac{1}{1+\frac{1}{\tau}}} \qquad \tau = \sqrt{1+\sqrt{1+\sqrt{1+\tau}}}$$

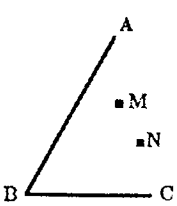
We now have the simplest continued fraction and root expansion known to man; each of which can be continued on *ad infinitum* in a systematic, patterned way. And as soon as one thinks of patterns, one thinks of symmetry. So here we have a systematic (symmetric) way to represent an irrational number, that in and of itself is almost a contradiction of terms, for irrational numbers do not have systematic decimal representations; yet with the notions of continued fractions and infinite square root expansions there is a certain symmetry to them.

Whether or not we have a subconscious gravitation towards symmetry, and special numbers such as the golden ratio, is admittedly a hazy area and perhaps best left to psychologists to investigate. But like learning to appreciate art and music, where one learns what to look for with respect to a painting, and what to listen for with respect to a piece of music of a particular period, one must be taught how to look for symmetrical relationships; gravitation towards symmetry might happen naturally, but *learning how to utilize symmetry must be taught,* (Dreyfus and Eisenberg, 1990).

TEACHING TO UTILIZE SYMMETRY

At the most basic level it helps to look at texts like the one written by the Hargittais, where the ubiquity of point and line symmetries are vividly pointed out and awake our sensitivity to geometrical symmetry. But the notion of symmetry enters many domains of school mathematics other than geometry. One of these domains is problem solving, where symmetry must be seen or imposed on a problem to effect its solution. Another domain is in concept formation, where it is often advantageous to think of basic mathematical notions in terms of symmetrical properties which surround them.

Figure 3 lists three problems whose solutions depend on symmetry. Do you see how to solve them and to generalize the problem? It has been our experience that most students cannot solve these problems, because they do not use symmetry as a heuristic tool. (Partial answers are presented at the end of this paper.)



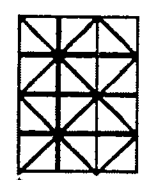
a

Find a point P on AB and a point Q on BC s. t. MP + PQ + QN is minimal.

Solve:

$x^4 - x^3 + x^2 - x + 1 = 0$

b



c

Given a billiard table 4 X 3. A ball stands in the lower left-hand corner and travels at an angle of 45°. This ball will end up in the upper left-hand pocket. Generalize.

Figure 3

But symmetry need not be limited to algebra and geometry. Let us consider the notion of Magic Squares.

Magic Squares: The idea underlying magic squares is to put n^2 integers into the cells of an nXn square so that the sums obtained by adding the numbers in each column, each row and each diagonal are equal. When we restrict the numbers to the positive integers 1, 2, 3, ... , n^2 we call the magic square a *normal magic square* of order n.

There are many problems associated with magic squares, e.g.,

- Show that it is impossible to have a magic square of order two.
- Show that the center cell in a 3X3 normal magic square must be 5.
- Show that it is impossible to have the number 1 in a corner cell of a normal 3X3 magic square.
- What must be the sum of each row in a 10X10 normal magic square?

Students seem to love problems with magic squares, but constructing the square itself, even a normal magic square, is not easy. There are two main classes of normal magic squares; those for which the order is odd and those in the order is even. And for those which are of even order, there is a further classification; those for which the order is divided by 4 and those for which it is not. For each of these there is a different algorithm to construct them. Because of space limitations we will only look at magic squares of odd order. (A comprehensive discussion of the underlying theory, construction and generalizations of magic squares is given in Ball (1949).)

The following algorithm can be extended to all magic squares of odd order; we exemplify it in Figure 4 for the 3X3 case. Start with 1 in the top middle cell. Now move along the diagonal, always moving upward and to the right. If in so moving, we find ourselves outside of the square's frame at the top, we continue filling in the cells of the square at the bottom of that column (Fig 4a). If we find ourselves outside the frame on the right, we move to the first cell on the left of that row (Fig 4b). If a cell is blocked within the frame itself because we have already filled it in, we drop to the cell immediately under the one from which we came, fill it in (Fig 4c) and continue filling in the square moving upward and to the right.

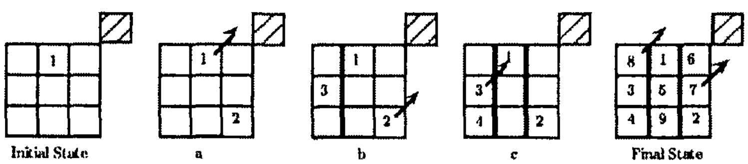


Figure 4

There are several ways to see that the square is really magic. One way is to actually check the sums in each row, column and diagonal. Another one is a little easier; e.g., to check to see if a 5X5 square is indeed magic, write each number in each cell in base 5. Now from the number in each cell subtract one from its base 5 representation. What remains is all combinations of two digit numbers in base five. It can be seen that each row and column and one of the diagonals are some permutation of the same digits. In other words, one need not do the actual computations, but simply check to make sure that all digits are there.

Students love to play with magic squares, but they are often unable to construct ones which are not normal. Here is where symmetry can enter because all magic squares of order 3 are of a particular form as shown in Figure 5.

a-p	a+p-k	a+k
a+k+p	a	a-k-p
a-k	a+k-p	a+p

Figure 5: Final State

Students are often amazed at how simple this is for them and they often try to generalize the method of building symmetry along the diagonals of higher ordered squares. This exercise opens up many doors for discussion, building in symmetry to effect a solution and degrees of freedom are just two avenues for deeper work.

Symmetry in cubic polynomials: There are two ways to prove that every cubic polynomial is symmetric about its point of inflection. A cubic polynomial is one of the form $f(x) = ax^3 + bx^2 + cx + d$ and its general graph is listed in Figure 6.

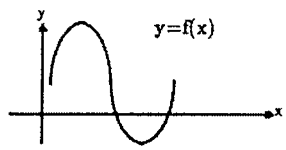


Figure 6

Its point of inflection is at $\left(\frac{-b}{3a}, f\left(\frac{-b}{3a}\right)\right)$ So, using its graph to guide us we wish to show that:

$$f\left(\frac{-b}{3a} - x\right) - f\left(\frac{-b}{3a}\right) = f\left(\frac{-b}{3a}\right) - f\left(\frac{-b}{3a} + x\right).$$

The algebra is not easy but straight forward.

Another way to do the problem is to play on its symmetry from the start.

First we look at the graphs of $y = x^3$ and $y = (x - 1)(x)(x + 1)$. Each of these graphs has its point of inflection on the y-axis, and the students easily prove that the graphs are symmetric with respect to their point of inflection. We then talk about the effect of adding a constant to the above equations, $y = x^3 + k$, $y = (x - 1)(x)(x+ 1) + k$. The students immediately realize that the effect of adding a constant to equations of this type simply pushes the graph up or down the y-axis; their points of inflection remain on the y-axis and thus the symmetry remains intact. The students thus realize that an equation of the form $y = x^3 + Bx + C$ has a point of inflection on the y-axis and that its graph is symmetric about it.

We then return to the general equation and ask if there is some way to transform it into one that will result in a cubic of the form $y = z^3 + pz + q$. The transformation $x = z - (b/3a)$ will move the graph horizontally so that its point of inflection is on the y-axis. In effect this gives another proof that cubic equations are symmetric about their point of inflection.

SUMMARY

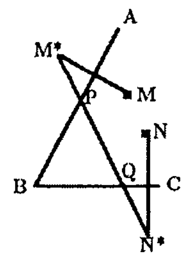
In this paper we have tried to show that symmetry is a concept which can be exploited and used as a red thread connecting different branches and skills in school mathematics. But the main message we have tried to argue is that symmetry must be taught. It is too useful and important a topic to let it develop casually, if at all, as one passes through the school curriculum.

ACKNOWLEDGMENT

We would like to thank Slavik Jablan for the wonderful animation of the static diagrams we submitted to him, and for the suggestions he made for improving the original draft of this paper as well as the encouragement he gave us along the way.

Notes on Figure 3

a) Let M^* be the reflection of point M through AB. Similarly, let N^* be the reflection of point N through BC. Let segment M^*N^* intersect AB and BC in points P and Q respectively. $MP + PQ + QN$ is minimal. Do you see why?



b) This is a symmetric polynomial. As such, we should focus on its middle term, namely x^2. Since $x \neq 0$ we can divide the polynomial by x^2.

$$x^4 - x^3 + x^2 - x + 1 = 0 \Longrightarrow x^2 - x + 1 - \frac{1}{x} + \frac{1}{x^2} = 0$$

Letting $t = x + (1/x)$ and noting that $t^2 - 2 = x^2 + (1/x^2)$ transforms the equation into a quadratic which gives a surprising solution.

c) Systematically consider special cases. A conjecture for the general case will soon be apparent. An elegant way to prove the conjecture is to use symmetry, by repeatedly reflecting a table of given dimensions about its edges. The path of the ball then becomes a straight line.

REFERENCES

Ball, W.W. (1949) *Mathematical Recreations and Essays,* London: Macmillan

Dreyfus, T and Eisenberg, T. (1990) Symmetry in mathematics learning, *Zentralblatt für Didaktik der Mathematik,* 1990/2, 53-59.

Dunlap, R.A (1998) *The Golden Ratio and the Fibonacci Numbers.*

Einstein, A. (1935) *The World as I See It,* London: John Lane

Hargittai, I. and Hargittai, M. (1996) *Symmetry: A Unifying Concept,* Bolina, Calif · Shelter Publications.

Herz-Fischler, R. (1998) *A Mathematical History of the Golden Number.*

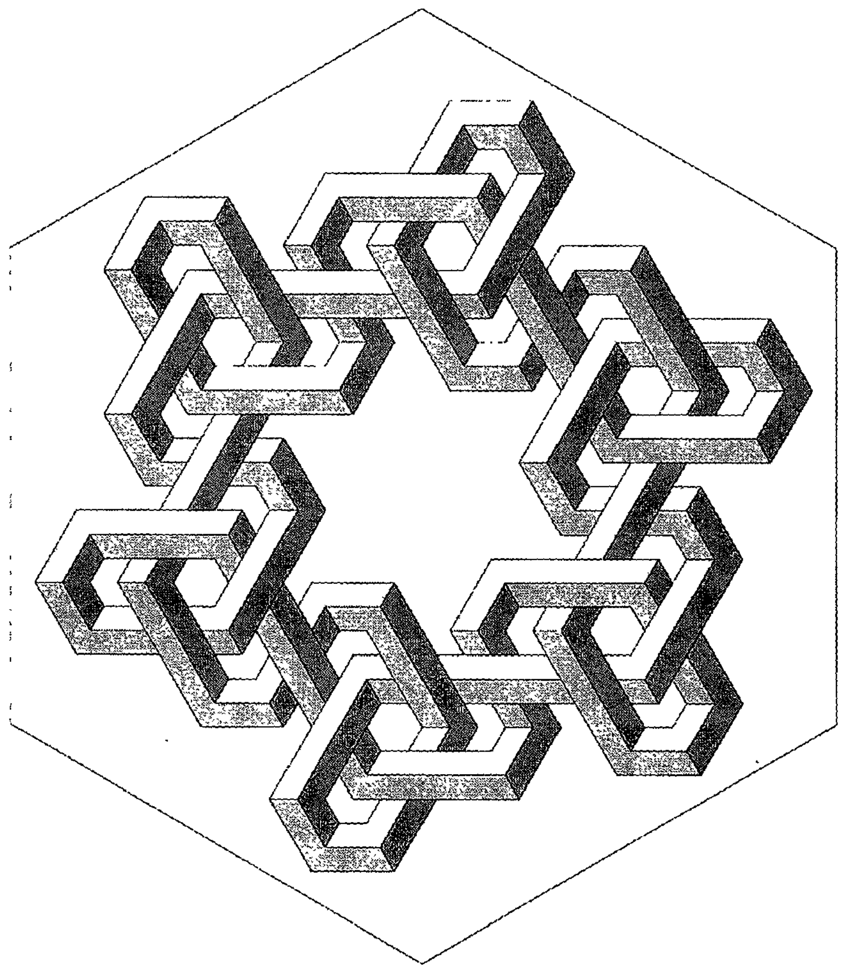
Huntley, H E. (1970) *Divine Proportion A Study in Mathematical Beauty,* New York, Dover.

Penrose, R. (1974) The role of aesthetics in pure and applied mathematical research, *Bulletin of the Institute of Mathematics and Its Applications,* 10(1974), 266-271.

Poincare, H. (1913) Mathematical creation, in· Newman, J R , ed , *The World of Mathematics,* vol 3., 2041-2050, New York. Simon and Schuster

Polya, G. (1967) *Mathematical Discovery,* vols. 1 and 2, New York: John Wiley and Sons.

Weyl, H. (1952) *Symmetry,* Princeton: Princeton University Press.

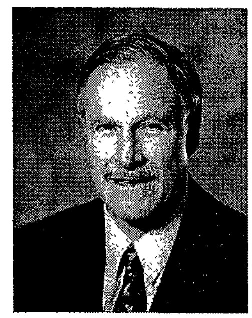


Symmetry: Culture and Science 199
Vol 9, Nos. 2-4, 199-205, 1998

MORPHOLOGICAL TRANSFORMATION OF HYPERBOLIC PATTERNS

Douglas Dunham

Mathematician, computer scientist (b. Pasadena, California, U.S.A., 1938)
Address: Department of Computer Science, 320HH, University of Minnesota-Duluth, Duluth, Minnesota 55812-2496, U.S.A E-mail: ddunham@d umn edu.



Fields of interest: Hyperbolic geometry, M. C. Escher, computer graphics algorithms, computer generated art.

Publications: Families of Escher patterns, presented at *M. C. Escher 1898-1998, An International Congress* (1998) June, Rome, Italy, *The Symmetry of hyperbolic Escher patterns, *Proceedings of the Third Interdisciplinary Symmetry Congress* (of the International Society for Interdisciplinary Studies of Symmetry: ISIS-Symmetry) (1995), August, Alexandria, VA; *Artistic patterns with hyperbolic symmetry, *Proceedings of Symmetry of Structure, An Interdisciplinary Symposium* (1989) August, Budapest, Hungary; *Creating hyperbolic Escher patterns, *M. C. Escher, Art and Science*, Coxeter, H.S.M. et al., eds., North-Holland, New York (1986) pp. 241-248; *Hyperbolic symmetry, *Computers & Mathematics with Applications* 12B (1986) nos. 1/2, pp 139-153.

Abstract: *The Dutch artist M. C. Escher is known for his repeating patterns of interlocking motifs. Most of Escher's patterns are Euclidean patterns, but he also designed some for the surface of the sphere and others for the hyperbolic plane, thus making use of all three classical geometries: Euclidean, spherical, and hyperbolic. In some cases it is evident that he applied a morphological transformation to one of his patterns to obtain a new pattern, thus changing the symmetry of the original pattern, sometimes even forcing it onto a different geometry. In fact Escher transformed his Euclidean* Pattern Number 45 *of angels and devils both onto the sphere,* Heaven and Hell *on a carved maple sphere, and onto the hyperbolic plane,* Circle Limit IV. *A computer program has been written that converts one hyperbolic pattern to another by applying a morphological transformation to its motif. We will describe the method used by this program.*

1 INTRODUCTION

M. C. Escher created many repeating patterns of the Euclidean plane. In a few cases he distorted or "morphed" these patterns to obtain new patterns in other geometries: spherical or hyperbolic. Escher's Pattern Number 45 of angels and devils is the only one that he converted to both the sphere and the hyperbolic plane. These three related patterns are shown in (Schattschneider 1990) on pages 150, 244 and 296; see Figure 1 below for *Circle Limit IV*. Professor Coxeter discusses the three patterns on pages 197-209 of Coxeter 1981.

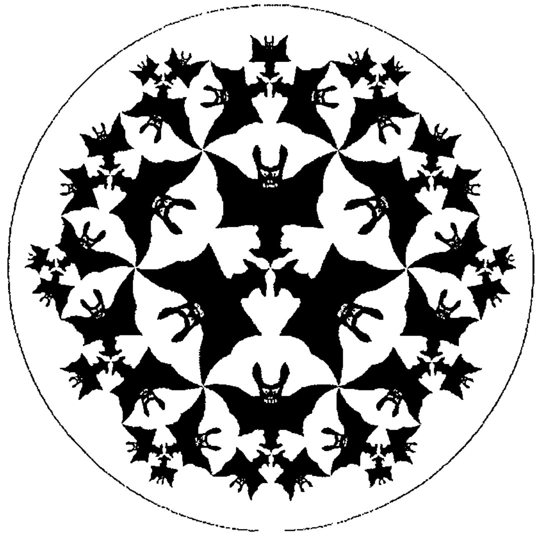


Figure 1: This is Escher's hyperbolic pattern *Circle Limit IV* of angels (the white background) and devils (foreground).

There are probably many ways to distort or "morph" one pattern into another. The method we will describe applies to repeating patterns based on the regular tessellations, $\{p; q\}$, composed of regular p-sided polygons meeting q at a vertex. Thus, given one repeating pattern, we could theoretically create a doubly infinite family of related patterns by morphing the original pattern into others based on different values of p and q. Many of Escher's Euclidean patterns and all of his spherical and hyperbolic patterns are based on $\{p; q\}$. For example, his Euclidean Pattern Number 45 and the related spherical and hyperbolic patterns mentioned above are based on the tessellations $\{4; 4\}$, $\{4; 3\}$ and $\{6; 4\}$ respectively. Figure 2 shows the tessellation $\{6; 4\}$ superimposed on *Circle Limit IV*. In these patterns, p is twice the number of angels/devils that meet at their feet and q is the number of wing tips that meet at a point. The meeting point of feet is the intersection of lines of bilateral (reflection) symmetry – hence the need to double the number of angels/devils to obtain p.

We will begin with a brief review of hyperbolic geometry. Next we discuss repeating patterns and regular tessellations, and the morphological transformation process, showing an example. Finally we suggest possible further directions of research.

2 HYPERBOLIC GEOMETRY

Unlike the Euclidean plane and the sphere, the entire hyperbolic plane cannot be isometrically embedded in 3-dimensional Euclidean space. Therefore, any model of hyperbolic geometry in Euclidean 3-space must distort distance. The *Poincaré circle model* of hyperbolic geometry has two properties that are useful for artistic purposes: (1) it is conformal (i.e., the hyperbolic measure of an angle is equal to its Euclidean measure) – thus a transformed object has roughly the same shape as the original, and (2) it lies within a bounded region of the Euclidean plane – allowing an entire hyperbolic pattern to be displayed. The "points" of this model are the interior points of a *bounding circle* in the Euclidean plane. The (hyperbolic) "lines" are interior circular arcs perpendicular to the bounding circle, including diameters. The sides of the hexagons of the {6; 4} tessellation shown in Figure 2 lie along hyperbolic lines as do the backbone lines of the fish in Figures 3 and 4.

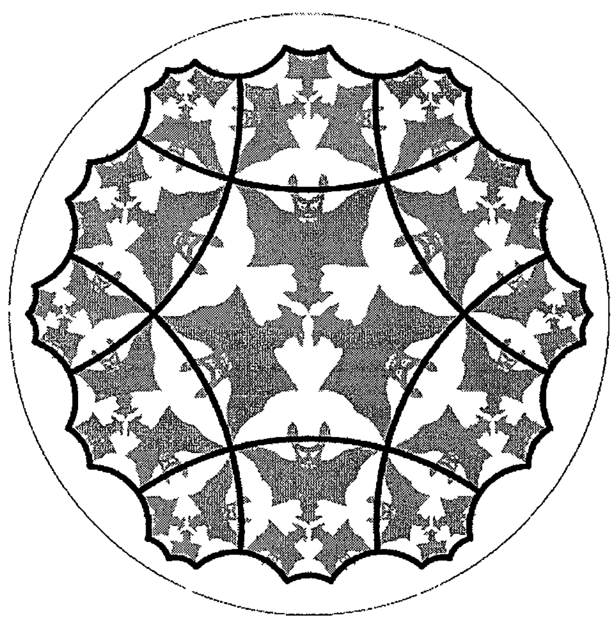


Figure 2: This is *Circle Limit IV* showing the underlying {6, 4} tessellation

3 REPEATING PATTERNS AND REGULAR TESSELLATIONS

A *repeating pattern* of the Euclidean plane, the hyperbolic plane, or the sphere is a pattern made up of congruent copies of a basic subpattern or motif. For instance, a black half-devil plus an adjacent white half-angel make up a *motif* for Figure 1.

An important kind of repeating pattern is the *regular tessellation*, $\{p;\ q\}$, of the plane by regular p-sided polygons, or p-*gons*, meeting q at a vertex. The values of p and q determine which of the three "classical" geometries, Euclidean, spherical, or hyperbolic, the tessellation lies in. The tessellation $\{p;\ q\}$ is spherical, Euclidean, or hyperbolic according as $(p\text{-}2)(q\text{-}2)$ is less than, equal to, or greater than 4. This is shown in Table 1 below. Note that most of the tessellations are hyperbolic. In the spherical case, the tessellations $\{3;\ 3\}$, $\{3;\ 4\}$, $\{3;\ 5\}$, $\{4;\ 3\}$ and $\{5;\ 3\}$ correspond to versions of the Platonic solids (the regular tetrahedron, octahedron, icosahedron, cube, and dodecahedron respectively) "blown up" onto the surface of their circumscribing spheres. One can interpret the tessellations $\{p;\ 2\}$ as two hemispherical caps joined along p edges on the equator; similarly $\{2;\ q\}$ is a tessellation by q lunes. Escher's only use of these latter tessellations appears to be the carved beechwood sphere with 8 grotesques (Schattschneider 1990, p. 244) based on $\{2;\ 4\}$. The tessellations $\{3;\ 6\}$, $\{4;\ 4\}$ and $\{6;\ 3\}$ are the familiar Euclidean tessellations by equilateral triangles, squares, and regular hexagons, all of which Escher used extensively.

q \ p	1	2	3	4	5	6	7	8	9	10	11	⋯
⋮		⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	
11		o	*	*	*	*	*	*	*	*	*	⋯
10		o	*	*	*	*	*	*	*	*	*	⋯
9		o	*	*	*	*	*	*	*	*	*	⋯
8		o	*	*	*	*	*	*	*	*	*	⋯
7		o	*	*	*	*	*	*	*	*	*	⋯
6		o	□	*	*	*	*	*	*	*	*	⋯
5		o	○	*	*	*	*	*	*	*	*	⋯
4		o	○	□	*	*	*	*	*	*	*	⋯
3		o	○	○	○	□	*	*	*	*	*	⋯
2		o	o	o	o	o	o	o	o	o	o	⋯
1												

□ - Euclidean tessellations
○ - "Platonic" spherical tessellations
o - spherical tessellations for which $p = 2$ or $q = 2$
* - hyperbolic tessellations

Table 1: This table shows the relation between the values of p and q, and the geometry of the tessellation $\{p;\ q\}$.

4 MORPHOLOGICAL TRANSFORMATION OF A PATTERN

The basic version of the computer program that performs the morphological transformation requires that the motif be contained in one of the p isosceles triangles formed by the radii of a p-gon. Figure 3 below shows the isosceles triangles within a 6-gon of {6; 4} that is the basis of Escher's *Circle Limit I* (shown in gray). A natural motif in *Circle Limit I* is composed of a black half-fish and an adjoining white half-fish, however such a motif has part of a white fish fin protruding outside its isosceles triangle. This motif can be modified to the required form by clipping off the protruding part and "gluing" it back between the tail and the back edge of the fin of the black fish.

The program has been extended slightly so that this modification is often not necessary. The extended program also seems to work reasonably well with a motif that overlaps two adjacent isosceles triangles (with roughly half the motif in each triangle) – as is the case with *Circle Limit IV* (Figure 1).

The basic morphing process makes use of the Klein model of hyperbolic geometry. As with the Poincaré model, the points are interior points of a bounding circle, but the hyperbolic lines are represented by chords. We let I denote the isomorphism that maps the Poincaré model to the Klein model. Then I maps a centered p-gon with its isosceles triangles to a regular p-sided polygon which also contains corresponding isosceles triangles. Different tessellations $\{p; q\}$ produce different isosceles triangles in the Klein model, but an isosceles triangle from $\{p; q\}$ can be mapped onto an isosceles triangle from {p'; q'} by a simple (Euclidean) differential scaling, since those isosceles "Klein" triangles are represented by isosceles Euclidean triangles. Thus the morphological transformation from a $\{p; q\}$ pattern to a $\{p'; q'\}$ pattern can be accomplished by (1) applying I to a motif in an isosceles triangle of $\{p; q\}$, (2) applying the differential scaling to that transformed triangle, and finally (3) applying the inverse of I to the re-scaled triangle containing the motif. The entire pattern can then be formed by replicating the morphed motif. Replication algorithms are discussed in Dunham 1986a and Dunham 1986b. Figure 4 shows the result of morphing the *Circle Limit I* pattern to a {4; 6} pattern – with a transformed isosceles triangle superimposed.

Using similar techniques, another program has been written to transform isosceles Euclidean triangles to isosceles hyperbolic triangles, and thus Euclidean Escher patterns (of which there are many) can be transformed to hyperbolic patterns.

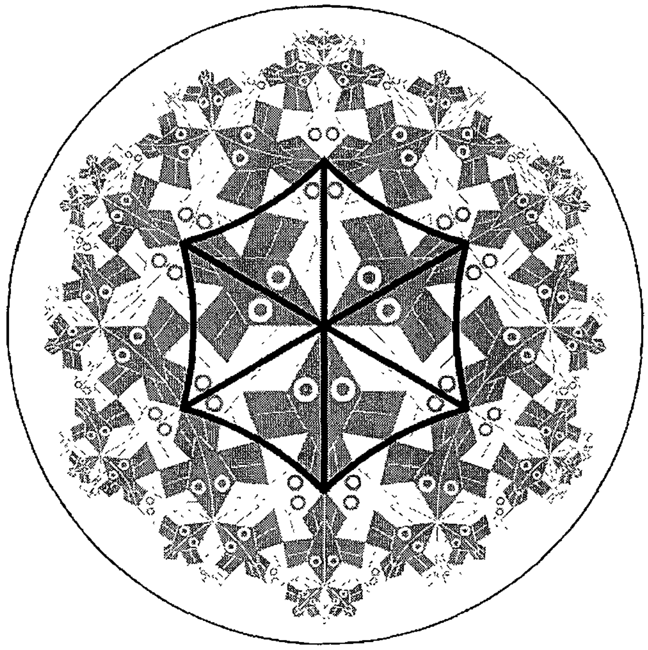


Figure 3: The isosceles triangles superimposed on *Escher's Circle Limit I* pattern.

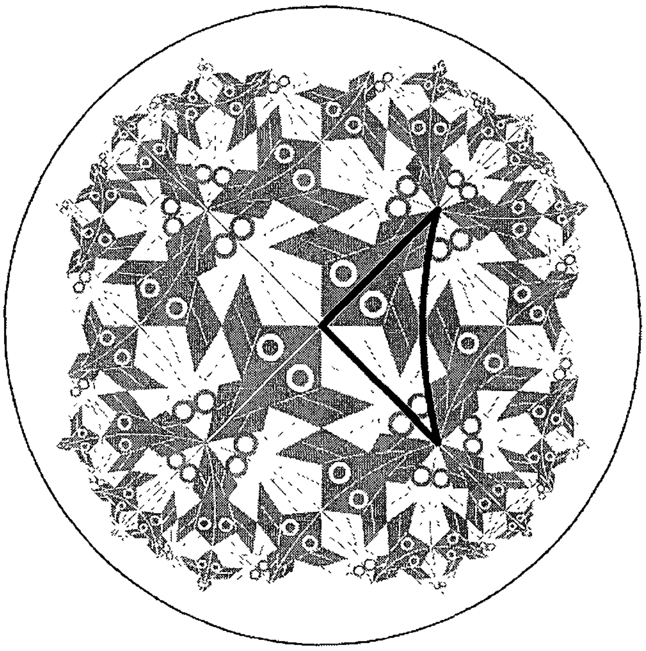


Figure 4: A morphed Circle Limit I based on {4; 6} showing an isosceles triangle.

5 FUTURE WORK

Directions of future research include: (1) finding different morphing transformations, (2) allowing the fundamental region to be a non-isosceles triangle or quadrilateral, and (3) transforming between any of the three classical geometries. The morphing transformation described above is not conformal. Theoretically, the Riemann Mapping Theorem says there is a holomorphic (and hence conformal) isomorphism between the Poincaré isosceles triangles of any two tessellations $\{p; q\}$ and $\{p'; q'\}$. A natural fundamental region for Escher's *Circle Limit III* is a quadrilateral divided into two triangles whose sides are two hyperbolic line segments and a segment of an equidistant curve – the above methods may extend to such triangles. Finally, transforming from spherical to Euclidean (and hence hyperbolic) patterns would only involve finding a mapping from isosceles spherical triangles to isosceles Euclidean triangles.

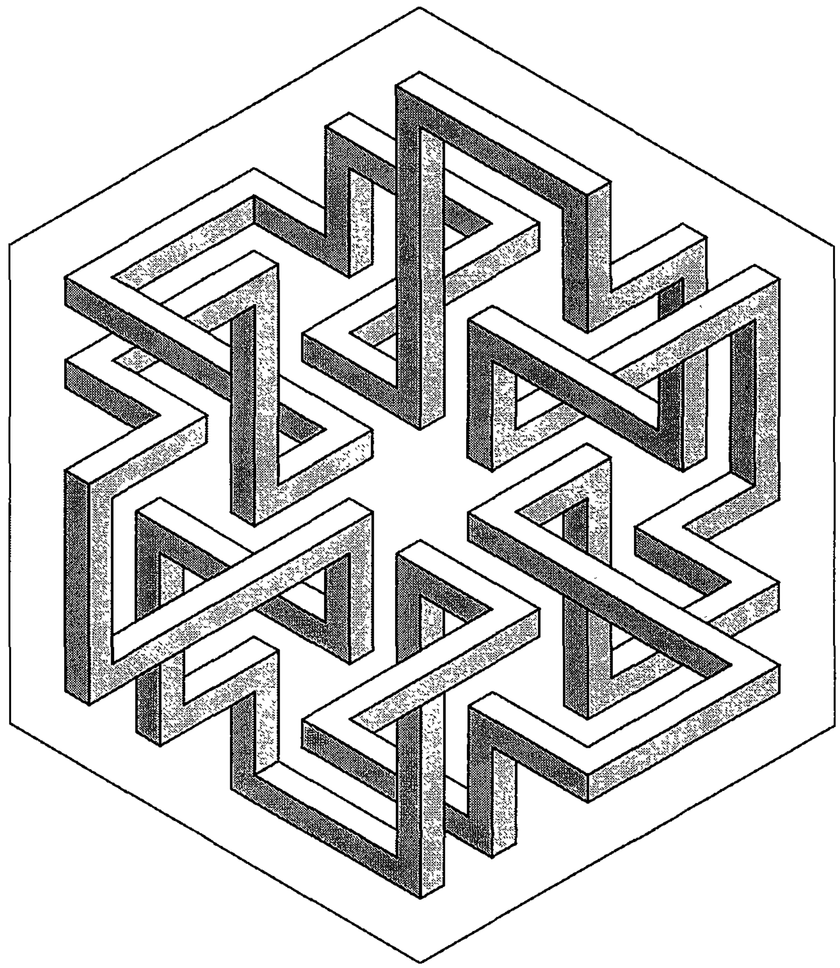
REFERENCES

Coxeter, H. S. M (1981) Angels and Devıls, ın Klarner, D A., ed. *The Mathematical Gardner*, Boston· Prindle, Weber and Schmidt.

Dunham, D. (1986a) Hyperbolic symmetry, *Computers & Mathematıcs with Applications* (1/2), 12B, pp. 139-153.

Dunham, D. (1986b) Creatıng hyperbolic Escher patterns, ın Coxeter, H.S.M et al., eds. *M.C Escher. Art and Science*, Amsterdam: North-Holland, pp. 24-248.

Schattschneider, D. (1990) *Vısions of Symmetry: Notebooks, Periodic Drawıngs, and Related Work of M. C. Escher*, New York. W. H. Freeman.



Symmetry: Culture and Science 207
Vol. 9, Nos. 2-4, 207-215, 1998

AN INTEGRAL SPATIAL APPROACH TO COMPARATIVE EVALUATION OF DENSE ENVIRONMENTS

Fisher-Gewirtzman D, Tzamir Y, Burt M.

Fisher-Gewirtzman, D.: Architect, (b. Haifa, Israel)

Address: Faculty of Architecture and Town Planning, Technion – Israel Institute of Technology, Haifa 32000, Israel. *E-mail* ardafna@tx.technion ac.il.

Fields of interest. Architecture design, Urban design, Interior design

Abstract: *Density measurements are conventional tools for space utilization and manipulation. All density measurements refer to ratios between dwelling units (number of houses or number of people) per area unit. A two dimensional density measurement can not reflect the spatial quality of an environment, particularly not of a dense environment. The paper describes an effort to develop a new conceptual approach and a method to 3-D density measurements and evaluation. This approach reflects the hierarchical-order, which is found in the urban environment. At the same time, it's disorder – the deviation from pure geometrical considerations. The target of this approach is to facilitate a reliable method for spatial quality evaluation, thus offering the basis of which an important urban planning and design tool can be constructed.*

INTRODUCTION

It is expected that around the year 2025 more than 60% of the projected 8-billion world population will live in cities. Furthermore, there is a tendency in many countries, mostly in Western Europe and the USA (as in Israel) to increase density in metropolitan areas. The reasons vary. In most cases this is due to environmental quality problems and of financial, transport and communication reasons. In Israel the main reason is to protect the decreasing reserves of land. In order to avoid dramatic damages to the environmental quality, this change necessitates new tools and strategies in urban design and urban planning. In the 21-century more planning efforts will be directed at coping with dense urban regions and, the interrelation of quality in dense built-up environments.

Environmental quality is an important component of our quality of life. It is related to many parameters, some quantitative and some qualitative. Quantitative parameters are related to climatic issues such as sun radiation, natural light intensity, wind forces and of course to conventional density measurements. Qualitative parameters refer to attractiveness, or repulsion, in which various features play a significant role, such as privacy, meaning, texture, colour, odour, quality of buildings, nostalgia, memory and more physical and psychological aspects. These aspects depend upon the cultural background of the population, socio-economic background and age; they sometimes depend on momentary state of mind.

This work focuses on the spatial configuration of the urban fabric, as we are convinced that "the physical environment and the configuration of the urban space, its qualities and characteristics have a major influence on the human perception and human behaviour" (Lang, 1994). In dense urban environments, spatial-organisations, has a great impact on perception, as they are more direct and more elementary experience than individual details. To anticipate spatial configurations that would reflect low perceived density and a way to identify and evaluate them, an objective index that would represent and evaluate the spatial quality in dense urban environments is required. This work introduces a new conceptual approach for an objective comparative evaluation of the spatial quality of dense urban environments, conditioned to human perception.

THE PROBLEM

Density measures are constantly used as design tools in many planning and design activities, such as architectural design, urban design, and urban and regional planning. The application of density measures is aimed at affecting the form and organisation of the built environment. All applied density measurements are ratios. The numerator may be the number of persons, families, households, habitable room, bedrooms, housing units or dwelling units. The denominator is a unit of area. It should be stressed that the present method of density measurements as a predicting and evaluating tool is quite adequate and relevant to the macro-scale planning (of regions and cities). As for buildings, city block or even neighbourhood scale, the spatial density perception could not be captured and adequately represented by the 'unit per area' ratios. On this scale the current tools are deficient to a point of irrelevancy.

For most of the built-up situations a two dimensional density measurement approach is bound to fail in representation of a three-dimensional spatial configuration, and clearly fail to predict spatial quality and evaluate a three-dimension perceived reality. A specific density measurement in itself *cannot* lead to a perceived quality of the environment. The same net-density measurement can be applied in many different spatial configurations, with different perceived densities and feeling of comfort. The problem increases with increasing densities, as the number of possible spatial organisations, in a given space, tends to infinity. Alexander (1988, 1993), points out that many density analyses raised the problems embedded with density measurements and their applications. These problems include indeterminacy and ambiguity, over simplification and possibly a weak relationship with perceived density, which after all is what measured densities are ultimately about.

THE NEW CONCEPTUAL APPROACH

The target of this work is to define an objective index to evaluate spatial quality, analogous to the human perception and evaluation. "We are aided in the *perception of space* by the effect of *movement in time*" (Gombrich, 1972). To develop a comprehensive evaluation for the built environment, it is necessary to visualise and perceive it from every reasonable viewing point, as it occurs in a space-time experience of the urban environment. Space–Time experience influences the viewer's perception and evaluation of that environment. Perception would result from the collection of views accumulated on the tour around the built environment. Hypothetically, the perception of space would be a function of all the views collected along a 'Space-Time' experience. The viewer's experience can be represented also as a collection of pictures from all possible angles on the tour, as illustrated on Fig. 1 and 2.

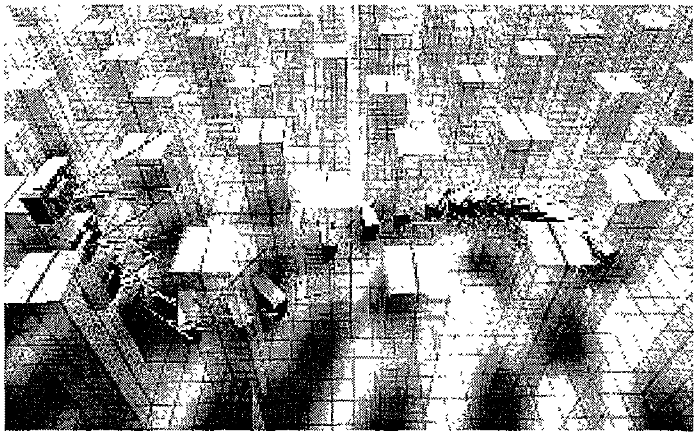


Figure 1: Illustrating "Space-Time" experience track

The collection of perspectives displays what is visually perceived through a spatial conical angle. The suggestion is that the sum of the overall 'Spatial Conical Angles' would reflect the human visual perception in the most appropriate way. This Spatial Conical Angle can be measured and used as a quantitative index, which reflects the spatial quality of the environment. Fig. 3 illustrates a spatial conical angle observed from a dwelling unit within a built complex. We name this index 'Spatial Openness'.

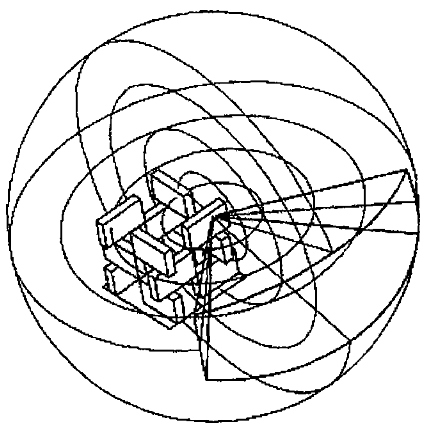


Figure 3: "Spatial Openness" Index (SOI)

The 'Spatial Openness' index measures the volume of open space, in pure geometrical morphological terms, defining openness to natural light, air, near and distant view. This quality index in quantitative terms can provide a comparative spatial quality evaluation of various spatial configurations. Our quality index actually refers to the volume of free space, which is observed from a specific viewers-position, as illustrated in Fig. 3.

As the urban environment is defined by the spatial organisation (composition and interrelation) of the built volumes and their density, there is high dependency among Spatial Openness on all urban fabric hierarchical levels. The dwelling unit, depends and influences the building compound, which also depends and influences the entire urban environment (on the neighbourhood scale). We assume 'Spatial Openness' on one level would reflect the integrated sum of 'Spatial Openness' of all lower levels. The 'Spatial Openness' of a specific building would be reflected by the sum of 'Spatial Openness' measured from all meaningful viewpoints (apartments and public spaces) relating to that building.

INDICATION TEST

Based on Rapoport (1977), who suggested that "Most times people see low perceived density as one of the characteristics of high quality environment" our hypothesis is: as the 'Spatial Openness' index would be higher it would indicate a higher spatial quality and lower perceived density.

We hypothised a correlation between the *'Spatial Openness' index* and a *'Perceived density'*. We assume that spatial configurations with a comparative high level of spatial openness would be perceived as less compressed and evaluated as more spacious. To obtain an indication for such correlation we synthesised five groups of alternative spatial configurations, all with the identical built-volume, and dispersed within a given volume of space. We measured an approximated 'Spatial Openness' to all alternative spatial configurations, so we could rank them within their groups. At the same time we asked a group of twenty-five participants to rank the alternative configurations by their relative "perceived density".

Architects and students of architecture were asked to participate in our indication test, as it was important that the participants would be able to make the analogy from models and computer representations to real environments. Fig. 4 illustrates part of the

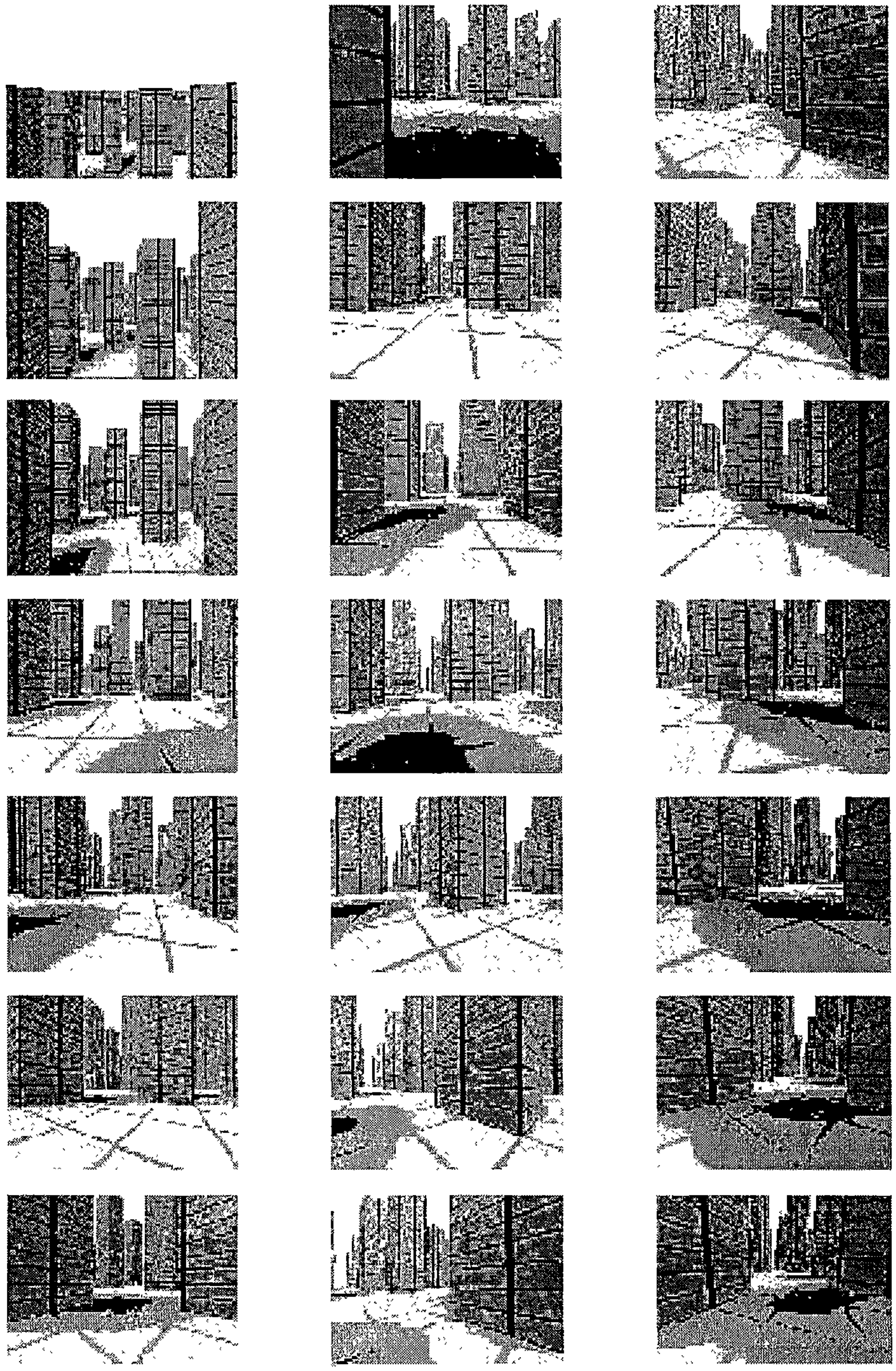


Figure 2: illustrating a collection of views from a 'Space-Time' experience.

information presented to the participants for one of the groups of alternative spatial configurations: it represents a basic spatial configuration, an urban fabric and a series of possible views taken within each alternative.

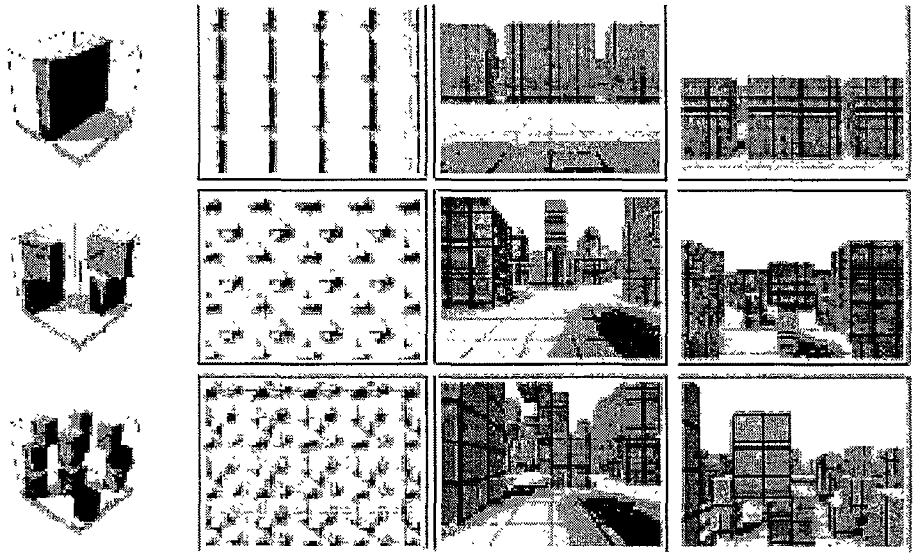


Figure 4: Comparative views from alternative spatial configurations.

High correlation was found in three out of the five groups, as demonstrated by the correlation graph, Fig. 5. This graph represents the correlation within objective measure of "Spatial Openness" with subjective "perceived density" evaluated by our participants. High correlation can give us good input for future tests. At the same time, group's (2) and (5), which had no correlation, may teach us more about the problematic aspects of subjective evaluation and propose the existence of extended aspects influencing the human perception and evaluation. These aspects introduce distortion and disorder into our system and must be taken into consideration in future development.

These preliminary tests gave us an important indication and encouragement for further development of conceptual background and a tool for objective spatial evaluation. We intend to develop this concept as a future objective spatial evaluation tool. This tool will relate in a better way to the structural and spatial organization of the urban environment and will be able to guide, predict and evaluate spatial quality of the dense urban environment.

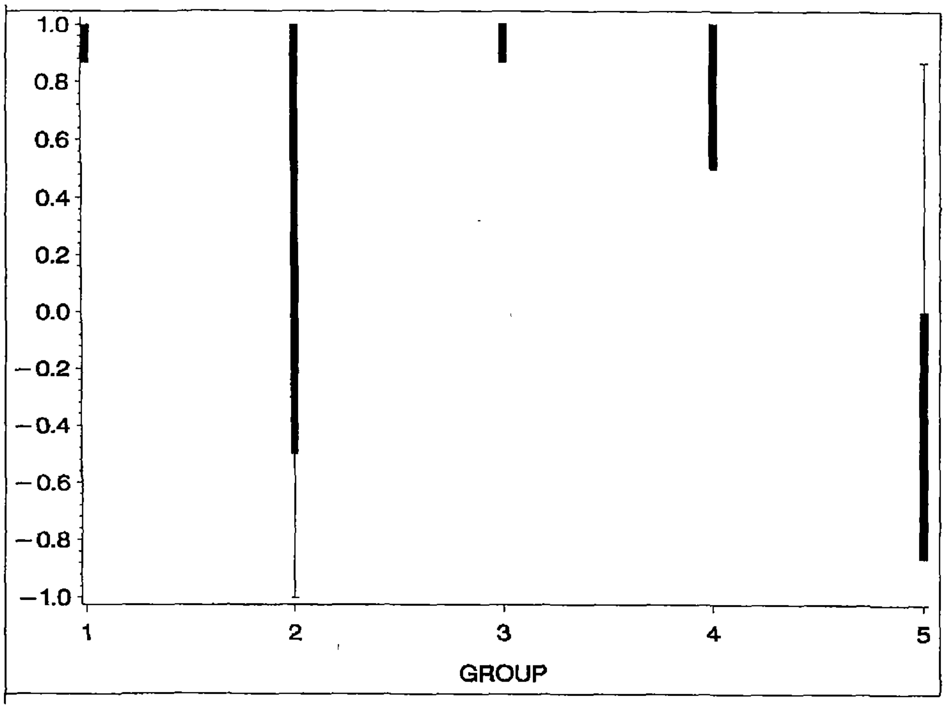


Figure 5: Graph representing "Spatial Openness" index correlation with "perceived density" evaluation.

SUMMARY

Density measurements are conventional tools for the examination of space utilisation and manipulation. A two dimensional density measurement can not reflect the spatial impression of the environment, especially not in dense environments. Our work, presented in this paper, proposes a new approach. We suggest a quantitative-index that applies comprehensive spatial quality information, which we name 'Spatial Openness'. We hypothised a possible correlation between the "spatial Openness" index and "perceived density" evaluated by people responding to alternative spatial configuration. Correlation was found among some groups of the tested spatial configurations. The correlation gave us indication and encouragement for developing the tool for objective spatial evaluation. At the same time the negative correlation had opened up many wonders, mainly concerned with subjective evaluation and emphasised the need for further research to be maintained in this field.

Acknowledgement: The statistical analysis, presented in this paper, was financed by the R. and L. G. Restone Research Fund in Architecture.

REFERENCES

Alexander, C. (1963) *Community and Privacy Towards a new Architecture of Humanism*. Double Day & Comp. Inc. USA.

Alexander, E., Reed, D and Murphy, P (1988) *Density Measures and their Relation to Urban Form, Publication in Architecture and Urban planning research*. University of Wisconsin Milwaukee, Wisconsin, USA.

Alexander, E R (1993) Density measurements: A review and analysis. *Journal of Architecture and Planning research*, 10, 3, 181-202.

Arnheim, R. (1977) *The Dynamics of Architectural Form*. University of California Press, Berkley, California.

Bitan, A. (1997) Climatic related planning strategies in compact urban environments. Methodology – Tools – Implementation. *Second International Symposium on Urban Planning and Environment* The Netherlands.

Fisher Gewirtzman, D., Tzamir, Y. and Burt, M. (1997) (Abstract book) Density-Quality interaction within the context of high density urban development. *Second International Symposium on Urban Planning and Environment*. The Netherlands.

Fisher Gewirtzman, D., Tzamir, Y. and Burt, M. (1998) (Abstract book) Systematic approach to urban density measurements as a basis for environmental quality evaluation. *Urban Sustainability and Cultural Processes International Conference*. Stockholm, Sweden

Gidion, S (1967) *Space Time and Architecture. The Growth of a New Tradition*. USA.

Gombrich, E. H., Hochberg, J. and Black, M (1972) *Art Perception and Reality*. London: The John Hopkins University Press.

Lang,, J. (1994) *Urban Design, The American Experience*. Van Nostrand Reinhold, New York, NY.

Steadman, J. P. (1983) *Architectural Morphology*, London.

Rapoport, A (1982) *The Meaning of the Built Environment*. Beverly Hills: Sage Publications, USA.

Rapoport, A (1977) *Human Aspects of Urban Form*, Oxford: Pergamon Press, USA.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 217-230, 1998

THE APPEARANCES OF THE DIFFERENT FINE ARTS IN MUSIC

Katalin Fittler

Musician, piano teacher (b. Mosonmagyaróvár, Hungary, 1952)
Addresses: (1) Hungarian Broadcasting, Budapest, H-1800; (2) Városkapu tér 3. Mosonmagyaróvár, H-9200.
Fields of interest: paleography, semantics, symbols in music.
Publications: Mozart's Freemasonic Compositions, *Magyar Zene* (1980) 2; Beethoven's Symphonies in Liszt's Transcription, *Magyar Zene* (1986) 1.

Abstract: *It needs hardly a proof that music occupies a special place among thr arts. So it is no accident, that musical commentary has got dilemmas, already historic ones.*
The division of arts has many branches: among then one of the most important trend is that, which makes a difference between arts accordingly to their approach of the description of phenomena, whether it is simultaneous or successive.
From this point of view one can make differences between timely and spacely arts – but this approach is true only in rudiments, because even spacely creations need time to accept, while the timely creations suppose some space as well.
Space and time, as ideas have different meanings in the different branches of art. The time of music is doubtle: closed and open. The same questions are worth to be examined also in connection with forms, with the analysis of visible and audible works. The problems of proportion-feeling offer a rich material too.

The Latin word *ars* has been meaning *trade* and *art* at the same time. While the branches of the former have been formed out of practical *necessity*, in consequence to a specialization, the development of the latter branches have become possible through a sort of surplus, the free time, remained after the activities needed for keeping life and race going.

The production of such activities, born from this richness, is a considerable part of our weekdays and holidays, for already more than a thousand years.

Looking back into the music-historical past, it is interesting, that music, not being yet an independent, autonomous art, still, has already become worthy of attention to those, who were interested in a theoretical aspect. For example around 1270, *Cardinal Bonaventura* summed up the characteristics of creation of art as follows: beautiful, useful and firm. He added: in music only the first two qualities may be applied, music cannot be firm, "*opus stabile*". So is its difference from the other arts is a well-known fact, for a long time.

At the end of the 15th century, Adam von Fulda, a German musician defined music just because of its momentary existence, as a meditation over death (*meditatio mortis*).

An important personality of the Caroling renaissance spirit, Alkuin, chancellor of Charles the Great, signified the *order*, *measure* and *clarity* as main characteristics of beauty. (The definition corresponds to Aristotle's Metaphysic: "The main forms of beauty are order, proportion and delimitation.")

So music, beyond all its reality is some sort of mirage, created (or recreated) from time to time. The new music of every age is the "crop" itself, and the resounding of historic music pieces and the record means the recreation.

Let us approach now this branch of art in the function of *time* and *space*, which play an important role in every branch of art (although differently interpreted).

From the musical point of view the primary dimension is *time.* Time in the music is closed and open. Closed because one can perceive the sounds right in the given moment, and open because each tone holds the memory of the preceding events and suggests the expectation for another one. So, listening to music, not only the real sense of time but the remembrance plays an important part, too. And just the simultaneous feeling of the different time levels makes it possible for us to "feel" the proportions in music.

It is a different thing to catch the sense of an art creation of made for eyes. In that case the *whole* is ready in advance before us and after having a sight, one can get the whole range and the sense of it only in the *very end* of a piece (or a movement). In this case it is the memory that helps us with recording each bit of details which are elements in a progression and make a comprehensive structure.

Looking for the connection of "forms" in fine arts and in music, let us stay at first in the plain.

There are illustrative figures, the sight of lines of which one thinks of a special phenomenon. It is no use knowing that the heart is no heart-shaped, seeing the symmetric line one always brings this sight into connection with a sensitive content.

How to characterize the connection which exists between the contour of the drawing and the real sounding phenomenon? The movement of many parts "*Belle, bonne, sage*" by Baude Cordier, has an amorous content. The connection between sight and content is concrete – but on the basis of sound one can scarcely reconstruct a fixed figure in his imagination. (The same chanson can be fixed with traditional notes, too.)

Figure 1

Another "heart": the contemporary Hungarian composer, *József Sári* aimed in his series of piano "Image-Music" at reviving such note-images that are on the basis of sound to rewrite, decode into a special geometrical formation, contour. One that knows that note-writing in our days is a coordinate system, in which the spatial (depth-height) relations are expressed in the direction of the perpendicular, and the timely relations in the direction of horizontal level, listening to the series of intervals, may easily find out that this means a note-image in the form of a heart (Figure 1).

Figure 2: József Sári – *Képzene* ("Image-Music")

It is a question of point of view which sort of heart-table one finds to be more direct. (As Mendelssohn has composed his piano pieces under the title „*Lieder ohne Worte*", it is understandable that the composer does not comment a concrete program either. This would be superficial, as the note-picture speaks for itself.)

Similarly talkative is the sound-shaping of the first human couple, Eve and Adam (Figure 3).

Figure 3

The "*homo ludens*" has worked in József Sári, too, as he has expressed in music a characteristic child's drawing: a house with chimney fuming (illustrated with a sign of trill), a fence at the sides (Figure 4).

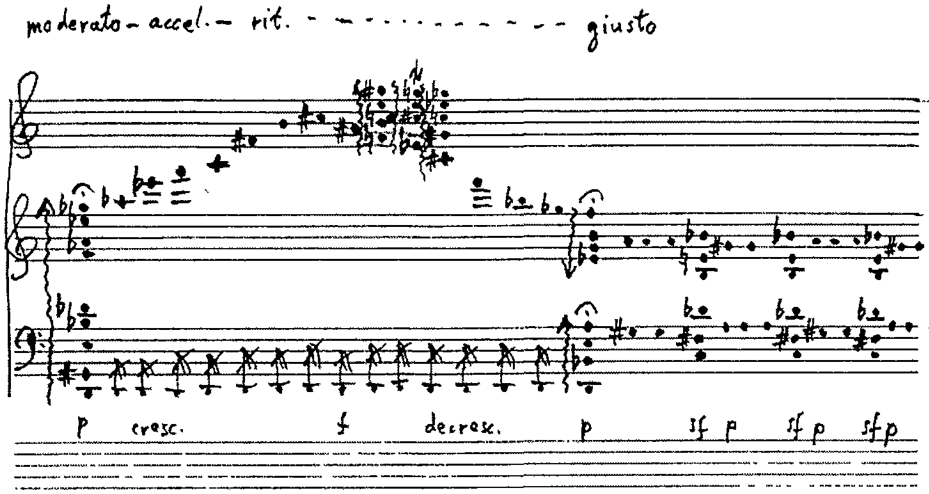
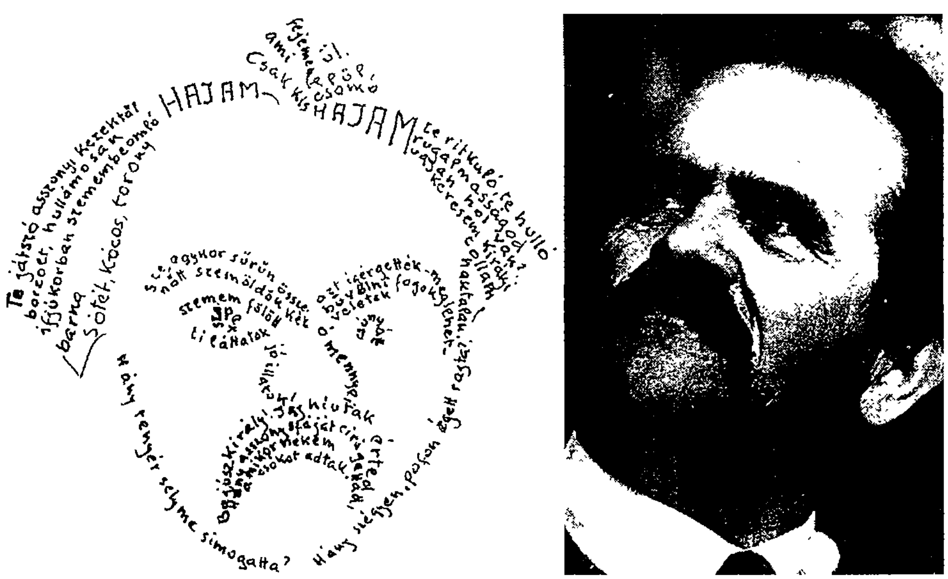


Figure 4

A valuable example from the Hungarian literature: the title of *Károly Jobbágy*'s picture-poem is: Selfportrait. Listening to its recitation, one can hardly guess what sort of written form it has (Figures 5).



Figures 5

The opposition of the left and right sides means often the feeling of "here" and "there". One can find an example for this in the painting. About such a sort of symbolic Ferenc Liszt wrote in his poetical letter about Raffael's painting Saint Cecily (in 1839): "The painter places Paul and John on the *left* of the picture: the former is deeply absorbed in himself, the *outer* world ceases to exist for him; behind his giant figure immense profoundities are lurking. John is a man of 'attractions' and 'feelings'; an almost feminine face looks out at us. On the other hand, Augustine on the *right* of Cecilia, maintains a cool silence... he abstains even from the most sacred *emotions* – constantly fights against his feelings. On the right edge of the picture stands Magdalena in the full splendour of her worldly finery; her whole bearing suggests worldliness, her personality radiates a *sensuousness* somewhat evocative of Hellas... Her love stems from the senses and adheres to *visible* beauty. The magic of sound captivates her ear faster than her heart is possessed by any supernatural excitement." (Lendvai 1993, p. 61)

Let us now come to the examination of space-forms. The Pantheon in Rome is a tipically "closed" building, a special "material" form. Such buildings are characterized by the harmony of relations, one thinks them to be the incorporations of "sensual beauty", independent of the onlooker.

On the contrary, on Christian buildings, the eye does not always find strict frontiers. There are spaces, in the case of which the building functions as a frame; frame of something which is more important than the beauty of the material form.

In such cases one feels that the space is quasi-"in movement."

In music several possibilities exists for the expression of the "open" character. The most obvious way is improvisation. It had its first golden age in the Renaissance when the composers fixed only the framework of the instrumental piece. From such a point of view the composition has been open, and – considering the player as a sort of co-author – is sounded differently on every occasion.

"The same, and always different yet" – this has been the fundamental principle of the ornamental praxis as well as of the forms of variation in the later ages.

A composition is similarly open, when one part or detail is aleatoric, the real sound being a work of momentary inspiration of the players.

In music there exists another open form expressly composed, named *perpetuum mobile* when the written music material comes several times, unforeseen by the listener.

This time, by the repetition, which is the most important and fundamental structural technique of formation, the listener "gives up" orientating himself in the music material.

I would like to give an illustration to this: the composition by *György Kurtág* "Scenes from a novel" in 15 movements to the text of *Rimma Dalos*, evokes in our memory – as its title makes guess – the world of Bergman. Each of the sort verses suggests – originally in Russian – some kind of attitude to life. The title of the 12. Movement is: Sundays. It has the subtitle: *Perpetuum mobile*. Behold the musical form grown among its strict limits, has got the task to express infinity, by a special time-play (if I may lend such a merry word to the serious topic). The movement repeats the text: "That's another Sunday over. That means the next will come."

The building, as scene of the sounds, come into being, could lend an idea for musical constructions, too. An example for this is the St. Marc Church in Venice. The structure of double choruses of which the opportunity is offered to compose works for two choirs. The art of the two Gabrielis is a beautiful example of musicians catching this opportunity.

One could think: it is obvious that the place at disposition is suggesting *in advance* sonore possibilities. But there is a controversial example, too. In the *Basilica "Misericordia" of Sant'Elpidio a Mare* there are two symmetrical choruses, but there is only one with an organ. They have set an organ only in the 80's of our century in the other chorus, too.

There is an example, too, that the composer is forming a special sound-space on a simple concert platform, accostumed to traditional sonority. In Cavalieri's oratorio "*La rapprezentatione di anima e di corpo*", the two soloists (woman and man) represent the soul and the body, the inner and the outer appearance. According to this, the continuo instruments are positioned at given parts of the scene, clearly separating the two worlds from each other.

Such an "organization" of the concert platform can be found also in one of Bartók's works. *Ernő Lendvai* has pointed out that Bartók's inner hearing was a "stereo hearing", and that Bartók knew those principles well the phrasing of which the directors of modern opera recordings had undertaken only at the beginning of the 60's. It is a question of guessing the so-called sonic stage in the case of the *Music for Strings, Percussion and Celesta*. There is used a double string orchestra in the piece, and among them the group of the piano, harp and percussions took place. The space-effect of the right- and left-sided strings are polarizing the scene of music. The sonority itself is polarized, extending from the drums to the aetherian celesta.

There are *proportions* to find equally in the nature and different branches of art, in space and time. The most important of those is the "golden section." (An example in nature is the sea animal called *Nautilus* where in whichever direction we drew a line through the middle-point of the snail, the middle-point is in the goldensection of every point of intersection of the snail line. In these cases the whole of it is proportioned to the larger part, as the larger to the smaller.)

The 13th century scientist, Fibonacci examined the appearances of organic nature, his series covers the simplest, in whole numbers to express the golden section sequence where every number equals the sum of the two preceding numbers. One can mention several examples from the folra, such is the organization of spiral lines on the pine-cones.

The same proportion is at home in every branch of art: We can recognize the proportion of the oldest Greek symbols, "Pax" and "Beginning and End", the Alpha and Omega in one masterpiece of the Renaissance building (Figure 6).

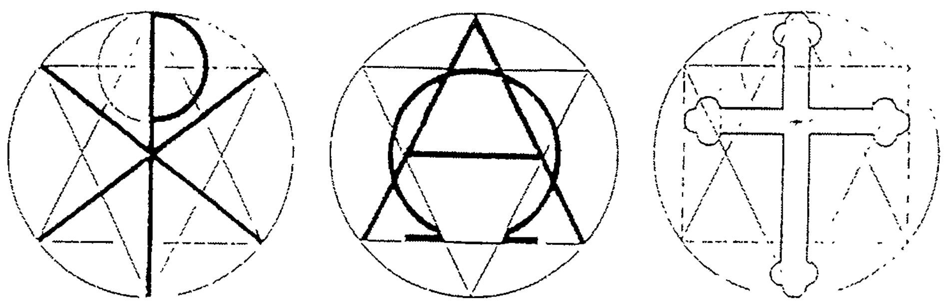


Figure 6

Here is the ground of the St. Peter Basilica (Figure 7).

The same proportion characterizes the Medici-sepulchre by *Michelangelo* and his painting *Madonna Doni*, where it is made easier to recognize the structure through the circular form of painting.

In the building art of our time it has been *Le Corbusier* who has applied most many sidedly the possibilities of the golden section, after having read the book by Manila Ghica, who published several books in the 20's about the connection of arts and mathematic or geometry. Le Corbusier has started from the proportions of the human body which show a double line of golden section. The title of his famous theoretical work is: Modulor.

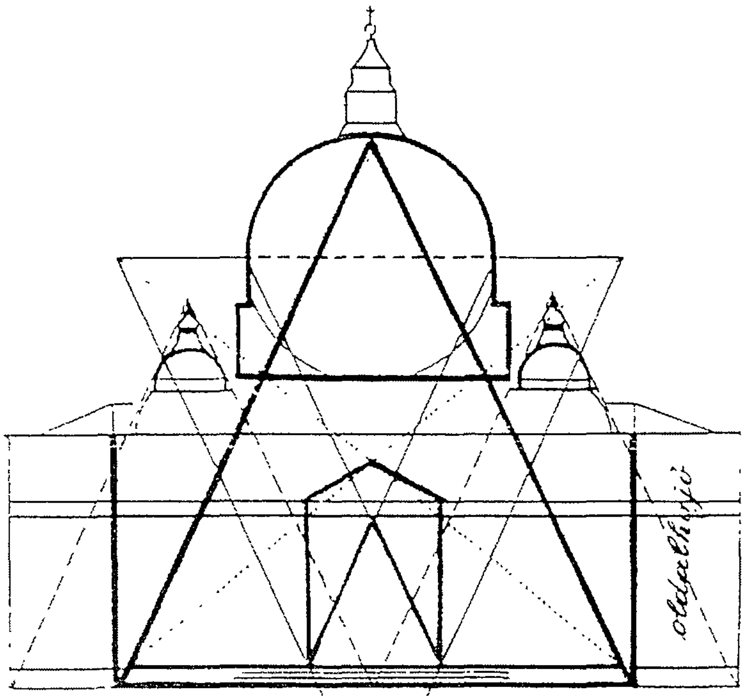


Figure 7

It was *Ernő Lendvai* who paid attention to the importance of the golden section proportion in the music, and examined this phenomenon first in Bartók's life-work.

The scientific results of E. Lendvai are interesting not only from the point of view of form. This importance is given by the fact that these proportions are always wearing a certain significance of *content.* At the points of golden section there always comes something important: either a turning point, or the highest moment of tension. As he continued these examinations, Lendvai has got similar results by the analysis of works by other composers, so it is obvious now: this connection also applies to the *timely* arts and timely proportions. Let me add that several composers have established consciously this system as a formal framework, under influence of Lendvai's researches published.

Now I'm showing a note page of Stones, an electroacoustic work by a Hungarian contemporary composer, János Decsényi. The lower line of the extravagant score signifies the time-axis. The uppermost line opens for us the secrets of several golden section-divisions, part-proportions, from the whole form down to the smallest form-cells; at the centre systems there are to find the realistically sounding tones or noises (Figure 8).

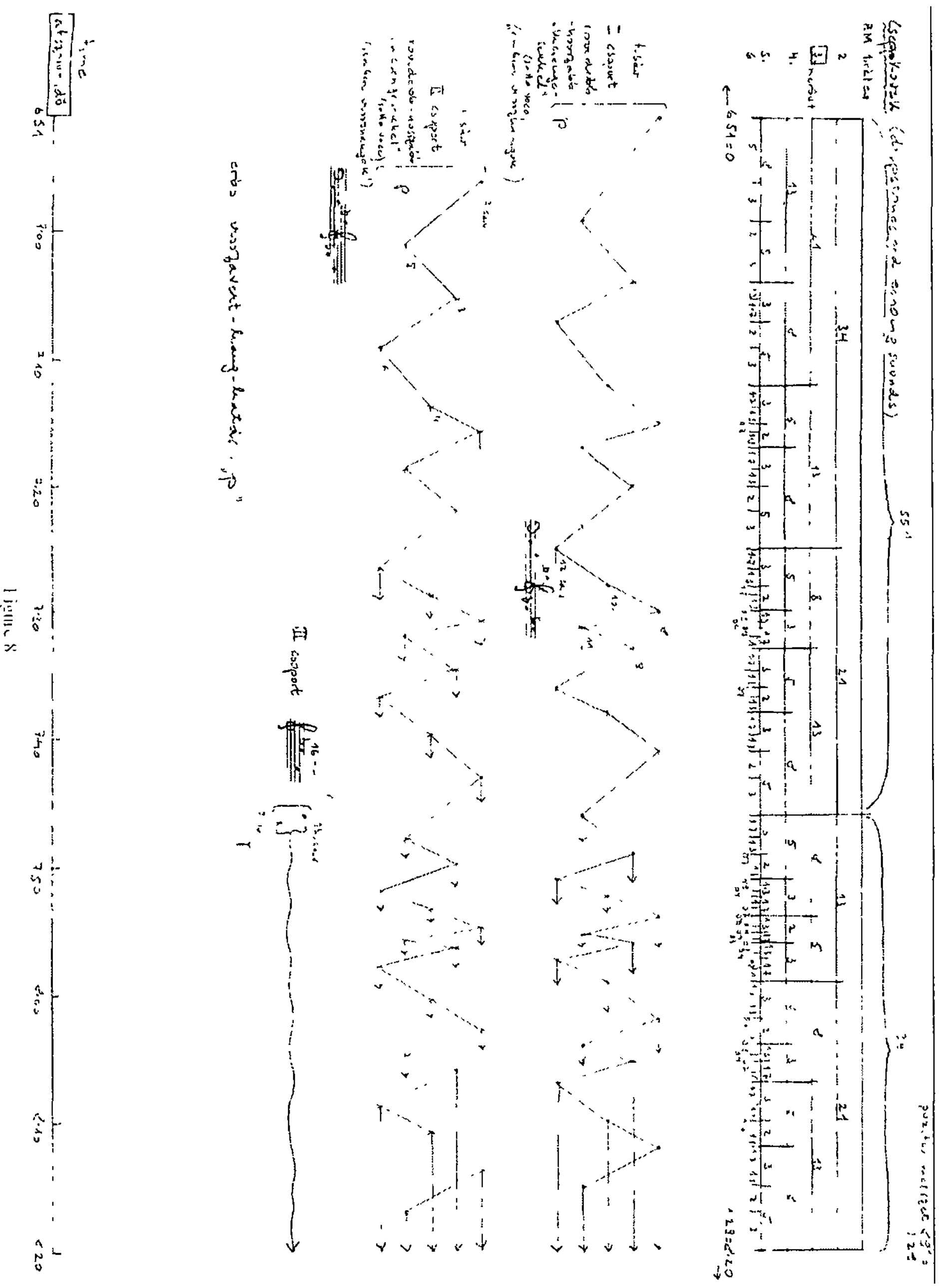


Figure 8

Till now I have spoken of an universal proportion, what holds its validity in a nearly infinitive circle. Now I would like to give a concrete example for a connection of old and new, between the spatial and timely composition.

The monumental motet *Nuper rosarum flores* by *Guillaume Dufay*, has been written 1436 for the inauguration of the Florence Cathedral. The knowers of Renaissance vocal-polyphony analyze it often. We can note that the same material sounds four times, by the tenor voices, at varying tempo, after the proportions 6 : 4 : 2 : 3. On the basis of this, *Howard M. Brown* defines the form of the motet as a free series of variations over a free isorhythmic canon. So, he registers the formal conception of the composition, and, what is more, evaluates this musicologically, stylistic, indicating that with the auditive conception of the constructional principle of numeric relations, it is reflecting the thought of the disappearing middle ages. Entertaining this interpretation, there is another one to mention here: namely, Dufay anticipated the spirit of the late popular puzzle-canons, as a *homo ludens.* For these proportions, which are to be found as constructional principles in many compositions anyway, by this time they overpass themselves in wearing symbolic meaning: they are timely appropriate of Brunelleschi's building. In David Fallows's interpretation: these proportions "correspond to the proportions of the nave, the crossing, the apse and the height of the cupola in the cathedral. That these two voices use the same melody – the Introit for the dedication of a church – at two different pitch levels and with interlocking rhythms itself symbolises the essence of Brunelleschi's structural feat, an inner shell and an outer shell with interlocking struts." (Fallows 1987)

And here is an example from the 20th century: There is a building, constructed after a concrete musical precedent, which has served later on as an impulse to the composition of another musical piece. This is the Philips Pavilion of the Brussels World Exhibition (1958). The architect of this pavilion is a composer and matematician, a close friend of Le Corbusier: *Iannis Xenakis*. He formed first the inner construction of the La Tourette Monastery and this work was one of the reasons, why he has got the opportunity to make the Philips Pavilion.

As Le Corbusier confirmed, the forms used by Xenakis had been *all* known before, Xenakis answered, acknowledging this fact: "I have applied the same right angles, surfaces and pillars which the architects use for thousands of years. The important thing is not whether they are new or not, but the way to applie them." (Varga 1980, [transl. Pándi, M.])

Figure 9 shows the first sketch of the pavilion.

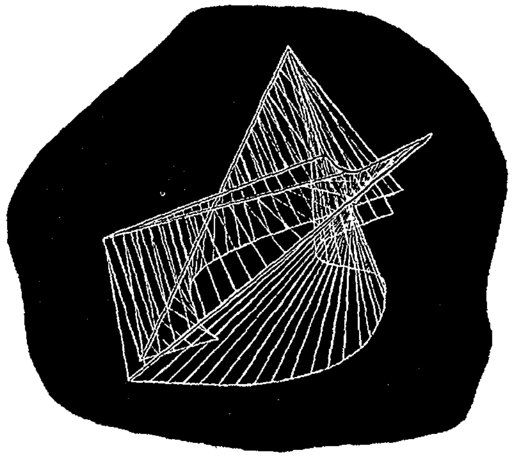


Figure 9

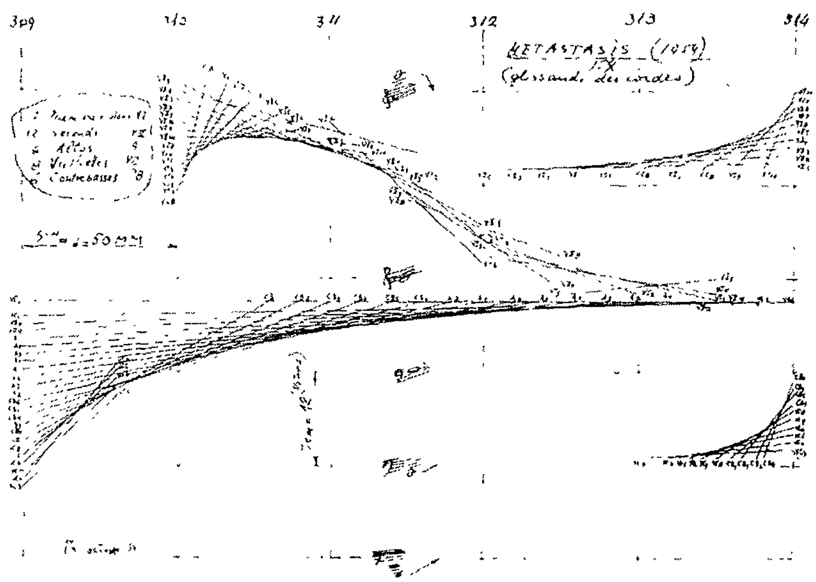


Figure 10

In the course of planning the Philips Pavilion, Xenakis himself has realized the same fundamental thoughts, as in 1953/54, at the composition of his musical piece, *Metastasis*. He was interested in the same thought: how to arrive from one point to the other, without beaking the continuity, both in the music and in the architecture. In Metastasis this aim resulted in the *glissando*, eliminating the abrupt, sudden change, in the pavilion in the hyperbolic paraboloid form.

After the picture of the pavilion, a detail from the music of Metastasis may follow. At the first sight it is obvious that in both cases there is the same question, the realization of the same idea.

There is another musical work connected with the Philips Pavilion, too, the Poeme electronique by *Varese*, which is imagined explicitly into this building by the composer.

Here, a thought of Xenakis is to be mentioned: "What is the straight line in the two-dimensional space? The constant changing of one dimension compared to the other. The same is happening in the relation of tone pitch/time: the straight line is the constant changeing of tone pitch in time. The difference between physical and musical space is that the former is homogeneous: both dimensions are lengths, distances. But in the music the natures of those two dimensions (tone pitch and time) are strangers to each other, they are connected only through the *possibility of organisation.*"

Dufay and *Xenakis*. Here are two examples picked out at random. How can we interpret the relation of archtectural and musical products? In Dufay's case the gesture is perhaps a mere metaphor. A proportion, a member of which has been related to a former one of different nature. This metaphor is supposing a strong affinity for symbols.

In Xenakis's case the aim was to realize adequate correspondence of a common phenomenon, thought, through different means.

The question emerges: are these special cases or is it worth looking for similar connections at other works, too?

In an age in which the existence of a special, uniform style is characteristic, there are many similarities to find, many common stilistic lines between the creations born near each other. The less we can consider it a musical common language, the more difficult it is to find individual correspondences among the works.

From this point of view Bartók's-life work seems to be in an exceptional position, as its theoretical-structural basis and the semantic concretisation of the phenomena has been recovered completely, thanks to the only one scientist: it is Ernő Lendvai.

In other cases the autoanalysis of the composer means a great help, although sometimes one mismanages the information at our disposal.

The cognition of the compositional technique of the New Vienna School would have required plenty of effort, if Schoenberg and his circle had not made public the process of their speculations, which has led at last to the serial technique. And what is the result? It is the doubt increased about the musicality of their music.

There is no standard method of salvation. Perhaps it would be desirable both composers and analysts to be multilaterally learned – if not in a polyhistoric meaning, but at least in sense of an interdisciplinar point of view. Then the analysis won't be meaningless, having the taste of paper, and the listener would not be obliged to be afraid, even in the case of excitingly individual, experimental works, to be victim of some "*épater le bourgeois*" that they abuse his naive interest – so he, who is looking for an artistic message; will lend his confidence to works, belonging to several types of organisation.

A REMARK

Goethe wrote in his *Farbenlehre* that colour is a creature of nature low, affecting distinctly the organ of the eye. *Anton Webern* transposed this definition at the terrain of music. As he said, music is a creature of natural law, connecting the organ of the ear.

But they are some people who are equally expert in several artistic branches. Besides that Arnold Schoenberg may be important from this point of view, E. T. A. Hoffmann, a writer, composer and painter in one person. He affirmed, that hearing is nothing else, but inner sight, and the sight is a more outside hearing.

The problematic of coloured hearing is a special question. Scriabin's thought, that a mystic music consists of light, can be traced back to old forerunners.

In the second part of Goethe's Faust, the sun is rising rumbling, and the celestial light is thundering like a trumpet.

Dante, speaking of the heavenly music in paradise, does not separate it from colours and light of the sky.

But this would lead us too far from our topic of „order – disorder".

REFERENCES

Fallows, D. (1987) *IC 0542*, Köln: EMI.
Lendvai, E. (1993) *Symmetries of Music*, Kecskemét: Kodály Institute.
Varga, B. A. (1980) *Beszélgetések Iannis Xenakisszal* (in Hungarian), Budapest: Zeneműkiadó, p. 155.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 231-247, 1998

THE MIRROR-SYMMETRY, THE FIBONACCI SERIES AND THE GOLDEN SECTION IN THE RENAISSANCE MASS COMPOSITION

Irina Guletsky

Address: Department of Musicology, Hebrew University, Jerusalem.

The perfect proportions counting, among others, the mirror-symmetry, the golden section and the Fibonacci series, that were newly discovered and held in especially great esteem at the Renaissance era, expressed themselves most vividly in architecture and visual arts. In the music of that period their reflection does not appear to be so obvious, which results from the specifics of the style and the expressive means of the epoch.

Let me remind you that in the strict style polyphony, the principle of a constant rejuvenation of the music texture and the latter's fluidity step forward and stipulate a more significant role of the process factor in the form-making. Any repetition – and it is a repetition that the appearance of symmetry is linked to – has a veiled character. It is especially true for the cult music and, specifically, the mass that is going to be the subject of this article.

At the same time, it is the presence of various forms of symmetry that provides for the architectonic form and permits to perceive a large cyclic work as a whole and clear composition.

It is necessary to note that, unlike their colleagues-architects, the music theoreticians of the Renaissance failed to leave any works or instructions concerning musical composition and its proportions. The theory of proportions that further develops the Pythagorean teaching on the link between the radiuses of the planet orbits and the structure of the musical scale referred, as it is known, to the field of the music pitch and the mensural rhythm. This theory penetrated architecture and literally reshaped architectural forms (Wittkower 1949, pp. 101-143). Whether it applied to the musical form and to what extent it was realized is open for speculation.

Only within the last four decades (since the late 50's) there appeared a number of research works on the subject. These are the works by M. van Crevel (1959, 1964), M. Henze (1968), N.W. Powell (1979) and others. True, they were all devoted to the analysis of structure and proportions of *cantus firmus* (c.f.) in the works of Franco-Flemish composers of the 15^{th} century. Let's recall that the structure of c.f. – the main thematic voice in the choral texture and the pivot of the composition – even though connected with the overall compositional proportions, cannot be identified with them. Besides, one ought not to ignore the following circumstance: at the early stage of its development – in the 14^{th} century – the mass was not yet a composition based on a single c.f., while on the pick of its development, in the 16^{th} century, it was already not.

Up to date, there exists not a single work devoted to the analysis of the overall compositional proportions and to the various ways of the symmetry manifestation in the Renaissance mass. This is exactly why these questions have become the purpose of my research.

In the present work I would like to stress the manifestation of the mirror-symmetry, Fibonacci series and the golden section in the overall mass structure that are expressed in the ratio of scales of parts and the whole. The overall structure is the first most significant compositional structure, sort of mathematical matrix, the most general plan, which the rest of the compositional structures depend upon to a certain extent. It is the overall structure that contains the chief principles of the compositional pattern. All the other structures – texture, ensemble plan, mode-cadence, polyphonic, etc. – are a kind of detalization of that basic plan of a composition. The higher the number of compositional structures that coincide with the overall structure in their proportional parameters, the clearer, the more crystal-like the musical form.

Before embarking on the analysis of concrete works, let us cast a look at the level of the musical form that is called *protoform*, as it is the protoform that contains, in a latent state, all these proportional ideas which will find their complete realization in the process of the genre development. The protoform is the most ancient and the most steady form level that appeared primarily and that can be defined as a form-idea. This stratum of the mass composition, in fact, gradually developed along the entire history of Christianity, together with the formation of Christian liturgy and as an integral part of the latter. The first known choral Ordinary that was found in the choir books of the St-Jaque monastery dates from 1254.

The protoform consists of seven structures that we are going to examine at length (see diagram 1).

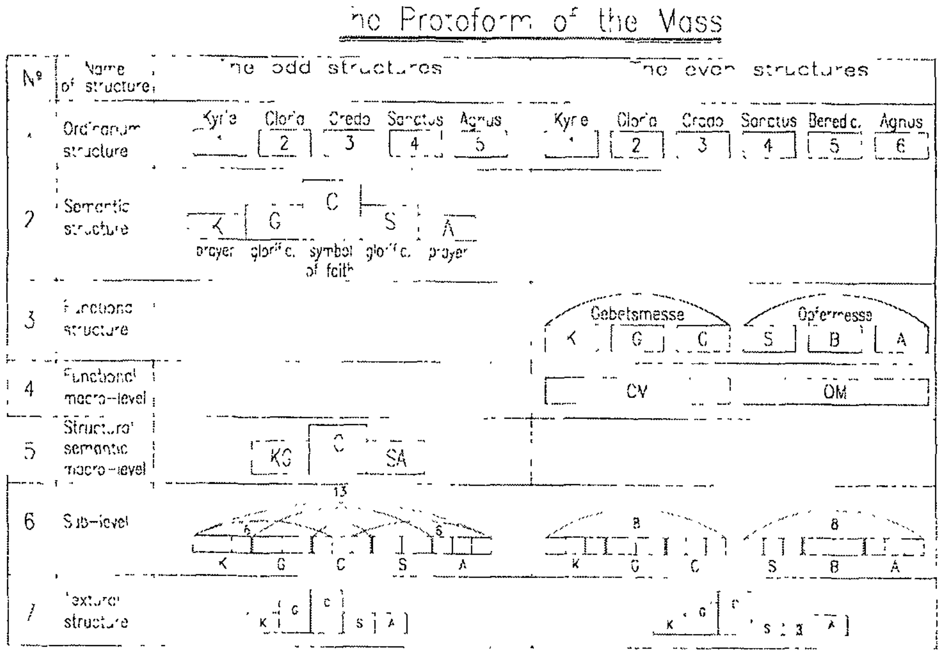


Diagram 1

The 1st structure: Ordinary (Ordinarium missae) is a musical cycle consisting of five parts (*Kyrie, Gloria, Credo, Sanctus, Agnus Dei*) and being a part of the Catholic office. In case *Benedictus* (a part of *Sanctus*) is treated as a separate part of the cycle, Ordinary will be a six-part composition. Besides Ordinary, the complete liturgy includes Proper (*Proprium Missae*), which consists of both musical and non-musical parts (e.g., readings, sermon, etc.). Unlike Ordinary, which is a permanent cycle, Proper varies dependent on concrete calendar days or event.

The 2nd structure - semantic - is defined by the contents of the verbal text, i.e., it reflects the meaningful functions of each part: *Kyrie* (K) and *Agnus* (A) - prayer, *Gloria* (G) and *Sanctus* (S) – glorification and gratification, *Credo* (C) – symbol of faith. It is important to note that this structure possesses a parameter of its own: it is hierarchic, for selfless glorification of the Lord (parts 2 and 4) stands higher as compared to the prayer (parts 1 and 5), while *Credo* is the climax, the central part of the entire liturgy.

The 3rd - functional - structure forms up as a result of the fact that the whole liturgy containing the five- or six-part Ordinary, falls into two functionally different parts: *Gebetsmesse* (GM) and *Opfermesse* (OM). *Gebetsmesse* is a prayer part of the liturgy, *Opfermesse* being an offering, or Eucharist. As a result, the first three parts of Ordinary

(*Kyrie*, *Gloria* and *Credo*) fall into *Gebetsmesse*, while *Sanctus* (*Benedictus*) and *Agnus Dei* - into the second part - *Opfermesse*. Thus occurs the unification of the parts of Ordinary into *two sub-cycles* with their respective titles and *the 4th structure* shapes up - *a two-part functional macro-level* of the form.

The formation of the *three-part macro-level - the 5th structure* - is also dictated by the distribution of the parts of Ordinary in the entire liturgy; however, it is not in the functional context already but rather in *the structural-semantic* one. What is meant is a specific order of placement of the parts of Proper and Ordinary. According to this order, *Kyrie* and *Gloria* sound directly one following the other; similarly follow each *Sanctus* and *Agnus Dei*. As far as the central part - *Credo* - is concerned, it is separated from the adjacent parts, *Gloria* and *Sanctus*, with a whole number of Proper parts.

Thus, *Kyrie* and *Gloria*, as well as *Sanctus* and *Agnus* join together forming macroparts. What appears as a result is a three-part macro-level, or a macro-structure, analogous to the semantic macro-structure (in the diagram they are united). The only difference is that the latter is hierarchic. *Credo* is singled out in it as the central, the main part of the liturgy.

The sub-level of the cycle is the 6th structure that results from the division of each part of Ordinary into sections. As we know, *Kyrie* and *Agnus Dei* were traditionally divided into three sections each; this tradition originated from the text of the Gregorian choral. *Sanctus* was traditionally divided into five sections, only two out of which being necessarily singled out: *Sanctus* and *Benedictus*. In *Gloria* and *Credo* a more liberal division was permitted, even though here too one can notice certain tendencies: in *Gloria* there were two strophes (*Et in terra* and *Qui tollis*) that were necessarily emphasized, while in *Credo* there were three (*Patrem*, *Crucifixus* and *Et in spiritum*) symbolizing the Holy Trinity. So, the protostructure of the sub-level contained 13 (and in cases of a more fractional division - 16) sections.

The last, *7th structure is textural.* It correlates with the number of syllables in each part and defines therefore the initially given proportions of the cycle.

As we can see, almost all the structures, with the exception of the textural one, are symmetric. Let us, for the sake of convenience only, call them odd and even. The odd structures are constructed hierarchically, the singled-out central part being the axis of the symmetry. In the even ones, the axis of the symmetry is imaginary and lies between *Credo* and *Sanctus*.

Such a division of the structures allows showing the sphere of influence of the two vectors of the musical form. Thus, in the odd structures, the vertical vector is emphasized which is responsible for the space-constructional aspect of the form, while in the even ones - the horizontal vector is emphasized, the one that reflects the time and process aspect of the composition. In this case it is possible to state the quantitative equality of the odd and even structures that testifies to a completely balanced protoform.

On the bases of the protoform one can distinguish five types of mirror-symmetry that will find their reflection in concrete compositions.

These are, first of all, the perfect types:

1) macro-symmetry – symmetry of a three-part macro-structure;
2) ordinary-symmetry – symmetry of the five-part Ordinary;
3) complete symmetry – symmetry of all three levels (the sub-level among them).

The imperfect types include:

4) partial symmetry – equality of only one pair (textural protostructure gives a sample of specifically this type of symmetry, as *Kyrie* = *Agnus Dei*, but *Gloria* does not equal *Sanctus*).
5) symmetry of ratios of parts – ratio of the scales of parts in form of proportions: K/G = A/S etc. (in the protoform these ratios are equal, as are the corresponding pairs, but in compositions this type of symmetry will be expressed only as equal ratios, for *Kyrie* will not equal *Agnus*, while *Gloria* will not equal *Sanctus*).

As has already been mentioned before, some musicologists have written about the presence of additive series and golden section in the structure of c.f. in the masses of the Franco-Flemish composers. However, in the view of the imperfect methodology of calculating the mensural proportions and in view of the veiled form of these progressions, no consensus on the subject has been reached thus far. In this respect, criticism is also being levelled against musicologists and art experts, who are accused of being carried away by the beauty of these operations and the search for them (Busse-Berger 1990, p. 92).

However, if one examines this question from a different point of view, turning first of all to the analysis of the protoform, the protostructure of the five-part Ordinary presents, at a closer glance, the beginning of the Fibonacci series – the first five numbers (diagram 2):

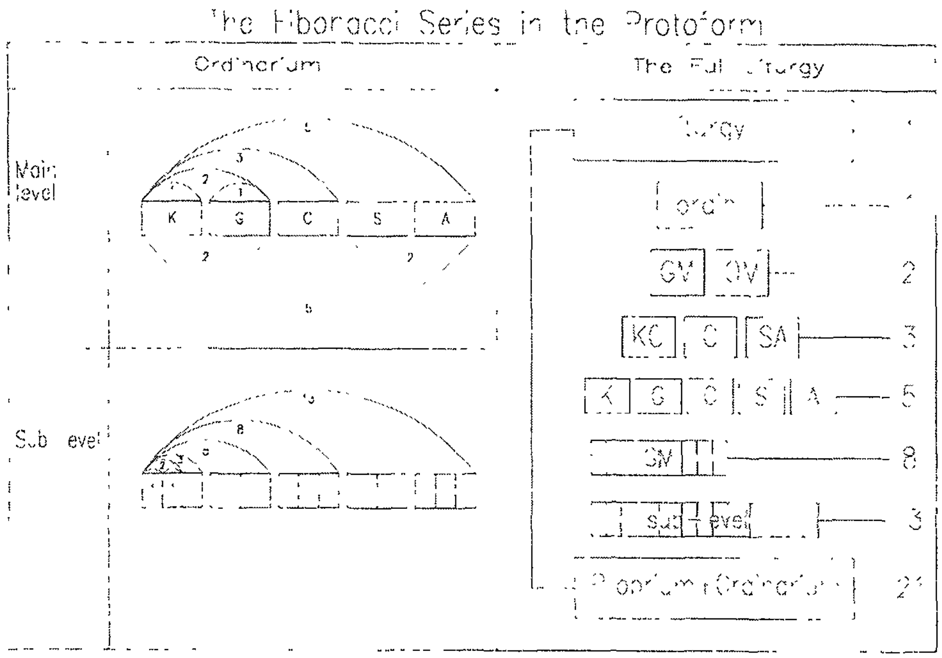


Diagram 2

1, 1, 2, 3, 5, ...

as well as the first two numbers of the Evangelist series, the latter being placed in any direction and order:

2, 5, ... or 5, 2, ... (as in the Gospels)

This kind of structural division of the cycle can be explained in the following way: the first and second parts of the mass *Kyrie* and *Gloria* as separate isolated parts are the ones. Unification of these two parts, as well as *Sanctus* and *Agnus* into macro-parts gives number 2, which in both cases acts as a monolithic structural unit. Number 3 means the unification of the first three parts into the first sub-cycle - *Gebetsmesse*. Number 5 corresponds to all the five parts of the cycle. This yields the sequence:

1, 1, 2, 3, 5.

The sub-level yielding a more fractional division of a five-part structure also creates series and continues the realization of this progression. Thus, the following number - 8 - corresponds to the number of sections on *Gebetsmesse*, while 13 is the number of sections in the entire mass. This forms up a succession of seven numbers:

1, 1, 2, 3, 5, 8, 13.

Another interpretation of this succession originates from the position of Ordinary in the liturgy: if the liturgical performance as a whole is taken for one, then Ordinary inside the liturgy is the second one. Further, 2 sub-cycles inside the mass, 3 macro-parts, 5 parts of Ordinary, 8 sections of *Gebetsmesse* (sub-level), 13 sections of the mass (a complete sub-level) and, finally, 21 - the complete number of parts in the whole.

Thus, the protoform is based on the multi-level progress of the additive series from the largest form units to the smaller ones, and vice versa.

The golden section, or a quantity close to it, appears in the protoform as a consequence of ratios of the neighbouring members of the series and usually as a cardinal ratio of the two sub-cycles: OM/GM. Correspondingly, in the compositional proportions the golden section may express itself as a steady correlation among the parts and their totals, as well as the focus of the composition, i.e., its climax zone that will fall on the final part of *Credo*.

Let us now examine concrete examples the manifestation of the overall compositional mirror-symmetry, Fibonacci series and the golden section.

The earliest type of symmetry of the overall structure that I managed to find is the symmetry of the macro-level in the *Barcelona mass* (the middle of the 14th century). The number of imperfect semibrevises (that are a common measure unit for all the parts) is more or less equal in the both macro-parts and the *Credo*. Therefore, the correlation of the scales of the macro-parts approximates one. Symmetry in such a case becomes stronger as a result of the similar modus in the macro-parts and contrasting - in the *Credo*. Symmetry of the macro-level is obviously projected without being fully realized in "*Notre Dame*" by G. de Machaut - the first authorized mass, composed probably in 1364. It is the only one, among the three complete masses of the 14th century that reached us, that demonstrates a clear and symmetrical structure of the sub-level (diagram 4-C).

In spite of the fact that the earliest masses of the 14th century can already be characterized as possessing a high level of symmetry of the overall structure, the 15th century presents but not so many samples where these symmetric patterns are being developed. Thus, for example, out of the ten masses by J. Ockeghem only two possess the macro-symmetry or the ordinary-symmetry; out of 21 Josquin's masses - 3, while G. Dufay and J. Obrecht do not have them at all. Cycles containing the imperfect symmetry are much more frequent. The earliest among them are, obviously, masses by A. Busnois - "*L'homme armé*" - and by Ockeghem - "*Sine nomine*", dating from the middle of the 15th century.

The real flowering of the symmetrical composition falls already on the period of a high and late Renaissance. In the productions of N. Gombert, a composer of the first half of the 16th century, the basic pattern is the symmetric one. However, it is only possible to talk about a full realization of the potentials of the mirror-symmetry contained in the protoform in connection with the masses of G.P. da Palestrina, the greatest Italian master of the second half of the 16th century. Suffice it to say that out of 104 masses of the composer, about 80% possess one or another type of symmetry and over half of them - one of the types of perfect symmetry.

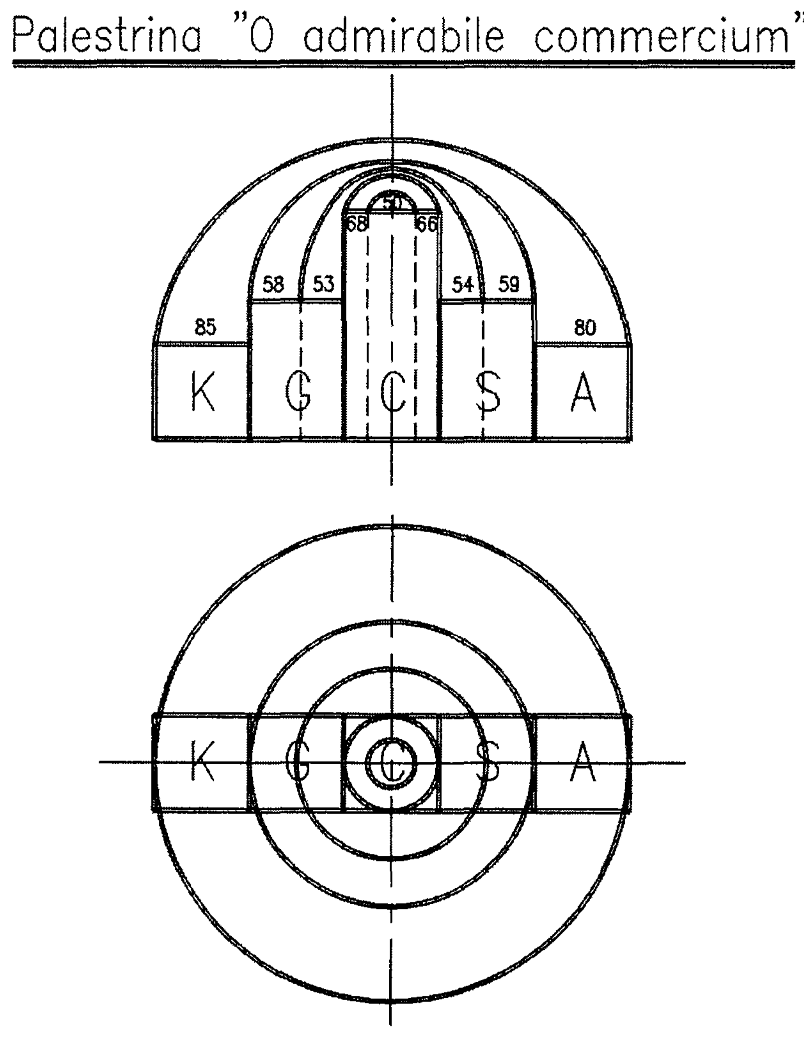


Diagram 3

One of the expressive examples of this type of construction can be found in the mass "*O admirabile commercium*" (diagram 3). As we can see, it is a perfectly symmetrical cycle wherein not only all the parts, but the sections too are corresponding. Equal number of sections is placed on each side of the central *Crucifixus*. In fact, it is only in Palestrina's mass that we find a fully symmetric construction that possesses a classic structure of the sub-level and, besides, is supported by a whole number of other compositional structures. Such a composition is, in the essence, a number of concentric circles. This is a qualitatively new stage in the development of the music form of the mass, its pick, and at the same time, its sunset.

The Fibonacci series find a realization in the cycles primarily in the structure of the sub-levels. The sub-level of any mass containing usually from 12 to 18 sections can be organized with the use of the Fibonacci progression, even though this is not always so. In diagram 4 we can see examples of the sub-levels of some masses consisting of 15, 16 and 18 sections. In them, the progressions are moving towards one another.

The Fibonacci Series in the Sub-levels of the Cycles

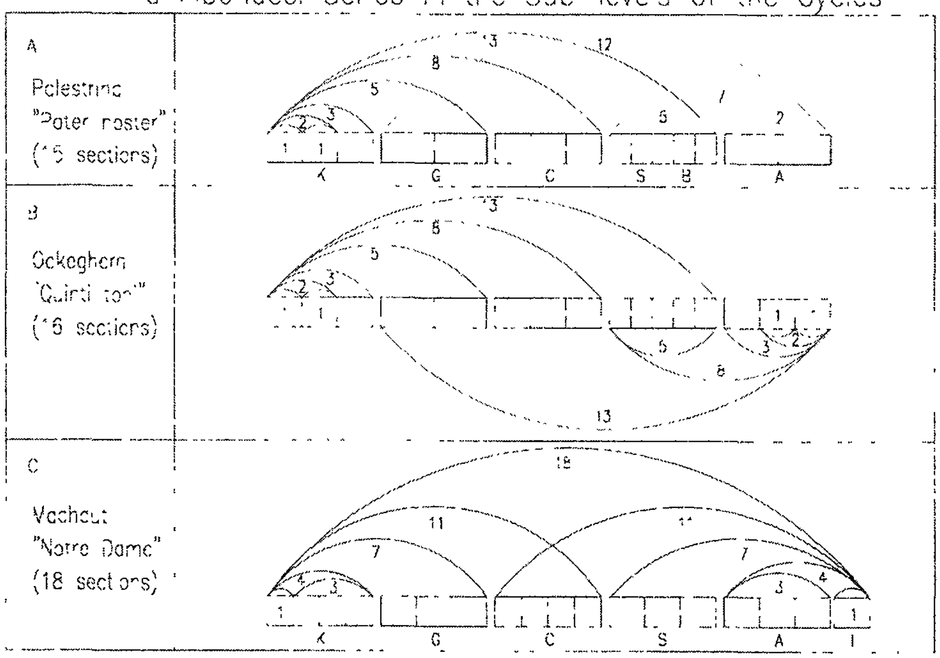


Diagram 4

Thus, if the Fibonacci progressions are programmed already in the protoform of the cycle and are *deliberately* introduced by the composer into the sub-level of the mass, it is hardly possible to suggest an "accidental" appearance of these series in the inner structure of the mass as well, in the c.f.; especially if one takes into account the fact that the hierarchy of precise calculations served as a sort of "ideal" basic plan of the would-be musical construction.

Along with the passing of the 15th century, the tradition of composing on the basis of c.f. disappears. The Palestrina's cycles, whose great majority has no c.f., possess a much more obvious form of progressions. Like the mirror-symmetry, they appear first of all in the sub-level structures and in the overall structure proportions. The sub-level structure of the Palestrina's mass possesses from 12 to 16 sections and is always constructed with the use of progressions.

Quantitative exponents of the scales of parts and their totals in the cycles of the Roman master frequently form additive series as well. Thus, for example, in the mass "*Nasce la gioia mia*" this series appears as follows:

	K	G	C	G+C	G+C+S+A
(bars)	67	106	173	279	448

And in the mass "*O virgo simul et Mater*" the sequence is even more precise:

K	G	C	G+C	M (the entire Ordinary)
70	124	194	318	512

As for the mass "*Salve Regina*" a progression shows up both in the direct order and backwards:

		⇒		
K	G	C	C+S	G+C+S+A
80	138	220	362	581

		⇐		
K+G+C+S	G+C	C	S	A
580	358	220	142	81

A most interesting sample of a form construction with the use of recurrent series, particularly the Evangelist series, can be found in the mass "*Inviolata*" (diagram 5). The cycle consists of 14 sections (7:7), a form of segmentation which, as noted above, recreates the second Evangelist sequence. At the same time, in the division of the mass into sections one can discover the first series too (diagram 5-A).

The number of bars in the parts and their sums also form the first Evangelist progression (diagram 5-B).

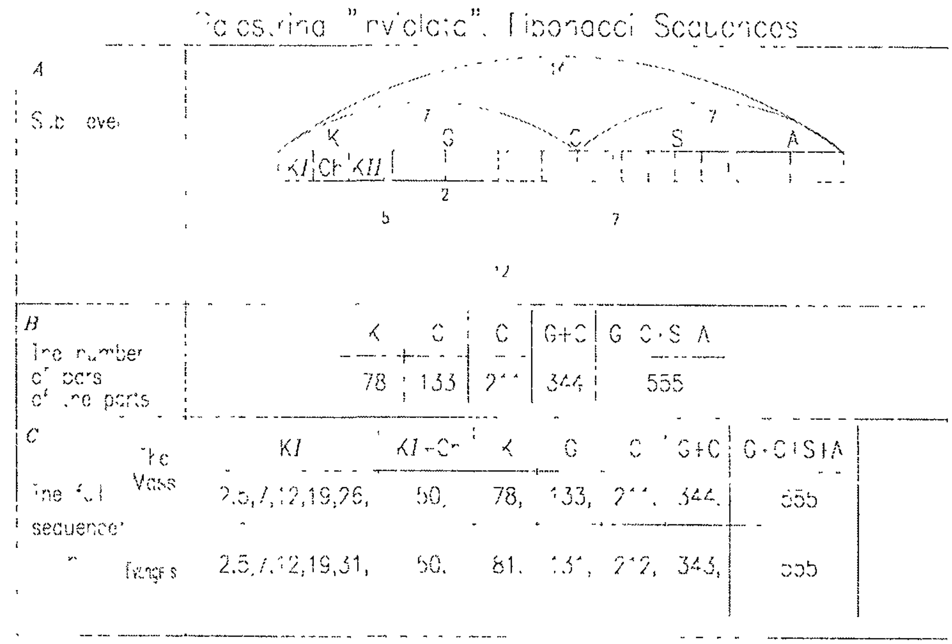


Diagram 5

As we can see, the sequence appears starting with the eighth member of the series, deviations being minimal. However, if we analyse the construction of the first part of *Kyrie –Kyrie I* we shall discover the first seven "missing" numbers of the sequence as well. Thus, the first section - *Kyrie 1* (KI) is constructed in the following way: in the initial 2 measures the tenor leads the theme (one voice sounds); then cantus joins and up until the *5*th bar included two voices sound; then the bass adds up and up until the *7*th measure included - three voices sound; starting with the *12*th bar starts a new theme (a full four voices); and finally, from the *19*th bar until the very end of the section, which is the 26th bar, there follows a conclusion of the second theme.

Thus, in *Kyrie 1* the following numbers fully corresponding to the Evangelist series are emphasized: 2, 5, 7, 12, 19.

Further on, the total of first two sections of *Kyrie – Kyrie 1 + Christe = 50* measures (this is the seventh member of the progression). We get a series of 12 numbers (with slightest deviations in a few cases – diagram 5-C).

MISSA: INVIOLATA

4 vocum.

As we can see, the entire mass is an unfolded progression embracing the whole of its cycle from the first until the very last measure. This resembles a mighty spring – a spiral half of whose coils (6 out of 12) appears already in the first section of *Kyrie*. In this case, by purely mathematical means is stressed the role of this initial section of the composition, its mighty energy potential as the one of a "mustard seed" out of which the "tree of life" grows.

One may note in the musical compositions of that period, a significant role of the golden sections which, being a consequence of the very same additive series, themselves form "golden sequences", as N. W. Powell (1979) puts it, in the c.f. structure.

Apparently, far from accidental is the fact that one of the earliest types of proportionality of the overall mass structure that I have found is connected with the golden section, which coordinates the duration of the cycle parts and their totals. These are the three masses "*L'homme armé*" by Busnois, Ockeghem and Dufay that appeared in the middle of the 15th century, between 1450 and 1460.

In the cycle of Busnois, the spiral of the golden section develops as follows:

K	G	OM	GM
A	K+G	GM	M
0.63	0.64	0.65	0.6

The same formula, with a slight deviation in numerical values, appears in Dufay's mass:

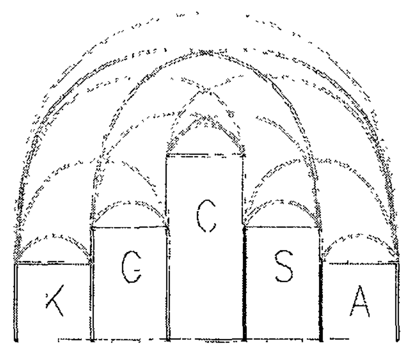
K	G	OM	GM
A	K+G	GM	M
0.618	0.66	0.67	0.6

A more extended version takes place in Ockeghem's mass:

A	S	K+G	G	K+G	C	OM	GM
S	C	OM	K+G	C	GM	GM	M
0.62	0.66	0.614	0.65	0.66	0.6	0.65	0.6

Proportions of the overall structure of Palestrina's masses are long series of golden sections, containing sometimes up to 18 members (diagram 6).

Palestrina
Mass "Pater noster"
(golden section)



$$\frac{K}{G}=\frac{G}{C}=\frac{S}{C}=\frac{A}{S}$$

$$\frac{G}{K+G}=\frac{C}{G+C}=\frac{C}{C+S}=\frac{S}{S+A}$$

$$\frac{K+G}{G+C}=\frac{K+G}{C+S}=\frac{S+A}{G+C}=\frac{S+A}{C+S}$$

$$\frac{G+C}{K+G+C+S}=\frac{G+C}{G+C+S+A}=\frac{C+S}{K+G+C+S}=\frac{C+S}{G+C+S+A}$$

$$\frac{K+G+C}{M}=\frac{C+S+A}{V}\qquad\boxed{=0.618}$$

Diagram 6

The most significant expression of the golden section proportion is in the climax zone of a composition, which is usually found in the closing section of *Credo*. In the mass by G. de Machaut, this is an unfolded solemn section *Amen* that is fulfilled in virtuous techniques of *goket*, as opposed to all the previous sections. In the mass "*Papae Marcelli*" by Palestrina, the closing section of *Credo*, starting with the 186th bar, there forms up a mighty culmination with a bell-like effect, which is due to the canonic sequences and imitations in all the voices.

Thus, the meaning of the perfect proportions inherent in the mass composition can be treated as follows.

Due to the mirror-symmetry the mass construction is analogous to the one of a cathedral as well as to the composition of an altar. Essentially, symmetry in the liturgical music formed up under a direct influence of the church architecture. Borrowing from architecture and further developing the laws of the mirror-symmetry, music assumes an obvious architectonic character, reflecting in its composition the main principles of a church building – its proportionality and multi-layer hierarchy with the centre in the dome.

So, the mirror-symmetry determines the space-constructional aspect of the form, which dominates in the five-part Ordinary.

The Fibonacci progressions reflect the unfolding of the form in time, i.e., a process. The semantics of these progressions in the masses is a hint at the evangelic episode of "the miracle with bread and fish" carrying a very important semantic message: the satiation with the spiritual food - the teaching of Christ. Besides, already in the text of the Gospels the expression of a miracle via progression symbolizes the Divine act of creation – the appearance of multitudes from One, a doctrine, which dates back to Plato's "*Timaeus*". Therefore, the Fibonacci series are interconnected with the six-part Ordinary expressing the time aspect of the form (let us recall the Six Days of Creation).

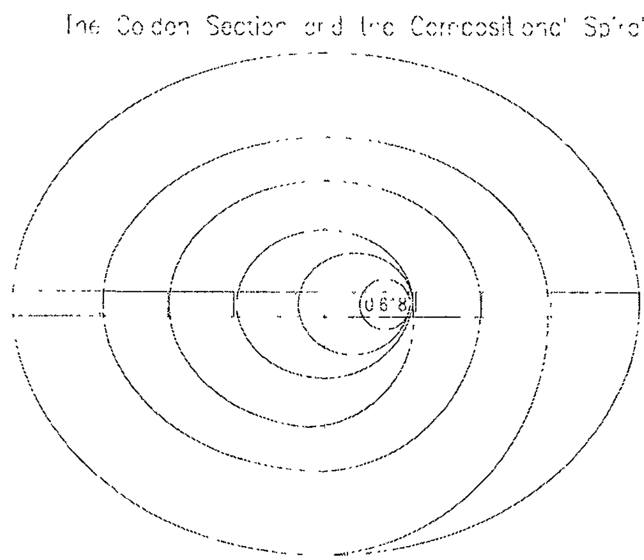


Figure 7

The five- or six-part Ordinary is one and the same form wherein the golden section unifies both vectors: vertical and horizontal, i.e., space and time. Being the centre of the broken symmetry and, consequently, of the highest tension, the golden section becomes the focus of the culmination zone, or the centre of the folding and unfolding of the compositional spiral (diagram 7). Analogous spirals appear in the famous *tondoes* by Botticelli and Raphael. This form is nothing else but a reflection of neo-platonics' most important statement that was formulated by N. Kuzanus as "*explicatio-complicatio*", i.e., the process of unfolding and folding of the world within the Allmighty.

Therefore, one can draw a symbolic parallel: Creator and His creation - the world; a Composer and his creation - the mass (a model of this world). While the three perfect proportions that are present in the mass in a complete unity are a mathematical symbol of Divine perfection and the threefold of Divine substance.

Botticelli: *The Madonna of the Magnificat*, Uffizi, Florence

REFERENCES

Aaron, P (1969) *Thoscanello de la musica,* New York.

Alberti, L.B. (1985) *L'architettura.* Bologna.

Atlas, A.W (1987) "Gematria, Marriage Numbers and Golden Sections in Dufay's 'Resveillez vous'", *AM,* vol. 59, pp. 111-126.

Bossuyt, I. (1994) *De Vlaamse Polyphonie,* Leuven

Busse Berger, A M (1990) "Musical proportions and arithmetic in the late Middle Ages and Renaissance," *MD,* 44, pp. 89-118

Crevel, M. van. (1959, 1964) *Jacobus Obrecht: Opera Omnia, Missae,* vi and vii, ed. by M. van Crevel, Amsterdam: Vereniging voor Nederlandse Musiekgeschiedenis.

Elders, W (1967) *Studien zur Symbolik in der Muzik der alten Niederlander,* Bilthoven.

Fellerer, K G. (1982) *Palestrina-Studien,* Baden-Baden.

Henze, M. (1968) *Studien zu den Messenkompositionen J. Ockeghems,* Berlin.

The Oxford Dictionary of Music (1994), p. 556.

Pacioli, L. (1974) "Divina Proportione" in *Quellenschriften fur Kunstgeschichte und Kunsttechnik des Mittelalters und der Neuzeit,* vol. 2, transl. and ed. by Winterberg, C., New York.

Powell, N W. Fibonacci and the Golden Mean: Rabbits, Rumbas and Rondcoux, *JMT,* vol 23, no. 2, pp. 227-273.

Sandersky, M.V. (1981) "The Golden Section in Three Byzantine Motets of Dufay," *JMT,* vol. 25, no 2, pp 291-306.

Schmidt-Görg J. (1968) *History of the Mass,* Köln.

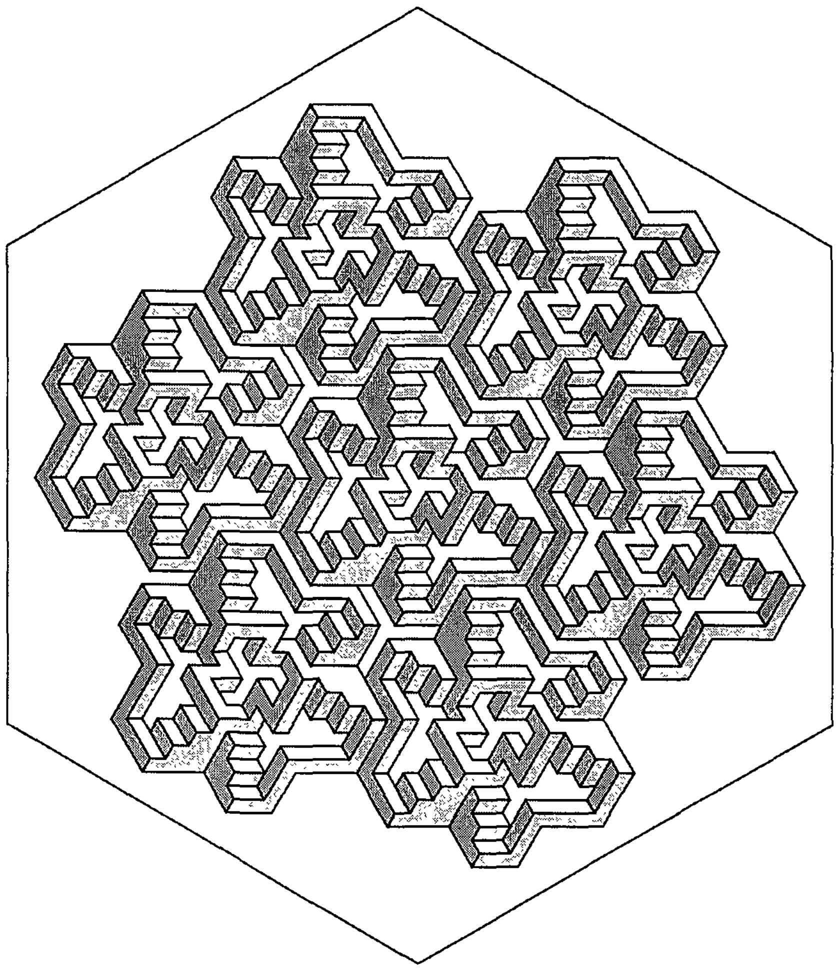
Tinctoris, J. (1979) *Proportions in Music,* New York.

Wagner, P (1913) *Geschichte der Messe,* Bd. 1 (bis 1600), Leipzig.

Warren, C.L. (1973) "Brunelleschi's Dome and Dufay's Motet," *MQ,* vol. 59, pp. 92-105.

Wittkower, R (1949) *Architectural Principles in the Age of Humanism,* New York.

Zarlino, G (1965) *Le istitutioni harmoniche,* New York.



Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 249-256, 1998

COMPUTATIONAL ASPECTS IN IMAGE ANALYSIS OF SYMMETRY AND OF ITS PERCEPTION

Hagit Hel-Or[(1)] and David Avnir[(2)]

Address: [(1)] Dept of Math and Computer Science, Bar-Ilan University, Ramat-Gan, 52900 Israel E-mail. hagit@cs.biu.ac.il.
[(2)] Department of Organic Chemistry, The Hebrew University of Jerusalem, Jerusalem, 91904 Israel. E-mail: david@chem.ch.huji.ac.il.

Abstract: *Symmetry as a characteristic of shape and form has been widely studied both in the artistic and esthetic aspect on one hand and in the mathematical and computational aspect on the other. Symmetry is typically viewed as a discrete feature: an object is either symmetric or non-symmetric. However visual perception and natural behavior and phenomena treat symmetry as a continuous feature, relating to statements such as "one object is more symmetric than another" or "an object is more mirror symmetric than rotational symmetric". With this notion in mind, we view symmetry as a continuous feature and define a Continuous Symmetry Measure (CSM) to quantify the "amount" of symmetry of different shapes and the "amount" of different symmetries of a single shape. The characteristics of the CSM provides the ability to analyze the symmetries of images, shapes and 3D objects while taking into account the hierarchical and continuous nature of these symmetries and allowing for noise in the data such as occlusion and fuzzy data.*

1 INTRODUCTION

Symmetry is one of the basic features of shape and form. It has been widely studied from various aspects ranging from artistic to mathematical. However the classical view of symmetry as a binary feature – either an object is symmetric or it is not - is inadequate to describe the symmetries found in the natural world. It is inconsistent with visual perception and natural behavior which treat symmetry as a continuous feature, relating to statements such as "one object is *more* symmetric than another" or "an object is *more* mirror symmetric than rotational symmetric". With this notion in mind, we view symmetry as a continuous feature. Accordingly, a Continuous Symmetry Measure (CSM) has been defined [11, 3] which quantifies the "amount" of symmetry of different shapes and the "amount" of different symmetries of a single shape.

2 THE CONTINUOUS SYMMETRY MEASURE (CSM)

We define the Continuous Symmetry Measure (CSM) as the minimum 'effort' required to transform a given shape into a symmetric shape. This 'effort' is measured by the mean of the squared distances each point is moved from its location in the original shape to its location in the symmetric shape. Note that no priori symmetric reference shape is assumed.

A shape P is represented by a sequence of n points $\{P_i\}_{i=0}^{n-1}$. We define a distance between every two shapes P and Q:

$$d(P,Q) = d(\{P_i\},\{Q_i\}) = \frac{1}{n}\sum_{i=1}^{n} ||P_i - Q_i||^2$$

We define the *Symmetry Transform* $\hat{P}$ of P as the symmetric shape closest to P in terms of the distance d. The *Continuous Symmetry Measure* of P denoted $S(P)$ is defined as the distance to the closest symmetric shape:

$$S(P) = d(P; \hat{P})$$

The CSM of a shape $P = \{P_i\}_{i=0}^{n-1}$ is evaluated by finding the symmetry transform $\hat{P}$ of P and computing: $S(P) = \frac{1}{n}\sum_{i=0}^{n-1} ||P_i - \hat{P}_i||^2$. This definition of the CSM implicitly implies invariance to rotation and translation. Normalization of the original shape prior to the transformation additionally allows invariance to scale.

A geometrical algorithm was developed to find the Symmetry Transform and the CSM of a shape (see [3, 11, 6] for details).

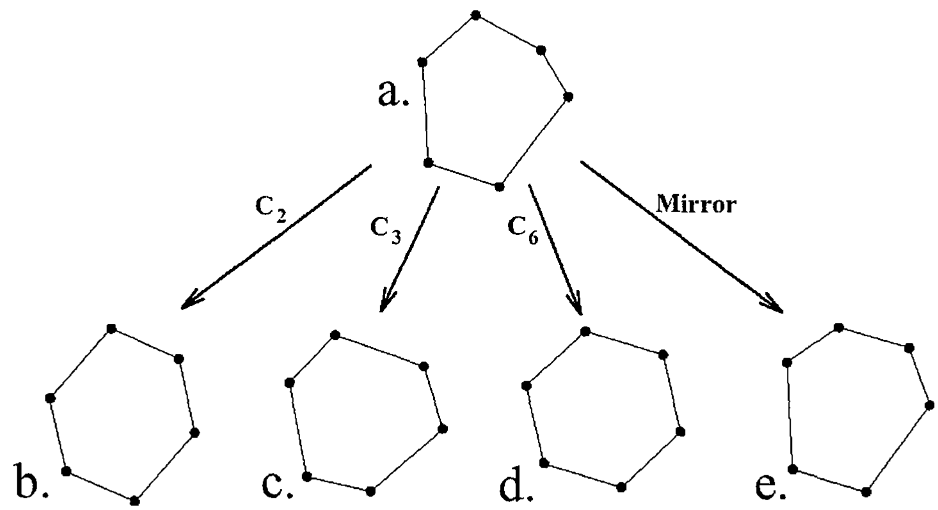
The general definition of the CSM enables evaluation of a given shape for different types of symmetries (mirror-symmetries, rotational symmetries, etc.). Moreover, this generalization allows comparisons between the different symmetry types, and allows expressions such as "a shape is more mirror-symmetric than C_2-symmetric."

An additional feature of the CSM is that we obtain the symmetric shape which is 'closest' to the given one, enabling visual evaluation of the CSM.

The CSM approach to measuring symmetry allows the hierarchical nature of symmetry to be expressed and quantified, as will be discussed below. Additionally, the CSM method can deal with noisy and occluded data, also discussed below.

3 MEASURING THE CSM OF SHAPES, IMAGES AND 3D OBJECTS

The versatility of the CSM method has induced its use in various fields such as Chemistry [9, 5], Psychology [4], Archaeology [2] and more. Underlying these studies is the ability of the CSM to evaluate continuous symmetry in shapes, images and objects. Previous approaches to evaluating symmetry in shapes, images and objects which mainly rely on the binary concept of symmetry are reviewed in [11, 3].



a) 2D polygon and its Symmetry Transform with respect to b)
b) C_2-symmetry (SD = 1.87)
c) C_3-symmetry (SD = 1.64)
d) C_6-symmetry (SD = 2.53)
e) Mirror-symmetry (SD = 0.66)

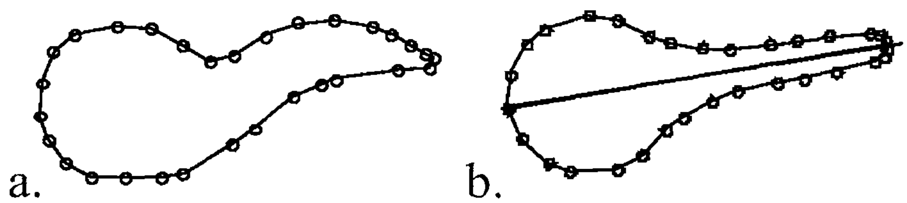
Figure 1: Symmetry Transforms of a 2D polygon.

3.1 CSM of a Set of Points

Given a set of points, possibly with connectivities between these points, the CSM is calculated using the geometrical algorithm previously developed [11, 3]. Given a shape such as a polygon, the vertices can be used as the input to the algorithms. Figure 1 shows an example of a shape and its Symmetry Transforms and CSM values with respect to several types of symmetries.

3.2 CSM of a Continuous Shape

Measuring the CSM of a general shape, requires the pre-process of representing the shape by a collection of points. Several methods have been suggested to sample the continuous contour of general shapes in order to obtain a collection of points as input to the CSM algorithm [8, 11]. Figure 2 shows an example of a general shape whose contour has been sampled, a collection of representative points obtained and the Symmetry Transforms and CSM values were calculated with respect to several types of symmetries.



a) The contour of the shape is sampled.
b) The Symmetry Transform of the shape is found.

Figure 2: Symmetry Transform of a general 2D shape

3.3 CSM of Gray-scale Images

Two approaches have been used in applying the CSM to gray-scale images. One approach segments areas of interest from the image and regards their contour as 2D shapes (see Figure 8 for example).

Another approach in dealing with images, lets pixel values denote elevation, and considers an image as a 3D object on which 3D symmetries can be measured. Figure 3 shows a range image and a gray-scale image for which the 3D mirror symmetry transform was computed, the 3D reflection plane was found and the 3D object rotated to a frontal vertical view.

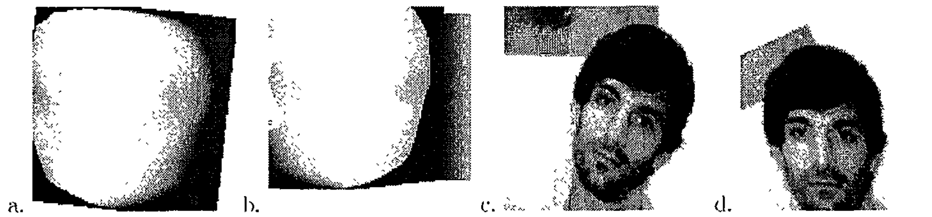


Figure 3: Applying CSM with respect to 3D mirror-symmetry to find orientation of a 3D object. a) Original range image. c) Original Gray-scale image. b,d) The symmetry reflection plane was found and the object in the image rotated to a frontal vertical view.

3.4 CSM of 3D Objects

Measuring the CSM of 3D objects is straightforward given a collection of 3D points as representatives [11, 3, 7, 9]. A variation of this scheme involves measuring the symmetry of the projection of a 3D object onto the image plane [12]. This approach is used in reconstruction of 3D objects from their 3D projections. Figure 4 shows a 3D object and its symmetry transform with respect to T_d-symmetry (tetrahedral). Figure 5 shows an example of 3 projections of a 3D object which was reconstructed using the CSM approach.

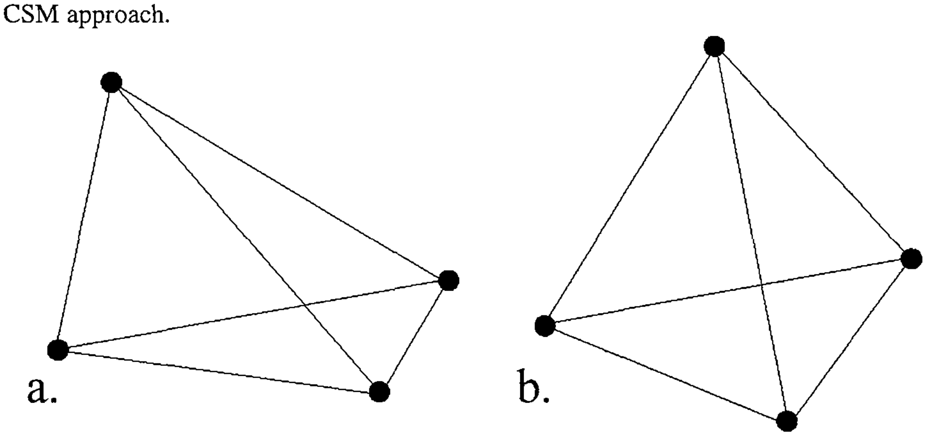


Figure 4: A 3D-object (a) and its symmetry transform (b) with respect to Tetrahedral-symmetry

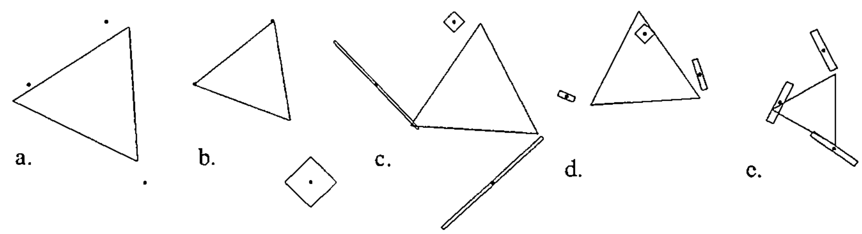
Figure 5: Reconstruction of a 3D-mirror-symmetric object from 2D images of different viewpoints.

4 EXPLOITING THE CHARACTERISTICS OF THE CSM

The CSM approach to measuring symmetry can be embedded in a scheme that takes advantage of the multiresolution characteristics of symmetry The symmetry transform of a low resolution version of an image is used to evaluate the symmetry transform of the high resolution version. This technique was used in estimating face orientation [11,7].

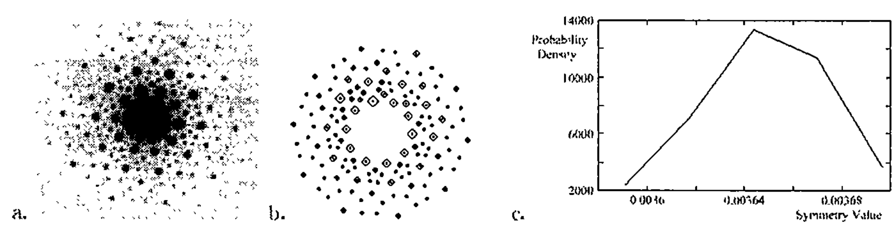
In many cases the source data is noisy. The CSM method can be exploited to deal with noisy and missing data. Considering noisy data, where the collection of representative points are given as probability distributions, the CSM approach can evaluate the following properties [11, 10]:

- The most probable symmetric configuration represented by the points.
- The probability distribution of CSM values for the points.



a-c) Changing the uncertainty (s.t.d.) of the measurements
d-e) Changing both the uncertainty and the expected location of the measurements.

Figure 6: The most probable C_3-symmetric shape for a set of measurements after varying the probability distribution and expected locations of the measurements.



a) Interference pattern of crystals. b) Probability distribution of point locations corresponding to a. c) Probabilty distribution of symmetry distance values with respect to C_{10}-symmetry was evaluated. Expectation value = 0.003663

Figure 7: Probability distribution of symmetry values

Figure 6 shows the effect of varying the probability distribution of the object points on the resulting symmetric shape.

Figure 7 displays a fuzzy image of points (a Laue photograph which is an interference pattern created by projecting X-ray beams onto crystals). and the probability distribution of the CSM values obtained for the pattern.

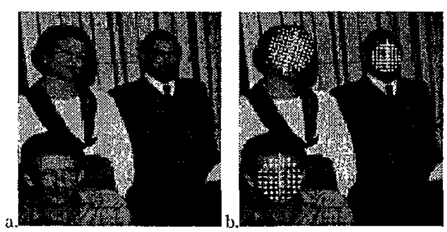
Finally, the CSM approach to measuring symmetry allows the hierarchical nature of symmetry to be expressed and quantified thus, global and local features can be evaluated for their symmetry content. Figure 8 shows an example where the CSM approach measures local mirror symmetry to find faces in an image. This issue of global vs. local symmetry is important in the case of measuring symmetry of human faces.

5 SYMMETRY OF HUMAN FACES

The human face is a complex structure comprising of a number of facial features. It can be modeled at several levels of hierarchy: at the top level is the global structure of the face (the face contour) and at lower levels facial details (such as facial features) are revealed. Accordingly, symmetry of the human face should be considered hierarchically. The CSM approach to measuring symmetry can be employed to evaluate mirror-symmetry of a face at several different levels (Figure 9):

- Face Contour - symmetry is evaluated by considering the contour of the face alone.
- Facial Features Centroids - the centroids of each of the facial features serve as the set of input points for symmetry evaluation.
- Facial Features Contours - the face contour and the contours of all the facial features are considered in the evaluation of symmetry.

Recent studies have proposed that there is a positive correlation between symmetry of faces and physical attraction [1]. However the methods used for evaluating symmetry do not capture the complexity of this characteristic. We propose that using the CSM approach in a hierarchical scheme will provide a more flexible reliable and meaningful measure of symmetry of human faces.



a) Original image. b) Faces are found as locally symmetric regions.

Figure 8: Multiple mirror-symmetric regions in images.

a) Face Contour. b) Facial Features Centroids c) Facial Features Contours

Figure 9: Measuring symmetry of human faces at 3 levels:

6 CONCLUSION

We view symmetry as a continuous feature and adopt the Continuous Symmetry Measure (CSM) to evaluate it. The CSM can evaluate any symmetry in any dimension and has been applied to shapes, images and 3D objects. The CSM can deal with noisy data, such as fuzzy and occluded data. The method of evaluating symmetry using the CSM can be applied to global and local symmetries. This can be extended to deal with symmetry in a hierarchical manner as in the case of measuring symmetry of human faces. The CSM is currently used in numerous fields including Chemistry, Physics, Archaeology, Botany and more.

REFERENCES

[1] Grammer, K. and Thornhill, R. (1994) Human facial attractiveness and sexual selection The role of symmetry and averageness, *Journal of Comparative Psychology* 108, pp. 233-242.

[2] Saragusti, I., Sharon, I, Katzenelson, O. and Avnir, D. (1997) Quantitative analysis of the symmetry of artifacts, *J. Archeol. Sci.*, submitted 1997.

[3] Zabrodsky, H. (1993) *Computational Aspects of Pattern Characterization - Continuous Symmetry*, PhD thesis, Hebrew University, Jerusalem, Israel.

[4] Zabrodsky, H. and Algom, D. (1994) Continuous symmetry: A model for human figural perception, *Spatial Vision* 8(4), pp. 455-467.

[5] Zabrodsky, H. and Avnir, D (1993) Measuring symmetry in structural chemistry, in Hargittai, I., ed *Advanced Molecular Structure Research*, volume 1, Greenwich, CT: JAI Press

[6] Zabrodsky, H and Avnir, D. (1995) Continuous symmetry measures, iv: Chirality, *J Am. Chem. Soc.* 117, pp. 462-473

[7] Zabrodsky, H. and Peleg, S and Avnir, D. (1992) *Hierarchical symmetry. In International Conference on Pattern Recognition, volume C: Image, Speech, and Signal Analysis*, pp. 9-12, The Hague.

[8] Zabrodsky, H., Peleg, S. and Avnir, D. (1992) A measure of symmetry based on shape similarity, in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 703-706, Champaign, June 1992.

[9] Zabrodsky, H., Peleg, S and Avnir, D. (1993) Continuous symmetry measures II: Symmetry groups and the tetrahedron, *J Am. Chem. Soc.* 115, pp. 8278-8289.

[10] Zabrodsky, H., Peleg, S. and Avnir, D. (1994) Symmetry of fuzzy data, in *International Conference on Pattern Recognition*, pp. 499-504, Tel-Aviv, Israel, Oct.

[11] Zabrodsky, H., Peleg, S. and Avnir, D. (1995) Symmetry as a continuous feature, *IEEE Trans. on Pattern Analysis and Machine Intelligence* 17(12), pp. 1154-1166.

[12] Zabrodsky, H. and Weinshall, D. (1997) Using bilateral symmetry to improve 3D reconstruction from image sequences, *Computer Vision and Image Understanding* 67(1), pp. 48-57.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 257-268, 1998

ICOSAHEDRAL B_{12} ARRANGEMENTS IN YB_{66} AND $YB_{41}Si_{1.2}$

Iwami Higashi(1) — Toshio Ishii(2) —
Takaho Tanaka(3) — Kimiko Kobayashi(4)

Address. (1) and (2) Department of Chemistry, Chiba Institute of Technology, Shibazono, Narashino, Chiba 275-0023, Japan
(3) National Institute for Research in Inorganic Materials, Tsukuba 305-0044, Japan
(4) The Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-0198, Japan

Abstract: *Icosahedral* B_{12} *structures of boron-rich solids show unusual and fine appearances which originate in icosahedral structural entities, and are quite different from conventional atomic configurations which are widely seen in most of other inorganic compounds. This paper describes the structures of yttrium-containing boron-rich solids,* YB_{66} *and* $YB_{41}Si_{1.2}$, *which are unique even among icosahedral* B_{12} *crystals in that they are respectively constituted of the* $B_{12}\cdot(B_{12})_{12}$ *giant-icosahedron and a similar giant structural unit,* $B_{12}\cdot(B_{12})_8(B_{12}Si_3)_4$. *The* $(B_{12}Si_3)_4$ *polyhedral unit within the latter is an icosihexahedron (a polyhedron with 26 faces) discovered in the* $YB_{41}Si_{1.2}$ *compound. By describing the two giant structural units with a large circle, the complicated structures of these boron-rich compounds are presented in an understandable way. Comparison of packing densities of* B_{12} *icosahedra and the giant-icosahedra in these two compounds are made by defining some hypothetical diameters for the structural units.*

1. INTRODUCTION

Boron, borides and related compounds differ from conventional solids in that it is impossible to interpret their structural features in terms of conventional rule of valence.

As a result, these materials manifest a number of unique properties which in many cases are of potential technological importance giving rise to a growing interest in their physical and chemical properties. In a previous paper (Higashi et al. 1995), we have presented structural arrangements of icosahedral B_{12} structures of some boron-rich solids. That is because they show unusual and fine structures that have originated in icosahedral structural entities, and are quite different from conventional atomic configurations which are widely seen in most of other inorganic compounds.

Although a great number of icosahedral B_{12} crystals have thus far been reported, they can be classified into eight types according to the modes of icosahedral B_{12} arrangements (Higashi 1986). It is impossible to utilize five-fold rotation symmetry in a two- or three-dimensional periodic network, and thus three dimensional arrangements of B_{12} icosahedra form open although rigid three dimensional networks in real boron-rich solids. Except α-rhombohedral boron, therefore, icosahedral B_{12} structures need additional boron polyhedral units, such as B_{28}, B_{22}, and B_{20} and isolated boron atoms to fill the openings within the icosahedral B_{12} frameworks.

In the present paper, we describe the structures of yttrium-containing boron-rich compounds YB_{66} and $YB_{41}Si_{1.2}$, which are unique even among icosahedral B_{12} crystals in that they are respectively constituted of so-called giant-icosahedron $B_{12}\cdot(B_{12})_{12}$ and a similar giant structural entity $B_{12}\cdot(B_{12})_8(B_{12}Si_3)_4$.

2. DESCRIPTION OF STRUCTURES

2.1 Structural units

An icosahedral polyhedron shows various outer appearances according to the direction along which it is seen. Therefore, to facilitate understanding the structural figures in this paper, the projections of icosahedron along the two-fold (a), three-fold (b), and five-fold axes are presented in Figure 1. The projection (d) in Figure 1 shows a directional nature of the linkage between boron icosahedra; it is mostly effected approximately along the five-fold axis. In Figure 2, a new structural unit $B_{12}Si_3$ is projected along the *a*, *b* and *c* axes of the $YB_{41}Si_{1.2}$ structure in which the $B_{12}Si_3$ unit occupies an important position as a structural unit. As in the icosahedron, this unit has a mirror plane or bisecting plane, on which three Si sites are located.

The $B_{12}\cdot(B_{12})_{12}$ giant-icosahedron and the similar structural unit $B_{12}(B_{12})_8(B_{12}Si_3)_4$ are projected along their two-fold axes in Figure 3 and Figure 4, respectively. In each figure, the central icosahedron was surrounded with twelve icosahedra (Figure 3) or eight icosahedra plus four $B_{12}Si_3$ units (Figure 4), showing that the mirror plane or bisecting plane of the unit is just on the projection plane.

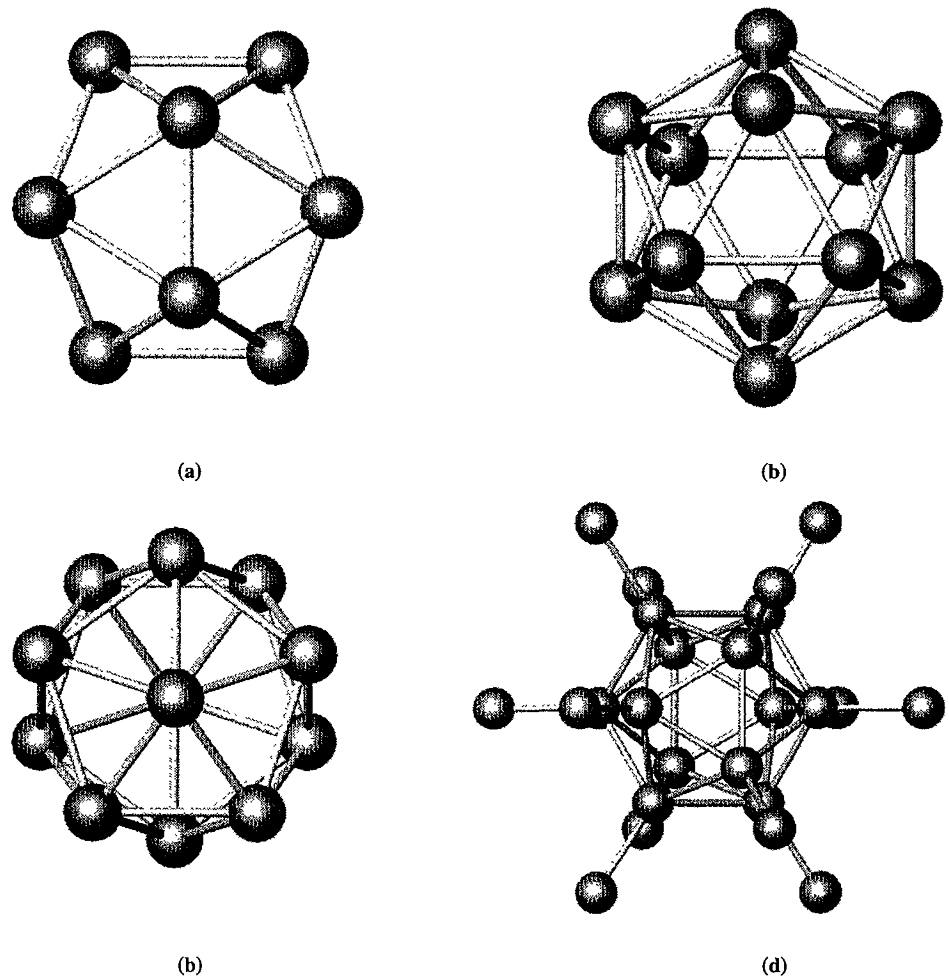


Figure 1: The icosahedral B_{12} unit as seen along the two-fold axis (a), three-fold axis (b), and five-fold axis (c). Figure (d) demonstrates a directional nature of linkages between the B_{12} icosahedron and neighboring boron structural units; the linkages are almost always effected along its five-fold axes.

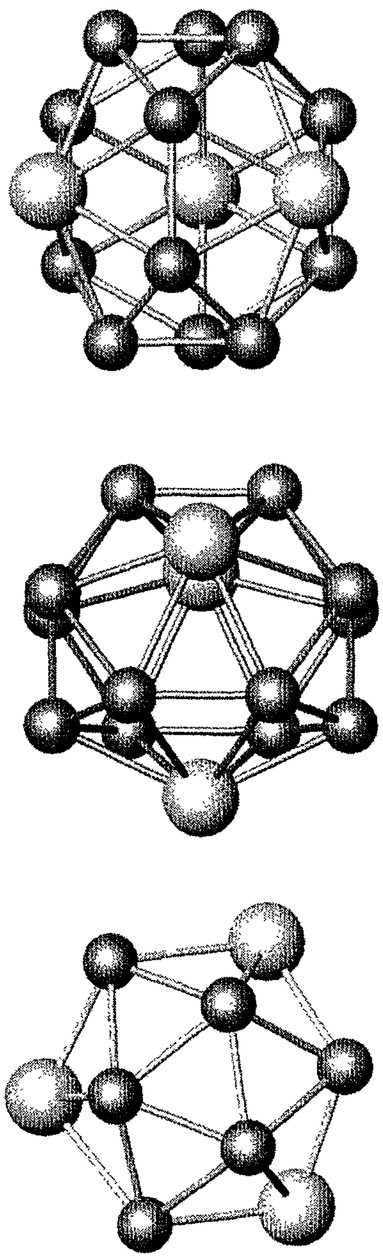


Figure 2(a)(b)(c): The $B_{12}Si_3$ icosihexahedron (a polyhedron with 26 faces) projected along the *a* axis (a), *b* axis (b) and *c* axis (c) of the $YB_{41}Si_{1.2}$ structure in which the $B_{12}Si_3$ unit occupies an important position as a structural unit

2.2 Structure of YB_{66}

The YB_{66} compound belongs to the face-centered cubic system (space group: Fm3c) with the lattice constant $a = 2.34364(6)$ *nm* (Richards and Kasper, 1969). The structure of this compound is basically constructed with the $B_{12}\cdot(B_{12})_{12}$ giant-icosahedron (Figure 3), which is located in one orientation at the face-centered cubic lattice-point (Figure 5). It also occurs at the centers of the cell and the cell edges rotated by 90° (Figure 5). In consequence, there are eight $B_{12}\cdot(B_{12})_{12}$ units (1248 boron atoms) in one unit cell. The relatively large hole at the center of each octant is filled with an irregularly shaped boron cluster consisting of 42 atoms, and smaller holes are filled with *Y* atoms. (In this paper, descriptions of the boron clusters and the distribution of *Y* atoms are omitted.)

In Figure 6, the arrangement of giant-icosahedra in the YB_{66} structure is presented by replacing the B_{12} icosahedral unit with a large circle. (In Figure 5, the B_{12} unit is depicted with twelve small circles.) The linkages between neighboring giant-icosahedra are formed by the inter-icosahedral B_{12}-B_{12} bonds approximately along the fivefold axes (Figure 1(d), Figure 5). The opening at the center of each octant of the unit cell is filled with the irregularly shaped cluster mentioned above.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 257-268, 1998

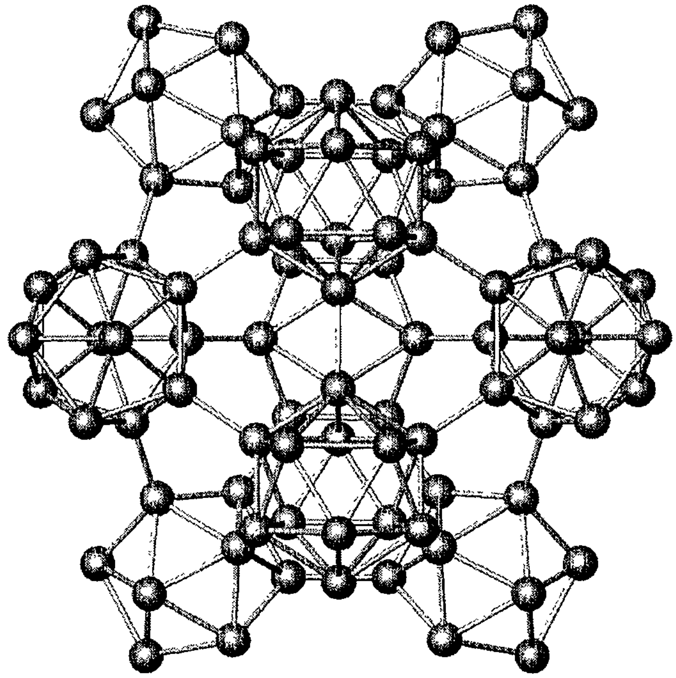


Figure 3: The B_{12}*$(B_{12})_{12}$ giant-icosahedron as seen along its two-fold axis.

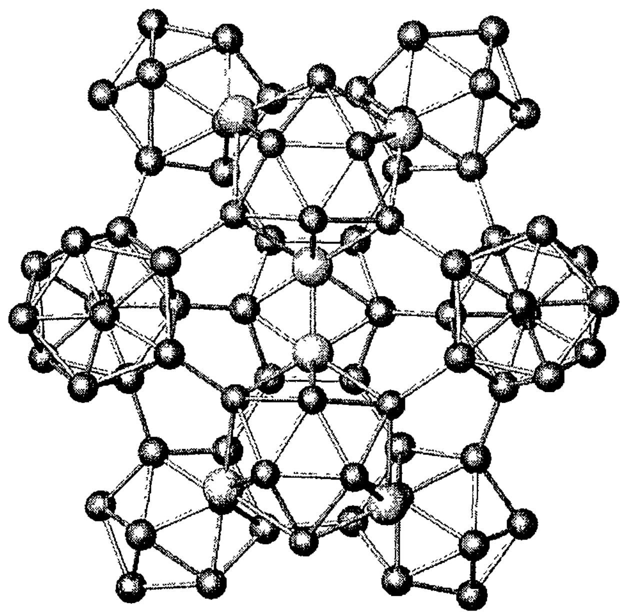


Figure 4: The B_{12}*$(B_{12})_8(B_{12}Si_3)_4$ giant-icosahedron as seen along its two-fold axis.

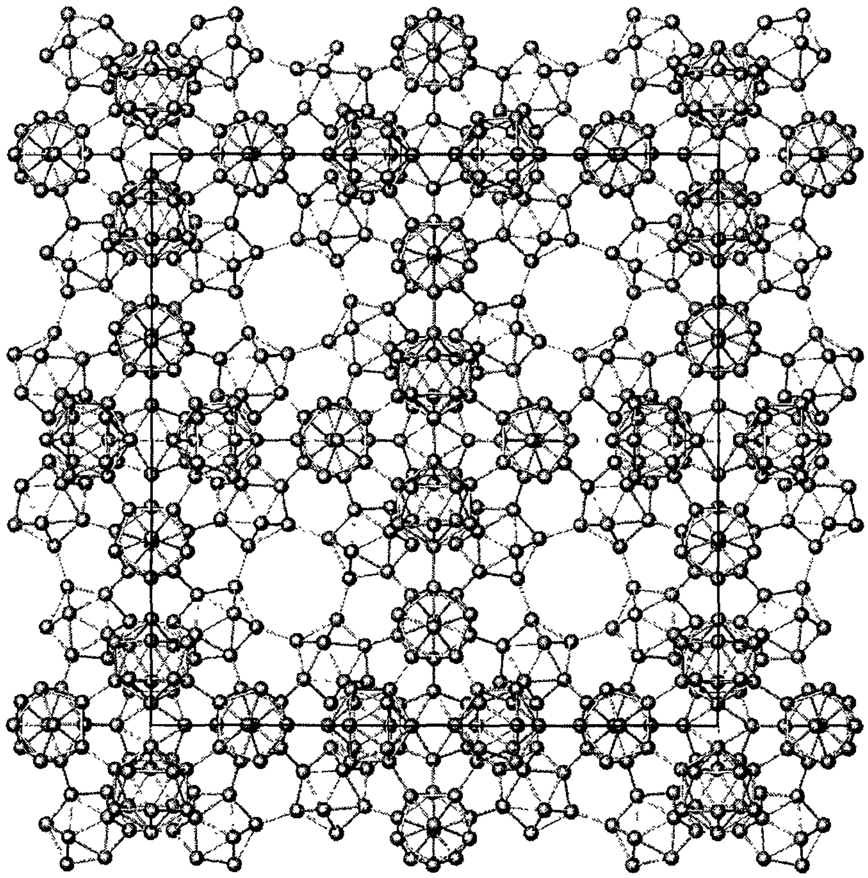


Figure 5: Arrangement of the $B_{12}{\cdot}(B_{12})_{12}$ giant-icosahedron (or B_{12} icosahedron) in the YB_{66} structure. In this figure, boron atoms are depicted with a small circle.

2.3 Structure of $YB_{41}Si_{1.2}$

$YB_{41}Si_{1.2}$ is a new compound with a new structure-type, which have recently been prepared and the structure determined (Tanaka et al. 1997, Higashi et al. 1997). The structure of this compound is made up of the B_{12} icosahedron and $B_{12}Si_3$ icosihexahedron (a polyhedron with 26 faces) (Figure 2). The crystal belongs to the orthorhombic system (space group: *Pbam*) with the lattice constants a = 1.6674(1), b = 1.7667(1), and c = 0.9511(7) *nm*. By analogy with the YB_{66}-structure, the complicated boron-framework can be interpreted by use of the $B_{12}{\cdot}(B_{12})_8(B_{12}Si_3)_4$ giant-icosahedron (Figure 4), which corresponds to the $B_{12}{\cdot}(B_{12})_{12}$ giant-icosahedron in YB_{66}.

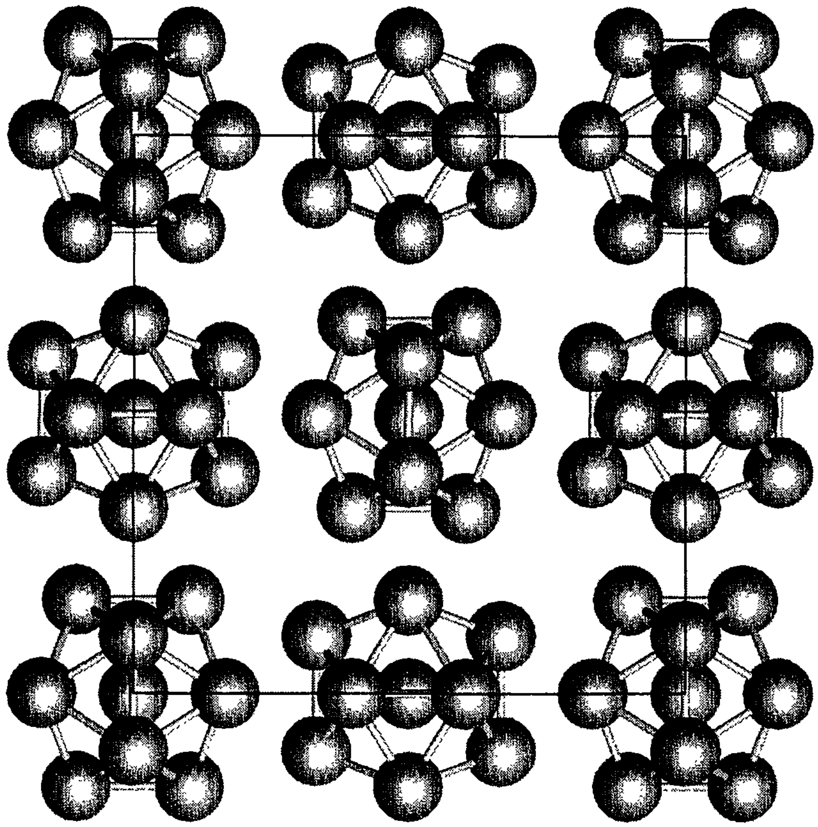


Figure 6: Arrangement of the $B_{12}\cdot(B_{12})_{12}$ giant-icosahedron (or B_{12} icosahedron) as seen along the *a* axis of the YB_{66} structure. In this figure, each icosahedron is depicteded with a large circle.

Figure 7 shows the arrangement of the $B_{12}\cdot(B_{12})_8(B_{12}Si_3)_4$ giant-icosahedra as seen along the *c* axis. In this figure, five B_{12} icosahedra within the giant icosahedron, which are projected along their respective two-fold axis (as seen in Figure 1(a)), are placing their bisecting plane on the crystallographic mirror plane at z = 0. On the other hand, the $B_{12}Si_3$ units, involved in the giant icosahedron, are having their bisecting plane on the mirror plane at z = 0.5. The two large openings being bisected respectively with the (100) and (010) planes are filed with icosahedral B_{12} pairs shown in the Figure 8. It is of interest to note that the intericosahedral bond of B_{12}-B_{12} pair is quite different from the conventional B_{12}-B_{12} bond which is formed approximately along the five-fold axis; the bond of icosahedral pair is effected by putting one edge of each icosahedron in parallel to each other, forming an approximately square plane which lies on a plane bisecting both icosahedra.

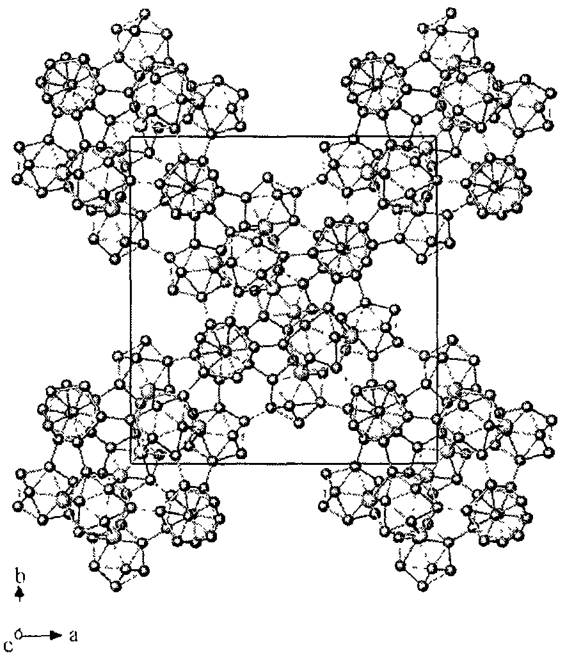


Figure 7: Arrangement of the $B_{12}{}^{\bullet}(B_{12})_8(B_{12}Si_3)_4$ giant-icosahedron as seen along the c axis of the $YB_{41}Si_{1.2}$ structure. In this figure, as in Figure 5, the boron atoms are described with a small circle, and the silicon atoms with a little larger circle.

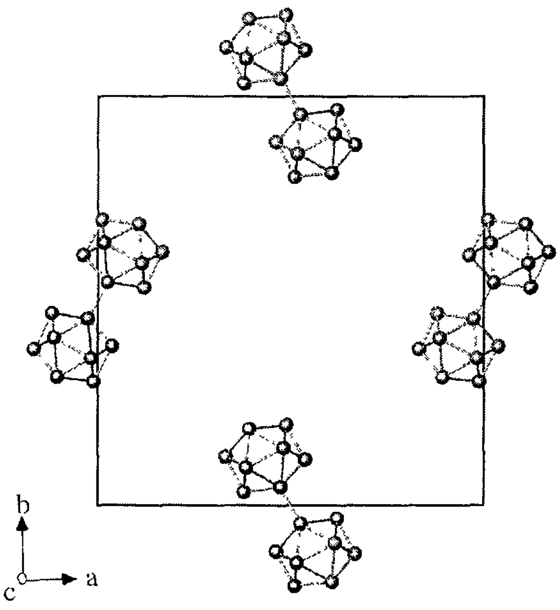


Figure 8: The B_{12}-B_{12} pair which occupies a large hole within the $B_{12}{}^{\bullet}(B_{12})_8(B_{12}Si_3)_4$ giant-icosahedral network (see Figure 7).

In Figure 9, as in the case of the YB_{66}, the arrangement of the $B_{12}\cdot(B_{12})_8(B_{12}Si_3)_4$ giant-icosahedra in the $YB_{41}Si_{1.2}$ structure is presented by replacing the B_{12} and $B_{12}Si_3$ units with large circles. As seen from the unit cell indicated with a rectangle drawn in thick line, the arrangement of the giant-icosahedra is crystallographically quite different from that in YB_{66}. The sequences of the giant units are, however, similar to that in YB_{66}.

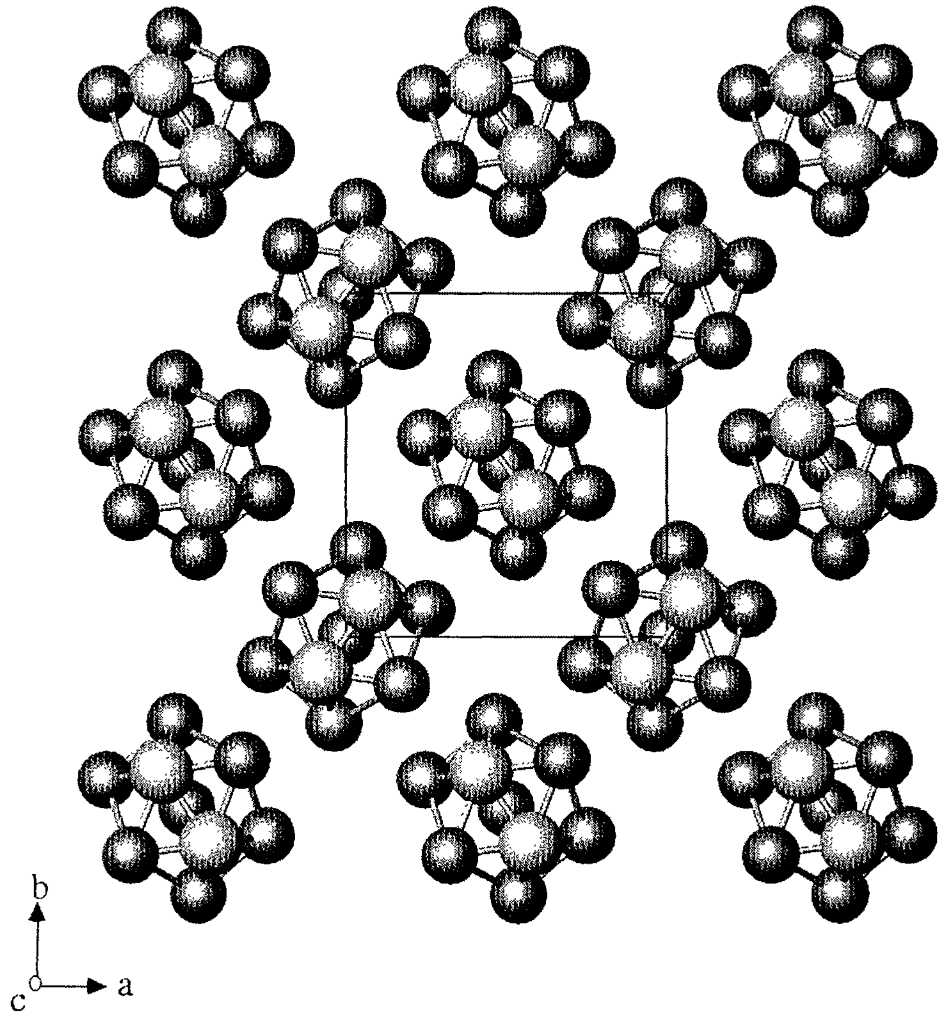


Figure 9: Arrangement of the B_{12}-$(B_{12})_8(B_{12}Si_3)_4$ giant-icosahedron as seen along the c axis of the $YB_{41}Si_{1.2}$ structure In this figure, each B_{12} icosahedron or the $B_{12}Si_3$ unit is depicted with a large circle. (The larger ones are the $B_{12}Si_3$ unit.)

Arrangement of the $B_{12}\cdot(B_{12})_8(B_{12}Si_3)_4$ giant-icosahedron as seen approximately along the $[\overline{110}]$ axis of the $YB_{41}Si_{1.2}$ structure is presented in Figure 10. In this figure, as in Figure 9, each icosahedron or the $B_{12}Si_3$ unit is depicted with a large circle. There are giant-icosahedral chains running along the *c* axis, in which each giant-icosahedron is connected to neighboring ones by sharing the $B_{12}Si_3$-$B_{12}Si_3$ giant icosahedral edges. Unlike the case that the linkages between $B_{12}\cdot(B_{12})_{12}$ giant icosahedra in the YB_{66} structure are always effected through intericosahedral B_{12}-B_{12} bonds along five-fold axes of B_{12} icosdahedra, the linkages formed by sharing the $B_{12}Si_3$-$B_{12}Si_3$ edges necessarily result in a dense giant-icosahedral framework in the $YB_{41}Si_{1.2}$ structure.

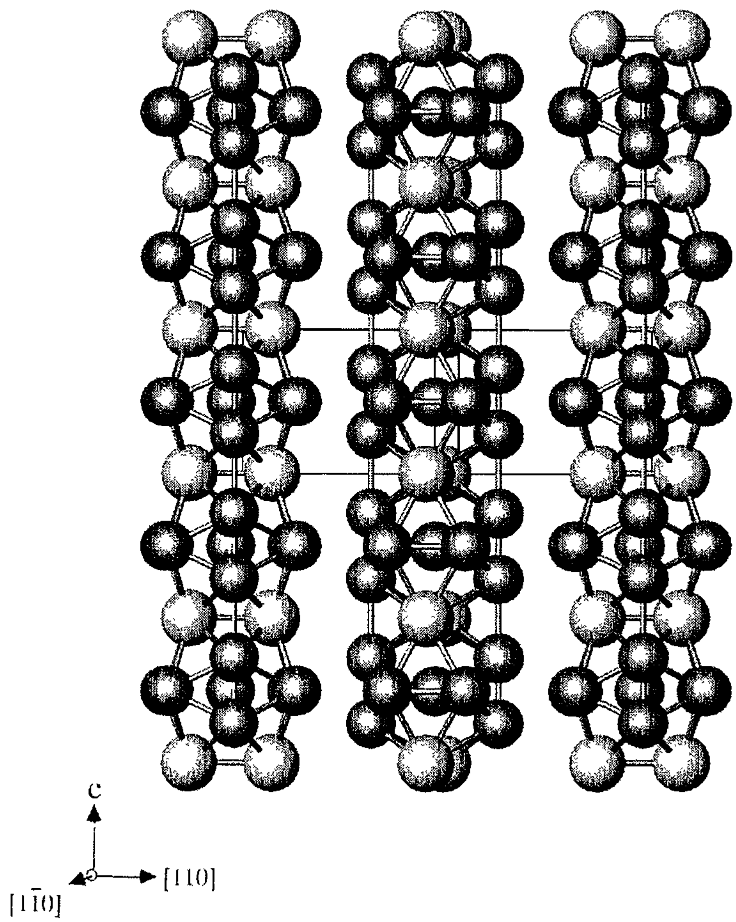


Figure 10: Arrangement of the $B_{12}\cdot(B_{12})_8(B_{12}Si_3)_4$ giant-icosahedron as seen approximately along the $[\overline{110}]$ axis of the $YB_{41}Si_{1.2}$ structure. In this figure, as in Figure 9, each B_{12} icosahedron or the $B_{12}Si_3$ unit is depicted with a large circle. (The larger one is the $B_{12}Si_3$ unit) There are giant-icosahedral chains running along the *c* axis, in which each giant-icosahedron is connected to neighboring ones by sharing the icosahedral $B_{12}Si_3$-$B_{12}Si_3$ edges.

2.4 Packing densities of structural units

Packing densities of structural units in the YB_{66} and the $YB_{41}Si_{1.2}$ structures are compared in Table 1. Calculations of packing densities were made assuming that the structural units are spheres with hypothetical diameters. The diameter of each unit was estimated by examining distances between the centers of directly connected structural units. The B_{12} icosahedron is thus proved to have almost the same size in both crystals; the estimated diameter of B_{12} icosahedron is 0.5 *nm*. The diameters of the $B_{12}Si_3$ unit and the $B_{12}\cdot(B_{12})_{12}$ giant icosahedron were estimated to be 0.53 *nm* and 1.17 *nm*, respectively. As to the $B_{12}\cdot(B_{12})_8(B_{12}Si_3)_4$ giant-icosahedron, the diameter can not be defined owing to its edge sharing with neighboring ones.

Units / Compounds [a]	B_{12}		$B_{12}+B_{12}Si_3$		$B_{12}\cdot(B_{12})_{12}$	
	n	d(%)	n	d(%)	n	d(%)
YB_{66}	104	53			8	52
$YB_{42}Si_{1.2}$ [b]	20	47	20 + 4	58		

[a] *n* is the number of the units contained in a cell, and *d* is volume fraction of the unit/units.
[b] As to $YB_{41}Si_{1.2}$, packing densities of B_{12} ($n = 20$) and B_{12} plus $B_{12}Si_3$ ($n = 24$) were calculated.

Table 1: Packing densities of structural units in YB_{66} and $YB_{42}Si_{1.2}$

The packing densities (52%) of $B_{12}\cdot(B_{12})_{12}$ unit in the YB_{66} structure is the same as that of spheres with equal size in the simple cubic close-packed arrangement, because the $B_{12}\cdot(B_{12})_{12}$ unit is located at the corner of each octant of the cubic structure of YB_{66}, and the hypothetical diameter of the unit is equal to the edge length of the octant. It is noteworthy, however, that the packing densities of B_{12} icosahedral units in the YB_{66} structure is almost equal to that of the $B_{12}\cdot(B_{12})_{12}$ unit. As to $YB_{41}Si_{1.2}$ structure, the packing density of B_{12} unit (47%) is smaller than that in the YB_{66} structure, while the packing density of B_{12} plus $B_{12}Si_3$ (58%) is larger.

3. CONCLUSIONS

Icosahedral arrangements in the YB_{66} and the $YB_{41}Si_{1\ 2}$ structures are presented and unique features in both structures compared. Although the B_{12} icosahedron is the fundamental structural unit of both compounds, the complicated structures are explained in an understandable way by using the $B_{12}\cdot(B_{12})_{12}$ and $B_{12}\cdot(B_{12})_8(B_{12}Si_3)_4$ giant-icosahedral units for description of the YB_{66} and the $YB_{41}Si_{1\ 2}$ structures, respectively.

Spherical diameters were assumed for the structural units, and packing densities of B_{12} and $B_{12}\cdot(B_{12})_{12}$ in YB_{66}, and of B_{12} and B_{12} plus $(B_{12}Si_3)$ in $YB_{41}Si_{1\ 2}$ were estimated; the obtained values were 53, 52, 47, 58%, respectively.

REFERENCES

Higashi, I (1986) Structure and preparation of boron-rich bondes, in: Emin, D., Aselage, T , Beckel, C. L., Howard, I. A. and Wood, C , eds., *Boron Rich Solids*, New York: AIP, pp.1-10.

Higashi, I., Kobayashi, K , Tanaka, T. and Ishizawa, Y. (1995) Icosahedral arrangements in boron-rich solids, *Symmetry Culture and Science,* vol. 6, pp.237-240.

Higashi, I., Tanaka, T., Kobayashi, K., Ishizawa, Y. and Takami, M. (1997) Crystal structure of $YB_{41}Si_{1\ 2}$, *Journal Solid State Chemistry,* vol. 133, pp. 11-15

Richards, S. M. and Kasper, J. S. (1969) The Crystal structure of YB_{66}, *Acta Crystallogr.*, B25, 237-251.

Tanaka, T., Okada, S. and Ishizawa, I (1977) Single crystal growth of a new family Compound· $YB_{44}Si_{1\ 0}$. *Journal Solid State Chemistry,* vol. 133, pp. 55-58.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 269-282, 1998

SYMMETRY AND SOCIAL ORDER: THE TLINGIT INDIANS OF SOUTHERN ALASKA

Margaret Holmes Williamson

Address: Department of Sociology and Anthropology, Mary Washington College, Fredericksburg, Virginia 22401 USA. E-mail. mhuber@mwc.edu

Abstract: *Symmetry is most commonly thought of in material or physical terms. A symmetrical arrangement is immediately apparent to the senses, most especially the sense of sight. Social symmetry also exists, but it is harder to discern because its pattern is usually manifest over time rather than instantaneously. Nevertheless the people constrained by such a system regard it as an atemporal structure, and discuss its temporal expression in such terms. A prime example of a socially symmetrical structure is the moiety system of the Tlingit Indians of Alaska, which affects society at all levels of inclusiveness from the tribe to the household. Tlingit geography, cosmography, and art express symmetry as well: material equivalents of the non-material social organisation. These seemingly disparate symmetries all reiterate the fundamental Tlingit notion that a functioning whole consists of mutually dependent halves.*

If we take a rectangle and bisect it we have an elementary symmetrical form, an automorphism (Figure 1). Elaboration such that any change on one side is carried out on the other reiterates the original balance: one side repeats or mirrors the other, and the two become interchangeable.

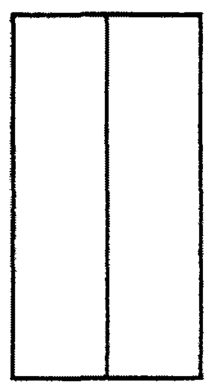


Figure 1: Bisected rectangle, a simplified model of Tlingit social organisation.

Such a figure is commonly used in anthropology to represent the social organisation of the Tlingit Indians of Alaska and British Columbia. This usage does reflect the Tlingit view of the world—for one thing, the great majority of their cultural elaborations reiterate an elementary dualism. But it also obscures what is most interesting about social symmetry, namely, that unlike physical examples of symmetry, which are apprehended immediately by the senses, it achieves itself over time and thus relies on memory for its recognition. Social symmetry, in other words, exists first as an idea in terms of which discrete social activities are interpreted; it is not the result of inductive reasoning. In this paper I discuss the social groups of the Tlingit, the most inclusive of which they viewed as conceptually equal to each other and mutually defining (i.e., in a symmetrical relationship), and then the means by which they recognised and maintained that symmetrical relationship.

For an idea to have currency it must have a collectively recognisable expression. The Tlingit expressed the idea of duality not only in social activities but also in physical forms. Their art, for example, most commonly takes the form of a bilateral symmetry, the left half mirroring the right (see Figure 2).

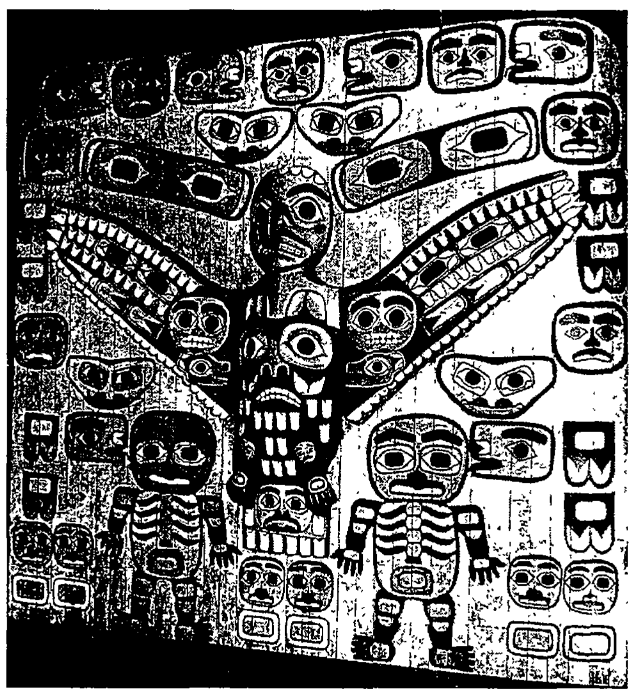


Figure 2: Thunderbird Screen from Thunderbird House, Yakutat. From de Laguna 1972: Plate 91. Photo reproduced by permission of Smithsonian Institution Press.

The plan of the Tlingit house displays a similar structure (see Figure 3). The back of the house was more honorable than the front, and the carved screen that separated the living spaces for the highest-ranking persons from the common area in the middle of the house was not duplicated at the front. Slaves slept on the floor at the very front of the house, the place of lowest honor. As Lewis Henry Morgan observed over a century ago, the architecture of American Indians reflects very exactly their social organisation (Morgan 1965). The Tlingit winter house is no exception to this astute observation. De Laguna says,

> The house of the Tlingit, that is, the real house usually occupied only during the late fall and winter, was more than a solidly built shelter against the cold. It symbolized for the inhabitant the whole social order, his place in lineage and sib, and his family ties with those of the opposite moiety (de Laguna 1972:294).

It thus provides us with a good introduction to Tlingit social organisation, the topic of this paper. Like their houses and their art, the society itself was a symmetrical structure whose component halves, in their various activities, endlessly represented their mutual relationship.

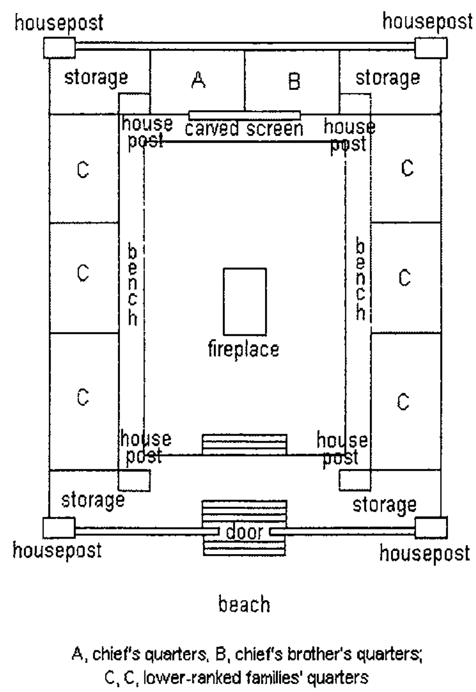
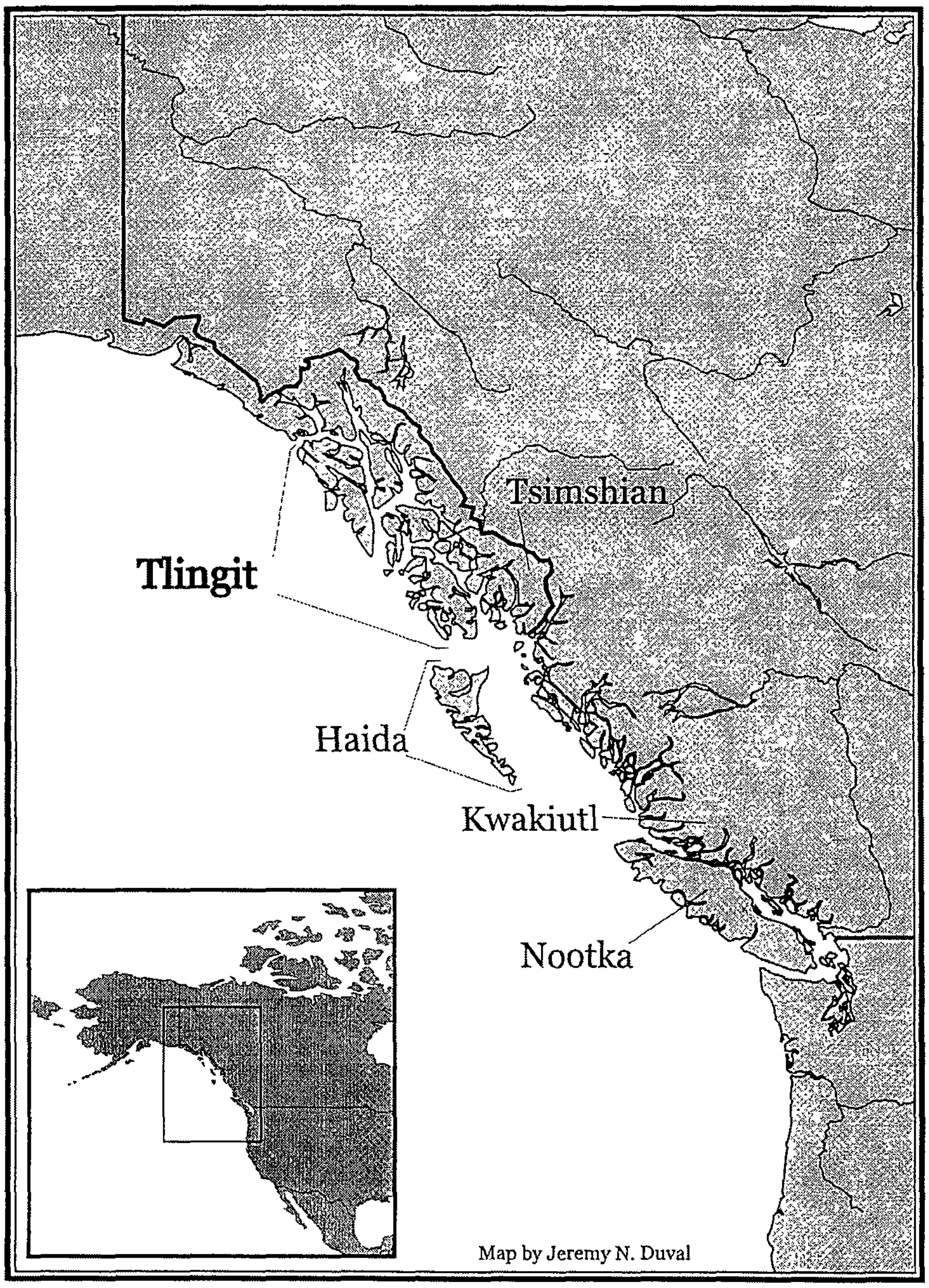


Figure 3: Plan of a Tlingit Winter House (after de Laguna 1972.Fig. 9)



Map: Northwest Coast of North America. Prepared by Jeremy N. Duval, Mary Washington College

The Tlingit are an Athapaskan-speaking group, numbering just under 7,000 in 1880 (Krause 1956:63), who have for several centuries inhabited what we call the Panhandle of Alaska and the very northernmost part of the coast of British Columbia. As the term "Northwest Coast" suggests, they lived between the western mountain ranges and the sea, a remarkably involuted coast characterised by fjords, rocky islands, and strong tides. The land, though unsuited to agriculture, nevertheless offered an enormous amount in the way of resources: a variety of game animals, equally varied berry bushes, giant cedars for construction, and copper. The Indians here became expert fishers, hunters, and gatherers, relying little if at all on domesticated foods for subsistence.

The winter house represented in the diagram was a communal house, several closely-related families living together under one roof Tlingit kinship was based on matrilineal descent, that is, reckoned through women rather than through men or through both sexes equally. The resident owners of the house were most usually a group of men descended from the same mother or mother's mother. Their sisters, resident elsewhere, were also considered owners of the house although they rarely lived there. Besides these matrilineally-related brothers (and it must be remembered that to the Tlingit the child of one's mother's daughter, what we would call a first cousin, was equally a "brother" or a "sister"), resident owners could include the sons of their sisters, who were the heirs and successors to the older generation. (In matrilineal descent property and status as well as kin designations descend through the female line; thus a man' s heir is his sister' s son and never his own.) In addition, the owners ' wives and children, including unmarried daughters, lived in the house; and one or more of their widowed mothers might live there too. A household of wealth and status had one or more slaves as well, acquired in the course of warfare (de Laguna 1972, 294).

Figure 3 identifies, at the back of the winter house, a "chiefs apartment" or room; as I mentioned above, the back of the house was the place of highest honor, the door the place of no honor at all. The spatial ranking reflects not just the social ranking of Tlingit individuals, but a pervasive idea that similar things must be ranked against each other. Thus not only were men ranked against other men, and women against women, but also house against house, lineage against lineage, clan against clan, village against village, and tribe against tribe. Only the two halves of the society, the moieties, were regarded as equal. What rank might mean in any circumstance, and how it was determined for any individual or group, are fascinating questions which are, however, pretty much beyond the scope of this paper. The important point for us here is that each house had a head, the highest-ranking man in the household. His duties included managing the economic resources of the group and maintaining its equally important ritual *paraphernalia*: the masks, rattles, dishes, costumes, and so on that represented the group's crest.

The term crest, or crest object, is the term de Laguna (1972, p. 451) applies to these representations in preference to the more widely-known "totem." Tlingit histories are almost entirely accounts of how they came to acquire these crests in the remote past, and thus why they are entitled to display them today. The Tlingit regarded such rights as exclusive to a particular kin-group, a fact which is significant here because rivalry between clans claiming the same crest or crests was an important aspect of the mutual relationship between the two halves of the society.

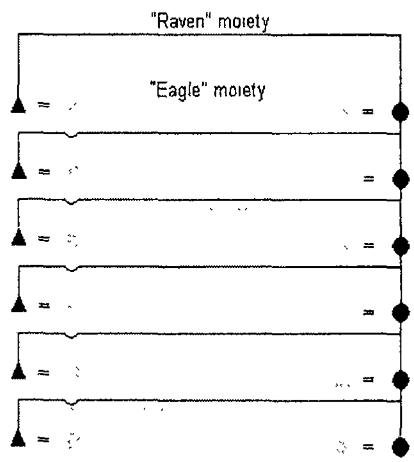
According to de Laguna, the crests "are, from the native point of view, the most important feature of the matrilineal [clan] or lineage " (1972, p. 451). The clans were the most significant social groups among the Tlingit, owning not only their specific crest objects but also tracts of land and seacoast, to which they had exclusive rights of fishing, hunting, gathering, and so on (Kan 1989, p. 23). In addition each clan had a single and common name by which its members identified themselves to each other and to others. Clan affiliation was important both economically and spiritually as well as socially.

Politically, though, it was much less important than the lineages that made it up. These acted together consistently in opposition to other, similar groups, and acknowledged the authority of a single man. The lineage was "like a [clan] in miniature" (de Laguna 1972, p.451, also 212; Krause 1956, p.77; Kan 1989, p.23). It usually acted independently in matters of marriage, display of ceremonial privileges, and economy at the same time as it claimed clan crests in common and recognised obligations of hospitality or help when called upon (de Laguna 1972, p.295; Kan 1989, p.23). A lineage was a co-residential group associated with a specific named house, although usually it occupied several adjacent houses in a village which was home to a lineage of at least one other clan as well (de Laguna 1972, pp.212, 292).

Several such villages made up a tribe, a very loose political group in which the possibility of social solidarity was considerably diminished by the overriding claims of matrilineal loyalty (de Laguna 1972, p.212; Kan 1989, p.3). Even so, persons in one' s own village or the other villages of the locale assisted in such activities as putting on the important potlatch ritual signifying the end of mourning (de Laguna 1972, p.606).

The most inclusive social grouping among the Tlingit was the moiety. As the term suggests, there were necessarily two, one called the Ravens and the other variously the Wolves or the Eagles (de Laguna 1972, p.450; Kan 1989, p.24). Each Tlingit clan belonged to one of these but never both, and some of its crests reflected its moiety association. The moiety was not in any sense a unified group, much less a political body. Its major function was to regulate marriage and the mutual ceremonial services that were the consequence of the marital alliance. Marriage within the moiety was prohibited (de Laguna 1972, p.490; Kan 1989, p.24). Thus every person's spouse came not only from a different clan but also from the opposite moiety.

The result was a system of marriage known in anthropology as bilateral cross-cousin marriage, or sister-exchange. (Cross cousins are the children of siblings of opposite sex. In a system such as the Tlingit, where kinship is reckoned through only one sex, such cousins necessarily belong to a clan other than one's own—in this case, to the opposite moiety—and are very commonly favored as spouses.) As Figure 4 shows, in every generation each descent group both loses a woman to the other group and receives one of *their* women in return: the sisters of each line become the wives of the opposite line. One consequence of this pattern is that every person marries someone who stands in the relation of both mother's brother's child and father's sister's child to him- or herself, since, in the previous generation, the mother's brother has married the father's sister (see Kan 1989, p.24). In doing so the spouses of any generation replicate the relationship established in the previous generation; their children, in turn, replicate it again, and so on.



A "Raven" must always marry an "Eagle," and vice-versa.
Systematically, in each generation a man's wife is also his sister's husband's sister; a woman's husband is also her brother's wife's brother.
Each side loses a woman (sister) and gains a woman (wife) in return, the exchange is equal and so are the sides

Figure 4: Tlingit moiety system, demonstrating marriage (idealised).
Circles = women; triangles = men; equal sign = marriage; vertical line = descent; horizontal line = siblingship.

The perpetual renewal of these relationships is usually a stated aim of marriage systems such as this, and the Tlingit were no exception (de Laguna 1972, p.590). The preferred spouse for either a boy or a girl was someone from the father' s own clan, specifically the daughter or son, as the case might be, of the father' s own sister—what we would call a first cousin (Kan 1989, p.24). In fact families of high rank tried to increase their influence by marrying into a number of other houses (Kan 1989, p.24), but with the idea that a satisfactory relationship would be perpetuated in succeeding generations as well. Figure 5 shows in very simple fashion the outcome of such a scheme of marriages.

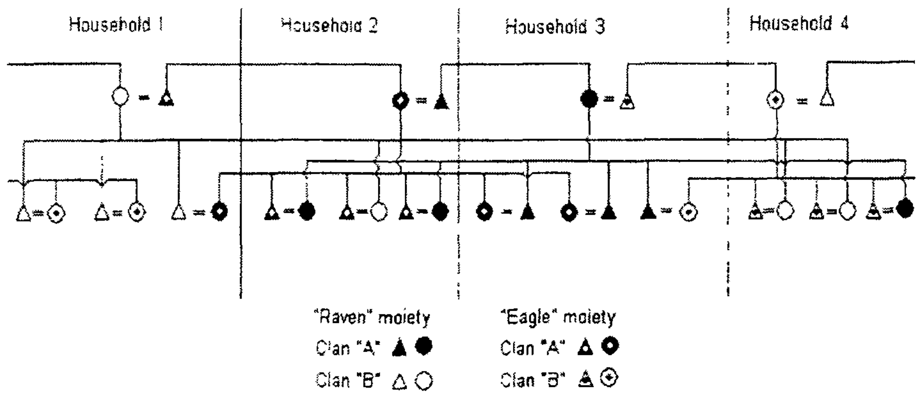


Figure 5: Marriages among four Tlingit households (ideal)

The Tlingit system is an example of a dual organisation, a form of social organisation of which Lévi-Strauss has written,

> This term defines a system in which the members of the community, whether it be a tribe or a village, are divided into two parts which maintain complex relationship varying from open hostility to very close intimacy, and with which various forms of rivalry and co-operation are usually associated (1969, p. 69).

More particularly, it is what he elsewhere (1968, p.139) calls a "diametric" duality, "the result of a balanced and symmetrical dichotomy between social groups, between aspects of the physical world, or between moral or metaphysical attributes." Typically the dual social organisation does indeed reflect a cosmic dual division as well. That is, the collective perception of the world is that it is a whole composed of two parts that are equal in value, if not in size; nothing in the world, therefore, can belong to one or the other of these parts. As this is true of natural phenomena it must also be true of social, or cultural, phenomena, since society is part of this cosmic whole and must therefore be in harmony with what are perceived to be the laws of that whole. The Tlingit associated the Raven moiety with the sea, and the Wolf or Eagle moiety with the land. They regarded both the sea and the inland forest warily, as dangerous places at the periphery of the social world whose core was the village; but they considered the land more comprehensible, and therefore safer, than the sea (Kan 1989, p.118). To this extent only did they impose any sort of rank on these opposed aspects of their physical world; and none at all on the moieties associated with them.

The universal means by which people express equality is exchange. Giving and receiving are integral to all societies, which is to say that in a very real sense society could not exist without exchange (Mauss 1990). In order to understand the thoroughgoing symmetry of Tlingit moieties a review of Mauss's argument is in order. It will be seen, in particular, that just as the symmetry of the bisected rectangle with which I opened this discussion depends on the mutual relationship of the parts in the whole, so also did Tlingit moieties depend for their definition—indeed, their very existence socially and physically—on each other.

Mauss's most significant observation about gifts is that they are not merely economic. That is, exchanges between peoples are not to be understood only in utilitarian terms, as the means to the satisfaction of people's material needs or wants, or to the enrichment of one's self, or to the better distribution of unevenly distributed resources. No one can deny that gifts may and do have these consequences; what we cannot then argue is that the consequences are causes for either the institution of these customs nor their continuation. In particular it is often the case that the completion of a cycle of exchanges leaves the parties materially no better off than they were before.

Socially, however, it may be otherwise. "To give is to show one's superiority, to be more, to be higher in rank, "*magister*" (Mauss 1990, p.74). By the same token, the recipient is placed, or places himself, in a subordinate position; he becomes *minister* to the other. He can rescue himself from his inferior status only by making an at least equivalent return gift. By this means is equality established and, over the long term, maintained despite temporary imbalances of gift-giving, during which one party is in debt (and therefore subordinate) to the other.

To anticipate the more detailed discussion below, Tlingit moieties were in a relationship of balanced reciprocity such that each moiety gave to the other exactly what it got back. This summary remains valid even though the moieties were not corporate groups acting in concert. At any particular location in Tlingit territory the lineages of each moiety exchanged things equally with each other, so that neither ever established superiority over the other. The total effect of all these individual, local exchanges was, similarly, the equality of the moieties. The balance in their exchanges both recognised that this was so, and ensured that it was.

We have seen how this worked with marriage: since one's spouse must come from the opposite moiety and never from one's own, each moiety ceded and gained an equal number of people. But one of Mauss's most important insights regarding the gift is that it constitutes what he calls a "total social phenomenon." This has two meanings. A gift is total in that it is simultaneously juridical, religious, aesthetic, political, and structural as well as economic (Mauss 1990, pp. 3, 38, 78-9). It is juridical, in that it arises from and imposes obligations on the parties to the exchange; religious, even if not offered to spirits, in that its transfer has an effect on and in the spirit world; aesthetic, in that both the gift and the circumstances of its bestowal should make an appropriate impression on the senses; political, in affecting relations of status and between groups; structural, in expressing and symbolising the connections among significant social groups; economic, in being produced and put to use. But a gift may be "total" in another sense as well, part of what Mauss calls a "*system of total services*". Such a system includes not just material items and conventional services (e.g., helping someone roof a house), but "banquets, rituals, military services, women, children, dances, festivals, and fairs" (1990, pp.5-6; emphasis original)

Such totality—in both senses—characterised the gifts given between moieties among the Tlingit. The pre-eminent example of this is the mortuary ritual (see Kan 1989, pp.8-9), a highly complex ceremonial lasting several years and involving great numbers of people from both moieties as well as considerable effort in preparation. Despite its complexity, the basic notion is quite clear: the lineages of each moiety exchange mortuary services and mortuary potlatches with the net effect of maintaining their equal status with regard to each other.

Unfortunately the limitations of space mean that only the most summary account of the mortuary customs may be presented here. [Those interested in reading fuller accounts of these rituals are directed to Emmons 1991, Kan 1989, Krause 1956, and de Laguna 1972.] What is important to the present discussion is that the deceased's father's and wife's relatives were crucial to the completion of the mortuary observances because they, and not the deceased's own matrilineal relatives, accomplished the practical business of the funeral, which had as once consequence the survival and eventual reincarnation of one of the deceased person's souls. They washed, painted, and clothed the body immediately following the death, and they assisted the mourners to assume their proper appearance too; they cooked the food on which they and the mourners lived during this time; they offered comfort to the bereaved; they built the pyre on which the body was cremated and kept it going long enough to accomplish its purpose; they rescued and wrapped the bones and ashes and placed them in the grave house, which they either built or refurbished for the occasion. If the deceased were a particularly important chief, his successor would erect as his memorial an entire new house decorated with clan and lineage crest carvings. This, too, would be built not by the people who would live in it but by their opposites.

The dependence of one moiety on the other as displayed in mortuary observances was not, however, unique to the funeral; it was evident at all life-crises among the Tlingit. A baby was delivered by women of the woman's husband's or father's clan (which might be the same, of course) rather than by the mother's own sisters or mother; and the father performed the birth ritual (Kan 1989, p. 107). The same relationship is apparent at the very important ritual that marked a girl's first menstruation. She was shut away in a small room at the back of the house for a period of time varying from four days to four or eight months or even two years, during which ordeal she was in the care of her father's sister .At the end of the seclusion the girl was provided with completely new clothing and ornaments, and she was tattooed and pierced for the labret (a plug, ranging in size from a collar stud to a bread-and-butter plate, inserted in the lower lip; only Tlingit women wore these)—all by members of her father's clan (de Laguna 1972, pp.518-522). This was, ideally, the same group from whom her husband came (de Laguna 1972, pp.524-525).

In short, no Tlingit, man or woman, could be a complete social person without the active participation of his or her opposites: those related through the father or the spouse. First we must realise that "person" is to be distinguished from "individual." The European concept of the *individual,* a being both biologically and morally independent of others except insofar as he or she wishes to be associated with them, is almost never to be found in non-western cultures. Instead we find the idea of the *person,* a physical and moral being made by, and making, the others in his or her society.

Tlingit saw human beings as made up of a number of spirits and of skin, flesh, and bones, of which they reckoned the eight long bones to be the most important (Kan 1989, pp.49-50). These components did not by themselves constitute a "person" in the Tlingit view. For that one needed a social identity, which one acquired (in part) from the matrilineal ancestors of one's clan. The material aspect of this social identity was the crest objects the clan owned and had the right to display; the non-material aspect included stories, songs, dances, names and titles. These formed a coherent system such that inheritance of an ancestral name gave one the right to tell certain stories and sing certain songs about the ancestor, perform the dances related to these, and wear the crest objects that transformed one from an ordinary Tlingit into the ancestral power itself.

Tlingit held that most of what made up a person came from the mother's side of the family, but they also attributed certain aspects to the father's side, and in fact "fatherless children were seen as incomplete persons" (Kan 1989, p.68). The father's relatives were necessary to the production of social persons in that they carried out the rituals that established one's status as a baby and child of one's parents, as an adult woman, as a married person, and finally as someone deceased. Following Hertz (1960) and van Gennep (1960), we recognise that statuses such as these do not follow automatically from biological events or changes. Just as the appearance of menstrual blood did not in itself establish the girl as a woman among the Tlingit, so also the cessation of breath alone did not establish the sick man as deceased. The Tlingit looked to their opposites to perform the rituals necessary to make the social person, and for each such service a potlatch was due.

The potlatch that repaid the opposites for carrying out the rituals of death was the most important of these. At the potlatch the hosts, who were the matriline of the deceased led by his heir, presented their clan history and identity to an audience made up of local clans of the opposite moiety and at least one clan coming from some distance away. In this endeavor they were assisted by local clans of their own moiety. (Referring to Fig. 5, Household 1 might offer a potlatch to Household 2, to the sister of which one them is married; they would ask Household 3 for support because those men belong to another clan in their own moiety, and invite Household 4—to the sisters of which two of them are married—as the "out-of-town guests.") The symmetry of the moieties is clear here, but we see also that it is a fractal in that each moiety represented was itself divided into two: the primary clan, and the subsidiary clan or clans which were at least potentially opposed in that two clans from the same moiety could, and often did, dispute the ownership of crest objects. That the local clans were willing to support the hosts was a political gesture of solidarity, just as the willingness of the heir's kin to support him as their leader was a political statement that they recognised his right to that status.

For any Tlingit, the opposites performed the rituals and made the crest objects that established him or her as a social person; and in return the opposites received food, gifts, and entertainment at a potlatch. But at a mortuary potlatch, anyway, there was more going on than payment for services rendered. This was a ranked society, and so everyone claimed a certain precedence over others. But claiming by itself does not give high status; it has to be accorded as well. The conditional nature of a person's rank was most obviously the case when the head of a house or a lineage died. At the potlatch for his predecessor the heir invited his opposites to confirm his claim that he was, indeed, the new head of the lineage or of the house. That they would construct that house for him as well as the mortuary paraphernalia and indeed much if not all his ritual accouterments, testified to their agreement that his claim was just. They confirmed this at the potlatch by addressing him, for the first time, with his new names and titles, which he had inherited from the deceased whom they were honoring with the ritual.

With the conclusion of the potlatch the clans from either moiety felt that the unequal relationship between them, the result of the mortuary services being so far unreciprocated, was no longer in effect; they were once again each other' s equals (Kan 1989, pp. 191-197). This fact by itself might seem to confirm my initial assertion that the moieties were equals, equivalent to the two halves of a bisected rectangle. But in fact I have presented here only one-half of the relationship: one moiety performing a funeral, the other paying them for it. True, this did restore equality between them; but since the initial gift—the mortuary services—and the repayment—the potlatch—were not at all the same kind of thing, theoretically a hierarchy might still have persisted. That it did not was due, of course, to the obligation of each moiety to perform such services for the other, and to accept the gifts of a potlatch in return. In fact we have here a four-fold structure: Eagles bury Ravens, and Ravens potlatch Eagles; Ravens bury Eagles, and Eagles potlatch Ravens. It was this structure, much more than the potlatch-for-services exchange, that reflected and maintained the equality of the moieties.

Symmetry may consist in this sort of equality, or equilibrium; but, as I suggested at the beginning of this paper, it requires also a relation of mutual definition. Each half is a half because the other half exists. And we see that that relationship, too, existed between Tlingit moieties. Categorically it operated in a negative way, in that anyone must belong to one moiety and never to both; therefore anyone who was not one of Us must be one of the Others. There were no other possibilities. In practice the principle operated positively in that the social persons (and to a somewhat lesser extent the clans) of a moiety were made by their opposites in the other moiety. Thus the moieties depended on each other for their continued existence.

That said, there remain some unresolved points. Perhaps the most interesting is that unlike their art, indeed unlike most examples of what we classify as symmetries, Tlingit social symmetry had to be apprehended over time—sometimes a very long time. The situation, which is hardly unique to the Tlingit, nevertheless presents an important variation in symmetrical forms, in that the symmetry must be deduced by the action of collective memory rather than perceived immediately by an individual's senses. This is all the more striking in that at any given time every lineage was in debt to its opposites—owing them a funeral owing them a potlatch a girl's initiation, a carving. By the action of memory, local temporary imbalances are overridden by wider-ranging long-term balances; the equality of the moieties both governs and results from this action. The pervasive artistic and architectural representations of bilateral symmetry in Tlingit life presented the model of symmetry to which the contingencies of social life were seen to conform. In the process of establishing that conformation Tlingit in effect suppressed the temporal aspect of life, making the sequential nature of exchanges appear to be simultaneous and, thus, as obviously symmetrical as the paintings on their housefronts.

REFERENCES

Emmons, G. T. (1991) *The Tlingit Indians,* edited with additions by Frederica de Laguna and a biography by Jean Low, University of Washington Press.

Gregory, C. A. (1982) *Gifts and Commodities,* Academic Press, xiv + 242 pp.

Hertz, R. (1960) *Death and The Right Hand,* translated from the French by Rodney and Claudia Needham, with an Introduction by E. E. Evans-Pritchard, Cohen and West, 174 pp.

Kan, S (1989) *Symbolic Immortality· The Tlingit Potlatch of the Nineteenth Century,* Smithsonian Institution Press.

Krause, A. (1956) *The Tlingit Indians: Results of a Trip to the Northwest Coast of American and the Bering Straits,* translated from the German by Erna Gunther, University of Washington Press.

Laguna, F. de (1972) Under Mount Saint Elias: The history and culture of the Yakutat Tlingit, *Smithsonian Contributions to Anthropology,* 7 (3 parts), Smithsonian Institution Press.

Lévi-Strauss, C. (1969) *The Elementary Structure of Kinship,* translated from French by James Harle Bell and John Richard von Strumer and Rodney Needham, editor, Beacon Press.

Lévi-Strauss, C. (1968) *Structural Anthropology,* translated from French by Claire Jacobson and Brooke Grundfest Schoepf, *Allen Lane the Penguin Press,* xxii + 410 pp

Mauss, M. (1990) *The Gift: The form and reason for exchange in archaic societies.* Translated by w .D. Halls with a Foreword by Mary Douglas, W W. Norton, Inc

Morgan, L. H (1881/1965) *Houses and House-Life of the American Aborigines.* Introduction by Paul Bohannan, University of Chicago Press.

Spencer, R. F , Jennings, J. D. et al (1977) *The Native Americans,* 2nd ed., Harper and Row.

Van Baal, J. (1975) *Reciprocity and the Position of W omen,* Van Gorcum: Assen.

Van Gennep, A (1960) *The Rites of Passage,* translated from French by M. Vizedom and G. Caffee, University of Chicago Press

Weiner, A. (1992) *Inalienable Possessions: The Paradox of Keeping-While-Giving,* University of California Press.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 283-295, 1998

APPROACHES IN URBAN DESIGN: THE DISORDER THAT RESULTS FROM ORDERING THE DISORDER

Einat Kalisch-Rotem[1] and Iris Aravot[2]

[1] 1995 - B.Arch , Faculty of Architecture and Town Planning, Technion. I I.T., Israel. Currently studying for M Sc. and teaching urban design theory and practice.

Address. Faculty of Architecture and Town Planning, Technion - Israel Institute of Technology, Haifa 32000, Israel *Private address:* 26 Gut Levin St., Haifa 32922, Israel. E-mail: KEB@actcom.co.il.

Fields of interest and research: urban design theories and practice, characteristics of the post-industrial world.

[2] Senior Lecturer. 1978 - B.Arch. (cum laude), 1984 - M.Sc.; 1987 - D.Sc , thesis *"Toward a Theory of Architectural Knowledge"*, Faculty of Architecture and Town Planning, Technion. I I.T , Isreal; 1984-88 Philosophy studies at Haifa University. 1988- Postdoctoral studies, A.A , London. Currently head of the center for Architectural Research and Development.

Address: Faculty of Architecture and Town Planning, Technion - Israel Institute of Technology, Haifa 32000, Israel. E-mail: aria@techunix.technion.ac.il.

Fields of interest and research: urban design, architectural philosophy, architectural education.

Publications: Published numerous articles in Hebrew and English. Practices in Urban Design.

Abstract: *Theories in urban design, also known as "approaches", are part of the same category, despite having more differences than similarities. Nevertheless, they all have several denominators in common. This paper will attempt to explain what are urban design approaches in general, describe types of existing approaches, and introduce both their similar and the different characteristics.*

In order to clarify a complex subject appearing in most of the approaches, a process of comparison, which is one of many ways of clarification, can be used. This can be done by a "comparative chart", trying to reorganize an abstract material into a two dimensional study. This paper will illustrate a comparison between five chosen approaches - which are not necessarily typical, common, or more important than others- the approaches of Coop-Himmelblau, Koolhaas, Venturi, Cullen and Rossi.

The process of imposing order has advantages (such as easing comprehension) but also bears its costs: the complex subject becomes simplified, presents a partial picture and leaves a question mark on its reliability. This paper will examine the consequences of the comparison, with its advantages and disadvantages and present the result as a non-perfect order, a term which holds in itself a contradiction.

1. INTRODUCTION

This essay attempts to demonstrate a process of ordering literal material, in order to extract a specific issue. This ordering process, that occurs in the urban design discipline but which can be also identified in others, is a personal experience brought to cognizance. The essay will show that this examination turned out to be not so simple: a complex process on one hand with quite simplified results on the other.

At the beginning, this essay will attempt to explain the complexity of the urban design theoretical domain, define central terms in the field, and introduce their varied contents. After understanding the problematic character of urban design theories, a process of clarifying will gradually emerge but will still try to remain, as far as possible, within the abstract sphere.

Then, five different urban design theories will be introduced, and will illustrate the former argumentation. From this point onwards the complex comparison process will accelerate, and become more concrete with the support of these five approaches. At the end the results of the process will be discussed.

2. THE NATURE OF URBAN DESIGN THEORIES

Understanding the term 'urban design theories' may reveal two different complex terms: *urban design* (which has much in common, in many ways, with the term *architecture*) and *theory* (which can also be referred to as *approach, ideology* and even a *manifesto* in this field).

2.1 The complexity of the urban design theoretical domain

Architecture and urban design are often regarded as a combination of *art and science*. This is possibly one reason for the complexity in the structure and content of theories. Bill Hillier discusses the differences between architectural and scientific theories: the scientific theory is "a rational construct intended to capture the lawfulness of how the world *is"*, help understand phenomena, whereas architectural theory apparently seems to be "a set of guidelines as to how it *should be* [...] and by that express aspirations rather then realities". Hillier adds that "scientific theories wish to help us act on the world, because they rely on first describing the world independently of any view of how it should be". The essence of this argument represents the difference between the *analytic* and the *normative* intent - a description how the world is, or a prescription how it ought to be (Hillier 1996, p.57).

Hillier is trying to clarify his distinctions. "Do theories in architecture really only mean creating a formula for architectural success? [...] On a closer examination, this is not the case. Admittedly, architectural theories are normally presented through the normative intent, but at a deeper level they are no less analytic than scientific theories" (Hillier 1996, pp. 57-58). Nevertheless, the problem with most architectural theories is that the normative aspects come to dominate the analytic (p. 68).

The second reason of complexity, derives from the word "*theory*" with its numerous meanings, and from other *analogous terms* used in this field: *approach, ideology* and sometimes *manifesto*. Despite the clear linguistic differences between these terms, they are often used to express similar things in the urban design discipline.

The ancient Greek origins reveal that the verb *theoreein* means to be a spectator, and the product of this speculative activity, *theoremata*, were speculations. The word has a long etymological history, but one source of its ambiguity lies in the nature of theories themselves: theories are found in the realm of speculative thought, "because they are at root, speculations" (Hillier 1996, p. 68).

The ambiguity is also a result of the multiple and varied linguistic meanings to the four words, to which 15 relevant meanings were found in the Webster's dictionary. According to the dictionary, the boundaries of the terms are sometimes vague, therefore their usage is not obvious. Jencks introduces his own unusual interpretation to these terms, and emphasises especially the nature of the *manifesto*, which he finds unique, in relation to the *theory*, which he describes as something less violent - a "congealed manifesto" - in order to become acceptable in the groves of academy (Jencks 1997, p. 8).

All the four terms mentioned are commonly used in the urban design language. Nevertheless, the term 'approach' is used predominantly and is somehow more common then the more epistemologically appropriate term - 'ideology'. This essay will cling to the prevalent term.

The third and last reason of complexity is a result of the different and *varied contents* found within urban-design approaches, as expressing peoples' beliefs. They can be opposed to one another in many cases and produce a "varied environment, a maximum choice for society... Eisenman, the master theorist and polemicist, inscribed his tablets in the pages of his magazine *oppositions.*" (Jencks 1996 pp. 8-9.)

2.2 Different classifications in the urban design theoretical domain

The previous part of this essay attempted to demonstrate why urban design theories are complex, first by presenting the inconclusively *origins* of the urban design discipline (art and science), then by clarifying the varied definitions of the term *theory* that represent different theoretical *methods*, and finally, by showing the diverse and colorful possibilities, that cause inconsistent and even opposing *contents.*

This part will introduce different classifications of three experts who suggest their personal order on the urban design theoretical field.

Hillier, an urban design theorist, defines two kinds of typical architectural theories: *The broad and the narrow propositions*. Broad propositions are intended to be universal by attempting to express ideas about architecture which are held to be generally true. The narrow propositions are offered as possible techniques for realizing an abstractly stated aim, by trying to bridge between the abstract and the concrete. According to Hillier, the problem of most architectural theories is that they are over specific where they should be permissive, and vague where they should be precise. (Hillier 1996)

Jencks, an architectural critic, classifies four categories differentiated by the subjects that motivates them (the situations they are responding or reacting to), and by the models they are built on (their ideology - what they wish to change, solve or improve).

He tries to denominate each category or *Tradition*, as he calls it, by the most prevailing definers: *Traditional Architecture, Late Modern Architecture, New Modern Architecture, Post-Modern Architecture.* (Jencks 1997, p. 9.) Jencks indicates that his classification method does not lack problems: although most architects remain loyal to one approach, few architects jump between one tradition to another, and some "do not fit happily into any tradition."

Broadbent, an architectural educator and researcher, classifies three categories that represent three different ways of thinking, "which developed over the centuries, into coherent - and rival - philosophies: *Empiricism*, which puts its trust in the human senses; *Rationalism* which does not, preferring to work in logical steps from first principles and *Pragmatism* which prefers things which are known to work in practice" (Broadbent 1990, p. 79.) The three basic ways of thinking were known and distinguished from each other by the ancient Greeks.

2.3 The similar characteristics found within urban design approaches

The three attempts to classify urban design approaches, as demonstrated in the previous part, represent personal notions of enforcing order within this field. Despite the differences, these experts share a diagnosis regarding the structure of the theory. This part will try to formulate a shared classification according to these experts.

Post-modern approaches are *positioned alike in the world of theory*: they are connected to the existing reality in a way of a crosscut, and to precedents and traditions in a way of a vertical section, each one, of course, in a different way. Hillier finds that the need for architectural theory arises from the need to formulate principles from our past experiences in order to guide us how to build in the future. (p. 84.) He is definitely applying to the vertical connection (past-present-future) more strongly then to the present realities of the world. Nevertheless, "theorisation begins when we note a certain type of phenomenon and then make a certain type of presupposition. The phenomenon we note is that of *surface regularity* in the world as we experience it", and by that he refers also to the present reality. (Hillier 1996, p. 70.)

Jencks applies strongly to the present situation as he believes theories are a result and a cause: they are created as a reaction to a special situation (present reality), either spatial or intangible (architectural, political, social, philosophical, etc.), with a wish to change, solve, improve this situation (think about the future). Approaches are "a cause to new built spatial environments". Le Corbusier and Eisenman proved that theory is "an engine of architecture... a machine which invents new types of buildings and new responses to the city", as future oriented theories. (Jencks 1997, p. 8.)

Post-modern approaches have *resembling structure*: according to Hillier, theories contain precepts about what designers should do (the normative intent), and a prior framework which describes *how the world is* (an analytical method). Careful examination will show that this is always the case with architectural theories. Sometimes this framework is explicitly set out, sometimes it is much more implicit. (Hillier 1996, p. 58.)

Resembling structure to that of Hillier's, is diagnosed by Jencks, who describes its two major structural components in a more blatant way: a motive to change the world - a crisis or a feeling of imminent catastrophe, and the pure theory which is based on science and logic. Jencks cites the bible. The first component he names *'The volcano'* (explosion of emotion): "a motive for destruction - to inspire fear in order to create unity and orthodoxy... It is still a tactic of Modernists, Late Modernists, and Prince Charles with his decalogue of ten principles. Those who write manifestos are jealous prophets who call the class to order." (Jencks 1997, pp. 6-7.) Jencks adds that our age is responding to a changing world, to the global economy, ecological crises and cultural confusions. In effect, these are a "second type of volcano." (p. 8.) The second component, *'The tablet'* (the laws and theories) is a metaphor of pure theory, cited also after the Bible.

The last structural similarity is the *theories' constituents* which create the forms of expression: words, and some other formal expression. According to Hillier (1983, 1984) the formal expression is usually mathematical, but this is no doubt his credo. Cullen would have said that formal expressions are drawings and illustrations. (Cullen 1971)

According to Jencks, the manifesto is constructed with two additional elements to those of the approaches: *The personal element*: "the most effective manifestos constantly address the reader as 'you' and reiterates the joint 'we' until an implied pact is built up between the author and reader (Jencks 1997, p. 7. after L.C. 1923), and *the contrast* - a comparison between good and bad. Jencks adds that without the first component, "the volcano", the manifesto would not be written. All four strategies of the manifesto are evident in Coop Himmelblau's *architecture must blaze*, a new modern manifesto of 1980. Here we find the bad (Biedermeier), the good (architecture that 'lights up'), and the two are distinguished in the first person plural ('we are tired of seeing Palladio and other historical masks'). The tablet of virtues is architecture that is 'fiery, smooth, hard, angular' etc., and the volcanic violence is architecture that 'bleeds, whirl, break' etc. (Jencks 1997, p. 8. after Coop-Himmelblau 1984)

3. DEMONSTRATION BY FIVE APPROACHES

Up to this point, this essay introduced and discussed urban design theoretical field in general, without mentioning any concrete approach, and attempted to demonstrate that the definition of urban design theory is complex, in a way that leaves many open ended possibilities of theoretical variations. From this point onwards, this essay will continue the discussion with reference to five approaches - which are *not* typical, common or more important than others:

Venturi and colleagues identify a new phenomenon - the American commercial strip in its extreme form - the Las-Vegas strip, and believe we should understand it better (Venturi, Brown & Izenour 1979); *Cullen* is drawn to the way people experience a place emotionally through the sense of sight (Cullen 1971). He wishes to make pleasurable places, as a modern version of the old picturesque movement of the 18^{th} century (Broadbent 1990); *Rossi*, a neo-rationalist, is interested in studying the architectural discipline and finding the architectural essence of *artefacts* - historical urban buildings that despite changing their functioning did not change their architectural form (Rossi 1982); *Coop-Himmelblau*, a deconstructive group, are searching after a new improved but not a beatified world (Werner 1989). Their different aesthetics are based on a new personal thinking processes - breaking and reconnecting in an innovative way (Farrelly 1986a); *Koolhaas* is using the phenomena of Manhattan as a model to outline high-density, high-rise metropolis. He believes any ideology will eventually change in the future by a new one - a belief in a pluralist world. (Koolhaas 1994a,b)

3.1 A process of comparison

A need to scan the issue of 'street and square as constituents of the public domain' within urban design approaches arose. It began a process including several phases (and will no doubt include more) until receiving a fairly, satisfactory result. During the process, it was realised, that this scanning can be done only through a comparison between the relevant approaches, because it is the only way to understand the approaches' attitude towards a specific topic *definitively* and *relatively*.

The first topic that seemed relevant to this study was *the attitude of the approaches towards the street and square*, as something that might indicate the relationships between them.

1st topic: *the relationship between the approach and the S&S*

Gordon Cullen	Aldo Rossi	Robert Venturi	Rem Koolhaas	Coop Himmelblau
Urban events take place in the public domain - the street and square.	Urban events exist in the essence of buildings, and therefore in the public domain artefacts	Urban events take place on the American strip in Las Vegas.	Urban events take place inside the blocks, between the street and square system	Urban events take place near the traditional S&S system, and change it.
$A \approx S+S$	$S+S \in A$	$S+S \notin A$	$S+S \neq A$	$A+S+S \Rightarrow (S+S)'$

The attitude of the approach towards the street and square (S&S) is a qualitative variant that does not fully succeed in explaining the Street & Square situation. What perhaps can also be tested is the possible future of the street and square according to the approach: *Do they have a future, and of what kind?*

2nd topic: *the future of the S&S according to the approach*

Gordon Cullen	Aldo Rossi	Robert Venturi	Rem Koolhaas	Coop Himmelblau
His approach takes place in the streets and squares. This is the place where his ideas can happen. *S+S* must have a better future. They are most important.	The architectural *artefacts* are in search The street and square are the spaces in between but not less important *S+S* will have a future because they are as important as artefacts.	The street & square do not exist, only the commercial strip in its extreme form - in Las Vegas does. *S+S* might have no future, and are not significant in natural economical processes.	Between the buildings is the street system the only real element even though it has no existence of its own S+S will have inevitable future as leftover space in the city. The city life will pass to its buildings.	The city buildings have new aesthetic characteristics which will affect the street and square. S+S will exist, but in a new way and with new spatial configuration. The traditional *S+S* have no future.

The results of the comparison display a partial picture because we can not understand *if these were the desired conclusions or merely the identified and therefore 'reported' conclusions*. Does Koolhaas really wish to leave the streets of Manhattan empty, or is it merely the problem of focusing on other issues which he finds more important or intriguing, and by that neglects the Street & Square? Does Venturi ignore the traditional street because it doesn't exist, or because it is not in his interest? The answer to these questions has to do with *ideology: Why* exactly are these the possible futures, *how* are they intending to happen and *what* will they include- might be better understood through the process of thinking that created this conceptual future.

The following charts try to organize ideological segments that together assemble the basis of the approach - "the platform". This part of the comparison lights up the street & square issue, by explaining the process of reaching conclusions. Without it - the street & square issue can be interpreted in several ways, and not particularly in the way intended by the approach.

Gordon Cullen	**Aldo Rossi**	**Robert Venturi**	**Rem Koolhaas**	**Coop Himmelblau**

3[rd] topic: *the approach is a reaction to...*

Ill recent past reflected through the implementations of urban design projects (Cullen 1971) the failure of modernist cities and built environments.	Ill recent past reflected through the implementations of urban design projects (Rossi 1982b) the failure of modernist cities and built environments	Emerging new reality in the American post - modern society in general, and in Las-Vegas in particular (Venturi, Brown & Izenour 1979) Identification of the current phenomena- the "commercial vernacular".	New-York as a phenomena of a special metropolis that needs exploration (Koolhaas1994a) identification of the old phenomena that was never discussed before.	Ill past reflected in architectural projects and through the sequential theories that did not success either (Werner 1989) The failure of modernist and post-modernist architectural projects.

4[th] topic: *aiming to...*

learn and use the art of "townscape" that has proved itself in the past **Create better built environments** that of today's	discover the essence of historical urban architecture in order to (Vidler 1976) **Create historical built environments**	analyze the reality in order to understand and interpret the new phenomena. **Bring** the phenomena **to attention.**	interpret this phenomena with exaggeration to make a point and (Jencks and Kropf 1997) To **encourage pluralism** and tolerance towards all theories.	invent new architectural codes of aesthetics and become a neoterist (Farrelly 1986a,b) **Create new** built **architecture** and environment.

5[th] topic: *values, basic premises, world-view*

Sympathy to the picturesque. Planning should be done with a reference to what the eye sees (Broadbent 1990) Neo-empiricist	Marxism and rationalism - in their 1970s' modern philosophical version (Broadbent 1990) Neo-rationalist	Understand and interpret the world we live in, because whatever exists - is worthy (Venturi, Brown & Izenour 1979). Pragmatist	A metropolis is a manmade place where all myths and ideologies can co-exist. Every belief is legitimate (Koolhaas 1994a,b) Pluralist	Everything that has been done is ill. Now is a time for a new start, using the same material but in a new way (Farrelly 1986a) Idealist

6th topic: *argumentation for using the values*

Planning only with the aid of scientific research has failed It is not enough (Cullen 1971)	Architecture should stand for itself and not rely on other disciplines but itself any more (Moneo 1976)	Analysing existed phenomenon has long history - it is a way to look back in order to advance (Venturi, Brown & Izenour 1979).	a belief that the principles of everything external, can be found inside the human mind. (Koolhaas 1994a,b)	Personal ideology, inward processes and insights (Werner 1989, Farrelly 1986a)
Personal ideology	Personal ideology	Personal ideology	Personal ideology	Personal ideology

The self-evident conclusion, naturally, is that all theoretical 'platforms' are, in essence, personal ideologies, and therefore are not suitable for debating. The next three charts discuss the methodological foundation - how the ideology is proven or demonstrated.

Gordon Cullen	**Aldo Rossi**	**Robert Venturi**	**Rem Koolhaas**	**Coop Himmelblau**

7th topic: *references to science/knowledge system*

The relations between objects as seen by the eyes - "art of relationship" (Cullen 1971)	Study of plans, sections elevations of historical buildings that represent (Rossi 1982a)	Las Vegas existing commercial strip as (Venturi, Brown & Izenour 1979)	Manhattan's history and myths, together with psychological understanding is (Koolhaas 1994a,b)	A very inward looking process that searches for knowledge (Farrelly 1986a, Tschumi 1987)
the ability of the **human sight system** to see and consequently, evoke experience and feelings	the essence of the structural physical **architecture discipline**	a current extreme but representative **case study** tested through **analytical tools**	an extreme, exaggerated and a retroactive **case study** tested mainly through **psychology**	a **self-searching process** (part of which rely on Derrida's **philosophy**.)

8th topic: *method of research*

Drawings, illustrations and sketches that enable	Examination of buildings in order to find their essence, and achieve	Analysis of plans, sections photographs etc., results as	Exploration of Manhattan's historical narratives and myths through	Implementations of personal ideas and personal manifestos by (Libeskind 1988, 1987)
systematic exploration of urban situations.	list of artefacts through a rational **typification.**	a **conventional** architectural **analysis research.**	**written material** that constitutes Manhattan's retroactive manifesto.	**realization of** architectonic **projects in practice.**

9th topic: *a product as a result*

Demonstration of ideas - in a book and in a non realized project (Cullen 1976, Broadbent 1990)	Demonstration of ideas - in a book and in projects (Moneo 1976)	Demonstration of ideas - in books and in projects (Venturi, Brown & Izenour 1979)	Demonstration of ideas - in a book and in projects (Koolhaas 1994b)	Demonstration of ideas - in manifestos and mainly in projects (Jencks & Kropf 1997)
implementation by 'do and not do commandments' - "the scanner" (Cullen 1966)	implementation by illustrating examples	implementation by being aware to new phenomena	a change in the perception and understandings of any reality (Aravot 1995)	creating a change in conventions.

The first four topics explained the personal motivation and the process which stimulated the outcome. The next three topics described the outcome itself through the product intended to be left to mankind, in order to help others realize their ideology. The attempt to understand the street & square issue, as demonstrated in the current part, is introducing a process of a question-and-answer cycle, in which an answer to one question raises another question, and its answer raises another, and so on.

The chosen topics introduced in this essay are only part of the possibilities of further investigation. The street & square issue, as a wide subject in itself, can be fragmented once again into further topics, such as *the future spatial structure of the street & square* (containing style, historical references, character of space, etc.), *their future functional character* (activeness, measure of liveliness, social role), etc. Continuing this comparison process is perhaps a matter for another essay. What is relevant to this one, is discussing its results.

4. ADVANTAGES AND DISADVANTAGES OF THE ORDER IMPOSITION

The described comparison process is a sort of classification procedure that extracts the topics of comparison from within the approaches themselves. The method of extracting topics represents a typification process, where - the certain topics are found within one or few approaches, and the other approaches are forced inside into these predetermined topics. It is not always possible to insert them easily, sometimes it demands certain creativity and using different techniques such as skipping over few topics, because naturally, not all approaches relate to all topics. In other cases there are topics that, despite being identified in each approach, are *so* different in content, it seems they have nothing in common, and therefore are not comparable. But this, of course, is also a kind of comparison's conclusion.

As shown at the beginning of this essay, Jencks defined four types of traditions within the architectural and urban design approaches. He calls it 'capsule definitions' and indicates that "it is always reductive to define growing, complex movements, always foolhardy because it can never be done satisfactorily, and always necessary - in order to clarify the issue at stake. [...] Those definitions, however, are academic, theoretical, bloodless - not something to leave home for (the ultimate aim of a good manifesto). They are necessary for cool ratiocination and comparison, which is why they are included, but I defy you to repeat them verbatim, without looking" (Jencks 1997, pp. 9-10).

5. CONCLUSION

Although the above described theories are not easy to compare, it is somehow possible to force them into order through a comparison process. All comparison charts should be read and understood together, in order to clarify the certain issue wished to be examined in the first place - in this essay, the street & square issue. In many cases, the comparison chart inevitably stays partly empty. This too is a characteristic feature of the comparison. Ironically, this clears the compared issue, and enables us to see and understand the whole picture: the approaches themselves and the relationships between them.

The conclusions of this essay are already known for a long time: the ordering process of verbal material is reductive. There is partial information that either falls out or is forced in throughout the process. These two forcing actions are necessary in order to create a new, simplified arranged material, in a way of comparison charts. Nevertheless, the whole picture is now comprehended, a picture that introduces a non-perfect order, a somewhat self-contradicting term.

The importance of these results is not only in their mere existence and presentation but also and mainly in the methodology by which they emerged. Therefore the innovation of this essay lies in its methodology and its application in the urban design theoretical sphere.

REFERENCES

Aravot, I (1995) Narrative-Myth and Urban Design, *Journal of Architectural Education* 49, 1, Nov., pp.79-92.

Broadbent, G. (1990) *Emerging Concepts in Urban Design*, New York· Van Nostrand Reinhold International

Cullen, G. (1966) *The Scanner*, London· Alcan Industries Ltd.

Cullen, G. (1971) *The Concise Townscape*, London: The architectural Press.

Farrelly, E. M (1986a) Coop Himmelblau, *The Architectural Review*, 8/1986, pp 18-24.

Farrelly, E. M. (1986b) The New Spirit, *The Architectural Review*, 8/1986, pp. 6-13.

Hillier, B. (1996) *Space is the Machine*, xiii, 463, Cambridge University Press.

Hillier, B. (1989) The Architecture of the Urban Object, *Ekistics*, no 334, January-February.

Hillier, B. (1984) *The Social Logic of Space*, xiii, 281, Cambridge University Press

Hillier, B. (1983) Space Syntax - A Different Urban Perspective, *Architectural Journal* 178, no 48, pp 47-63.

Jencks, C. (1977) *The Language of Post Modern Architecture*, pp. 39-52, New York: Rizzoli.

Jencks, C. and Kropf, K. (1997) *Theories and Manifestos of Contemporary Architecture*, Chichester: Academy Editions.

Koolhaas, R (1994a) *Delirios New York - A Retroactive Manifesto for Manhattan*, NY: The Monacelli Press.

Koolhaas, R (1994b) Life in the Metropolis, *Architectural Design* 47, no. 5, pp. 315-369.

Libeskind, D (1988) City Edge Competition, Berlin, *A&U*, no. 215, August 1988.

Libeskind, D (1987) City Edge Competition, Berlin, *AA Files*, no 14, Spring 1987.

Moneo, R. (1976) Aldo Rossi: The Idea of Architecture and the Modena Cemetery, *Oppositions*, no. 5.

Rossi, A. (1982a) *The Architecture of the City*, Cambridge, Mass: M.I.T. Press.

Rossi, A (1982b) The Analogous City, *Lotus International* 13, pp. 5-9.

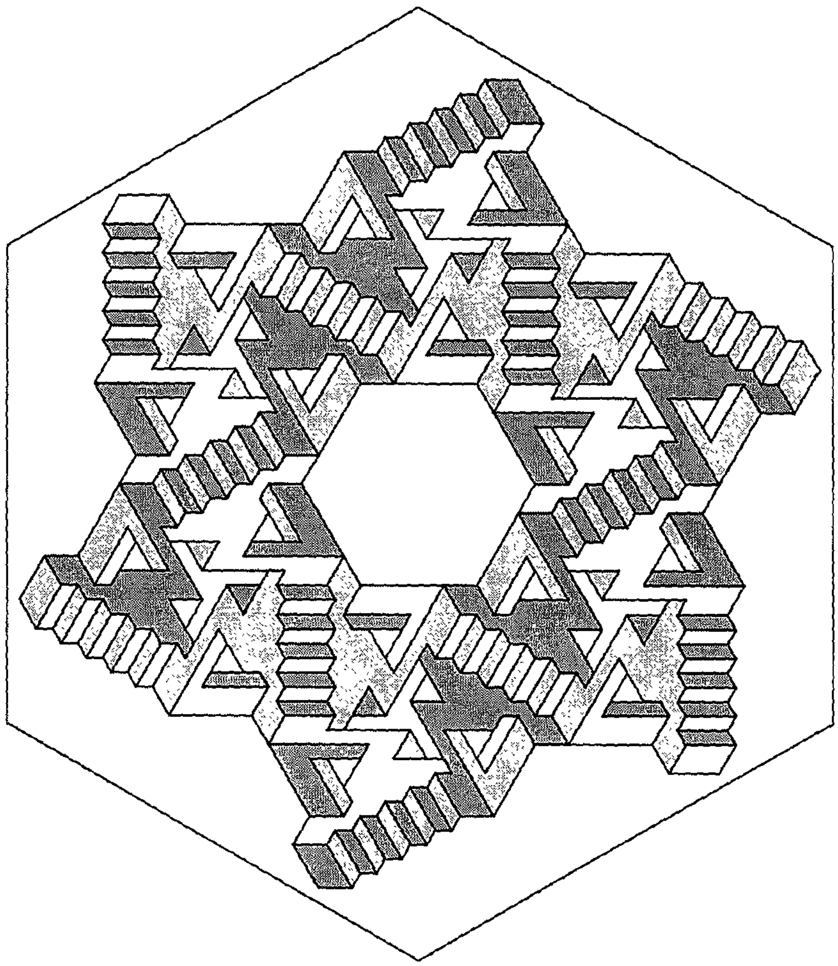
Tschumi, B. (1987) *Cinegramme Folie, Le Parc de la Villette*, Princeton· Princeton Architectural Press.

Venturi, R., Scott Brown D. and Izenour S. (1979) *Learning from Las Vegas*, Cambridge, Mass. and London: MIT Press

Vidler, A. (1976) The Third Typology, *Oppositions*, no. 7.

Webster's Ninth New Collegiate Dictionary (1991), by Merriam Webster Inc., pp. 98, 597, 724, 1223

Werner, F. (1989) Coop-Himmelblau, 1968-1989, *A&U* , 226, July 1989, pp. 15-133.



Symmetry: Culture and Science
Vol 9, Nos. 2-4, 297-305, 1998

THE ORDER OF STAIRS: ARCHITECTURAL CONSIDERATIONS IN STAIRCASE DESIGN

Eyal Karni

Architect *(b. Haifa, Israel, 1952).*

Address· Faculty of Architecture and Town Planning, Technion-Israel Institute of Technology, Technion City, Haifa, Israel 32000. E-mail: arkarni@tx.technion ac il

Fields of interest: Building Technology, Industrialized Building, Spatial Structures, Building Details.

Awards: Prize for Excellence for M Sc. research thesis of the School of Graduate Studies, Technion, 1983; *Rubinstein Prize for the project: "Membrane Structures for Roofing Free Wide Spans", 1988, Dori Prize of Haifa Municipality - Honourable Mention, received for the book. "Gypsum in Building", 1990;* Recipient of the "Joseph Levy Academic Lecturship in Architecture and Town Planning", 1993-1995.

Publcations· Karni, E., Burt, M. & Czamanski, D. (1990) "Approach to Comparative Evaluation of Construction Systems of Light-Weight Space Structures", *International Journal of Space Structures*, (England) Vol. 5 No. 1, 28-37.; Karni, E. (1994) "Membranes for Spatial Structures", *International Journal of Space Structures*, (England) Vol. 9 No. 1, 38-45., Karni, J and Karni, E. (1995) "Gypsum in Construction - Origin and Properties", *Materials and Structures (RILEM),* (France). Vol 28, 92-100.; Karni, E. and Werczberger, E. (1995) "The use of the Compromise Criterion for Evaluation of Multi-objective Alternatives", *Environment & Planning, (Planning & Design b),* (U.S.A.) Vol. 22, pp. 407-418 ; Karni, E. (1995) "Enhancing User's Flexibility in Adaptable Dwelling Units in High-Rise Public Buildings", *Open House International,* (U K). Vol 20, No. 2, pp 39-45., Karni, E (1996) "Plastic Cladding for Bar, Joint and Cladding Structures - Physical Properties and Performance", *Materials and Structures (RILEM),* (France) Vol 29, May, pp. 241-249.; Karni, E. (1996) "Retractable Spatial Structures for Swimming Pool Enclosures", *International Journal of Space Structures,* (England). Vol 10 No. 4, 231-236.

Abstract: *Man spatial movement within given volume of structures is a complicated fascinating phenomena. On the face of it, such 3-D movement seems to be a simple trival matter. However, careful examination reveals a multi-criteria design problem which should address a variety of architectural-engineering considerations. A proper result of sound architectural design of stairs is a biased order of programmatic goals designed to achieve desired movement in built space. However, malpractice could result in a chaotic- disorder, the result of which is an unsatisfactory structure.*

In this paper, architectural considerations of stair design are brought forward, together with optional geometric solutions, demonstrated by outstanding examples of fine architectural design.

1. STAIR DESIGN - THE MEETING POINT OF ART AND TECHNOLOGY

Stairs are one of the most fascinating building sub-systems when architectural design is concerned. Apart from being a tool for connecting various levels inside structures - without which no human activity could take place when the third dimension is concerned, stairs combine aesthetics as well as spatial geometry, building and construction technology, art and craftsmenship. A proper staircase is a work of art. It might be regarded as a 'construction statue'. Its spatial appearance inflicts upon personal feeling and human perception. Presenting staircases and exploring the immediate space around them to the ascending or descending people is one of the most valuable assets of designers trying to achieve structural quality.

Stairs have inspired people for ages and several eye-pleasing examples have lasted through building history both structurally as well as aesthetically. The first well-known stairs have probably been biblical such as Jacobs ladder to heaven and the Babylon tower, which were expressed in artistic works of Peter Bruegel and Abel Pan. These were later succeeded by the famous flights of stairs leading to the Parthenon, followed by those in the Epidarius theatre (350 B.C.) and the stairs designed in the Colosseum (72-82).

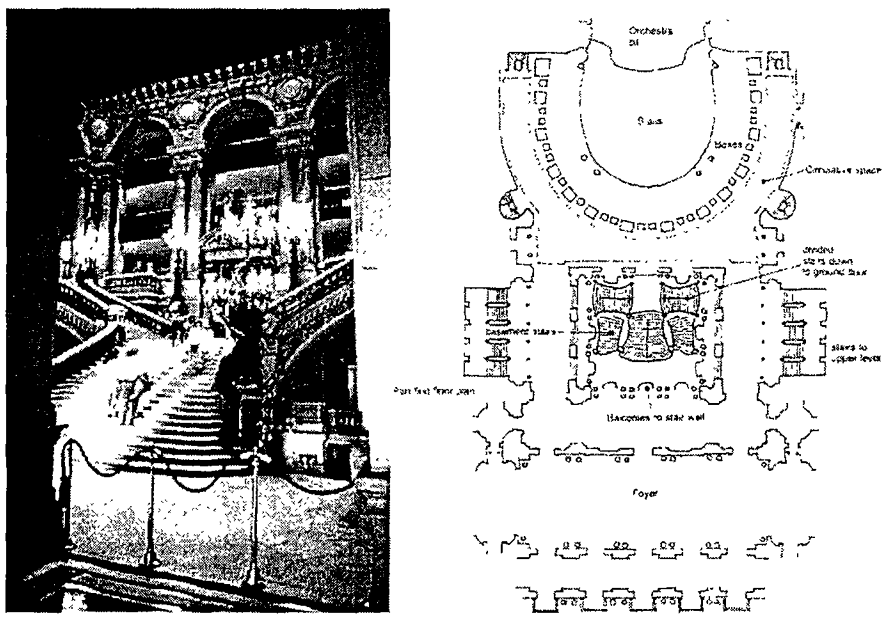


Figure 1: Pans Opera House

In recent times the double staircase designed in Chateau de Blois (1515-1530), the grand staircase built at the Paris Opera House (see Figure 1) by Charles Garnier (1861-1974), the double radial staircase leading to the Sistine Chapel at the Vatican, Gaudi's spiral stairs at the Sagrada Familia (1920), the well-known mule-path in Santorini (see Figure 2), the Spanish stairs in Rome (end of the 17^{th} century), Frank Lloyd Wright descending spiral ramp in the Guggenheim museum (1956, see Figure 3), the escalator system in Charles de Gaule terminal, the escalator in Pompidou centre (Piano and Rogers 1977), Lloyd's building entrance staircase (Rogers, 1986) and lately the spiral ramp in the Louvre glass pyramid (Pei, Cobb & Freed; see Figure 4). All these fascinating man-made inventions are fine examples of human spirit combined in architectural-engineering skill supported by superb building craftsmanship.

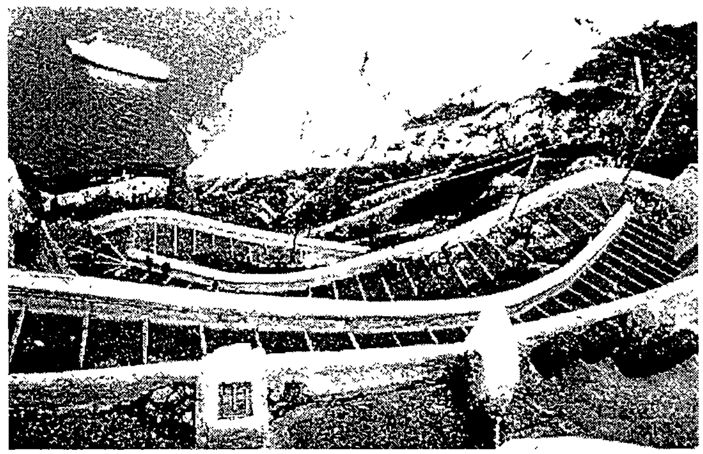


Figure 2: The mule-path in Santonni

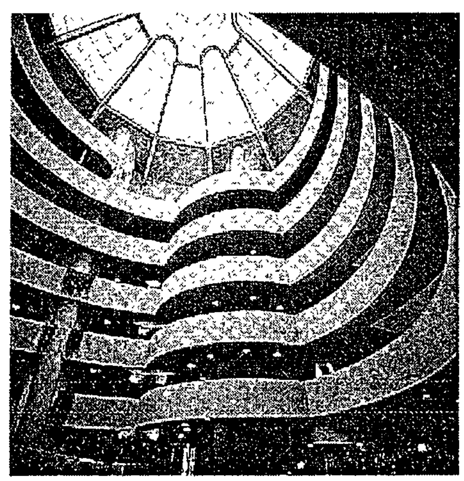


Figure 3: Frank Lloyd Wright Guggenheim museum, N.Y.

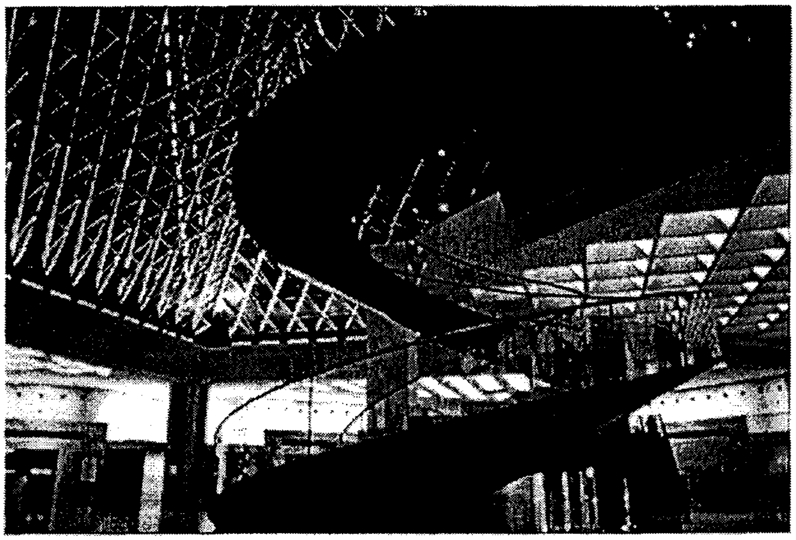


Figure 4: Louvre spiral ramp

Applying the appropriate staircase to buildings interior space, physical performance and man comfort presents first class professional challenge to designers. Objectives like aesthetics, spatial appearance, walking directions in the various levels connected to staircase, constructive system supporting the flights of stairs, sound selection of building materials, texture and colour, walking comfort, safety, maintenance, cost and industrialization - are all to be properly addressed by designers. Alternatives should be regarded leading to the best possible choice addressing all design goals. Such work is not an easy achievement and various solutions have proved numerous times to be unsatisfactory and even dangerous as Edward Bear - Christopher Robin play-mate experienced when making his way down in a most unorderly manner:

"Here is Edward Bear, coming downstairs, bump, bump, bump,

on the back of his head, behind Christopher Robin.

It is, as far as he knows, the only way of coming downstairs,

but sometimes he feels that there really is another way,

if only he could stop bumping for a moment and think of it..."

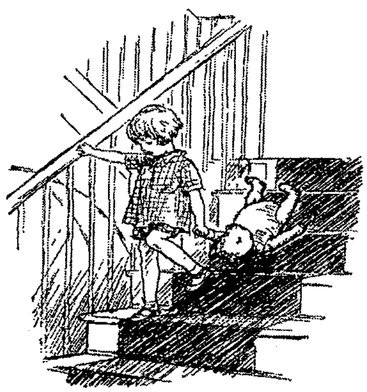


Figure 5: Winnie-the-Pooh by A. A. Milne.
Decorations by E. H. Shephard

THE ORDER OF STAIRS:
ARCHITECTURAL CONSIDERATIONS IN STAIR DESIGN

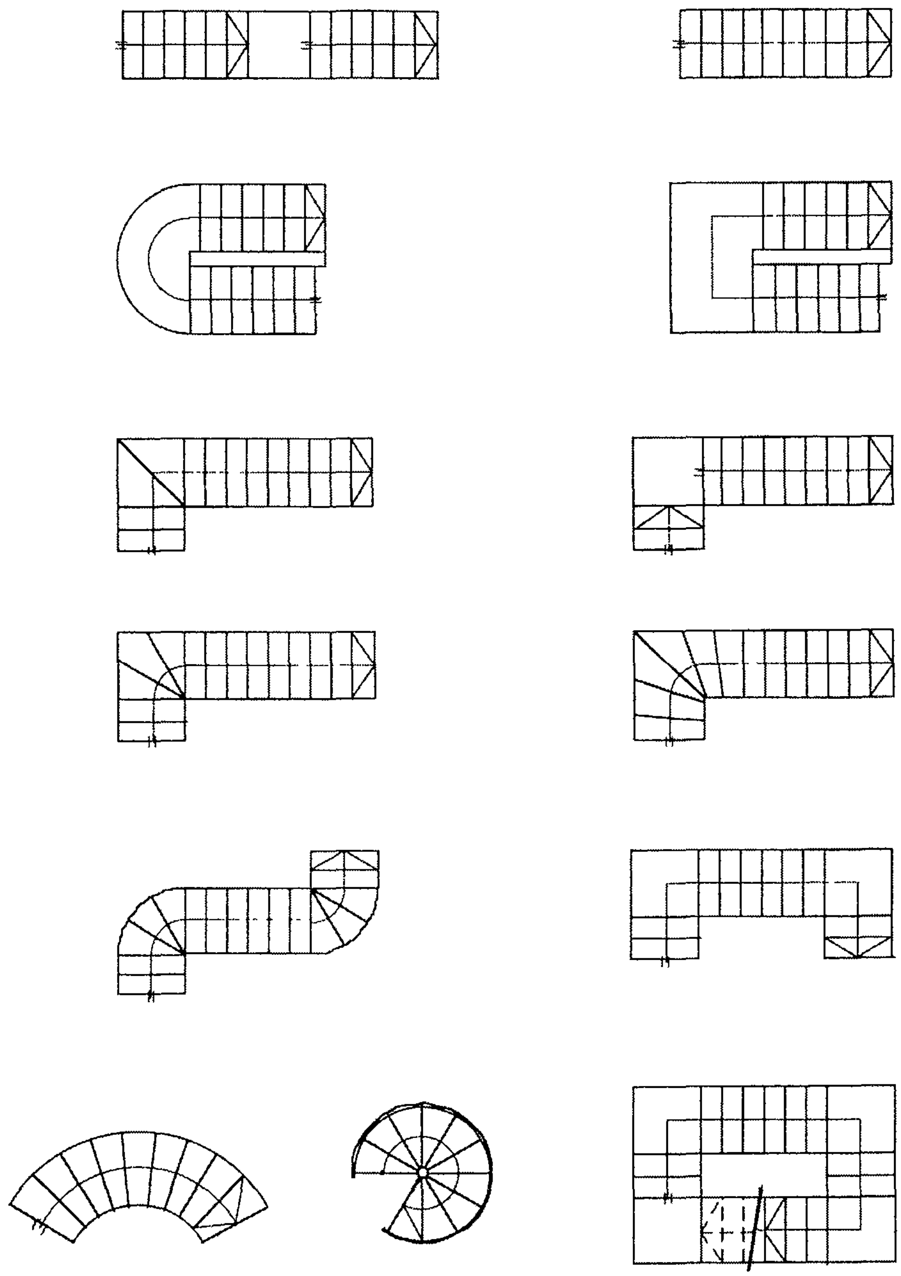


Figure 6: Various staircase geometries

2. TYPES OF STAIRS

2.1 Geometrical Classification

The variety of buildings has produced various types of stair design differing in shape, spatial organization, overall dimensions, construction methods and building materials.

In Figure 6 common types are presented according to staircase spatial geometry. Some staircases are linear, some *L* shaped while others may converge in any other degree, radial, elliptic and hybrid. Staircases may include resting areas *(podest)* or not up to a certain number of stairs in each flight. Dimensions may also vary according to specific stair dimensions (tread and riser).

2.2 Architectural Considerations of Stair Design

Choosing a suitable solution among such alternatives is influenced by several architectural/engineering constraints:

- Overall dimensions of specific space designed to accommodate the staircase
- Directions from which people are expected to approach or leave stairs
- Comfort expected to be achieved while ascending/descending
- Eye-pleasing appearance of the stairs in structure
- The effect the stairs have on the activity conducted near by
- Structural system needed to support the staircase and the building construction system to support stairs' system

3. ORDER-DISORDER IN STAIR DESIGN

Having presented the relevant considerations regarded when designing stairs, the question of order-disorder arises. In fact, the problem is how the overall geometry of a staircase and its specific place inside the building volume create order or perhaps cause disorder within structures. Can a professional decision create chaos or perhaps enhance building integrity concerning movement, user's flexibility and comfort, etc?

A fine example is Corbusier's design (see Figure 7) to the artist Ozenfant. Corbusier had to place a flight of stairs leading to an upper floor in Ozenfant's studio. Instead of accommodating the stairs adjacent to the wall in a straight or *L* shaped flight, he decided to use the diagonal for a linear flight. By doing so, an interesting space was achieved, as well as using the stairs as a semi-open partition dividing the small studio into two interior spaces which seemed to be needed for the artist's work.

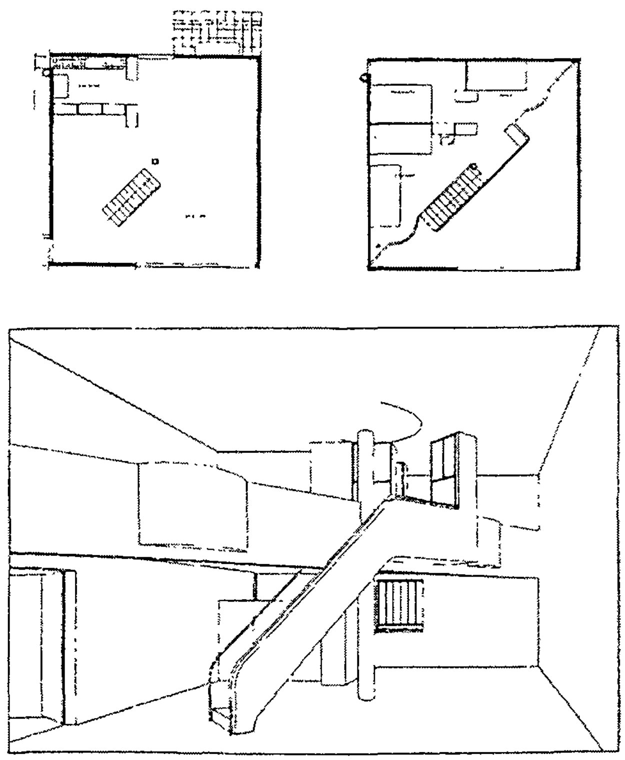


Figure 7: Corbusier maison pour artisans

Other cases may involve the question of symmetry. Would symmetrical stairs prove to be a better solution or should asymmetrical solution be applied? The grand staircase in Charles Garnier's Paris opera house (see Figure 1) seems to be the perfect solution for this unique, highly praised building. Here, a symmetrical solution was applied, offering opera lovers either right or left approach to the hall. Clearly the solution is a perfect result of the interior volume of the building. So is the case of the exterior stairs leading to the *Sacré Coeur* or the Spanish Stairs in Rome considered by many as the finest in the world.

When symmetry is involved several possibilities are relevant such as mirror (left-right) symmetry, radial symmetry etc. The sophisticated double spiral staircase at the Sistine Chapel in the Vatican (see Figure 8) is a puzzling architectural invention. Although space is small, the objective was to divide the ascending from the descending thus allowing better movement inside the entrance to the gallery. The idea was not new originating in Leonardo's double helical stair although it had been applied only few times before as in the *Chateau de Blois*. Double symmetry is achieved. First a radial one then a mirror type symmetry concerning the crowd movement. Symmetry is used again in Pei's East Wing of the Art Museum in Washington where two magnificent small radial stone stairs are situated at both ends of the building.

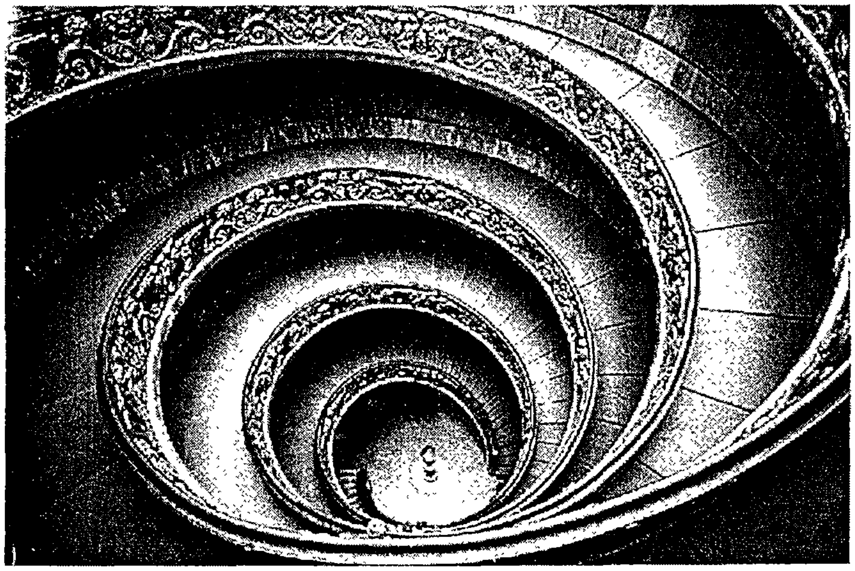


Figure 8: Vatican – double spiral staircase at the entrance to the Sistine Chapel

However, asymmetrical solutions seem to enjoy special qualities as well, providing their own contribution to building order. Perhaps, the famous of all is the mule ramp in the island of Santorini, which had served for years as the main climbing route to the top of the mountain. The escalators system in Trump Tower, N.Y. is not a symmetrical solution either as it is not placed in the middle of the interior space thus leaving enough room for other activities. Pei, Cobb and Freed Louvre spiral ramp is asymmetrical, too, serving only the descending crowd, while a side escalator carries the ascending leaving crowd out of the museum. Right's descending spiral ramp in the Guggenheim museum is another contribution to architectural ingenuity not only serving as descending route, but also creating a great interior for a relative small museum, allowing all-around view and creating pleasant atmosphere.

Another puzzling effect of order and space conception as well as user's comfort and safety is the turning direction when using stairs. Should the user turn left or right when ascending/descending? Professional experience proves that descending is more important to regard due to safety considerations. Descending to the left (anti-clockwise) is easier than to the right. The same problem is clearly experienced when running in a sport arena. All over the world the preferred turning direction is anti-clockwise and so is the case in most airfields for descending-landing airplanes. There is a hidden characteristic in mankind enabling better turning movement in an anti-clockwise manner than vice versa, and this inflicts directly on stair design although normally such design is created unconsciously by architects.

4. SUMMARY

Stair design has direct impact on order/disorder of architectural design. Stairs are work of art as well as technological inventions. Professional design of stairs provides users with sound understanding and biased perception of space, thus creating pleasant atmosphere and satisfaction. Malpractice concerning such design may prove unsatisfactory as well as hazardous and could lead to physical damage resulting from disorder in space perception and environmental behaviour.

Therefore special attention should be paid to architectural design of stairs providing users with an aesthetic building sub-system as well as with proper technological structural component.

REFERENCES

Blanc, A. (1996) Stairs, Steps and Ramps, *Butterworth Artchitecture.*

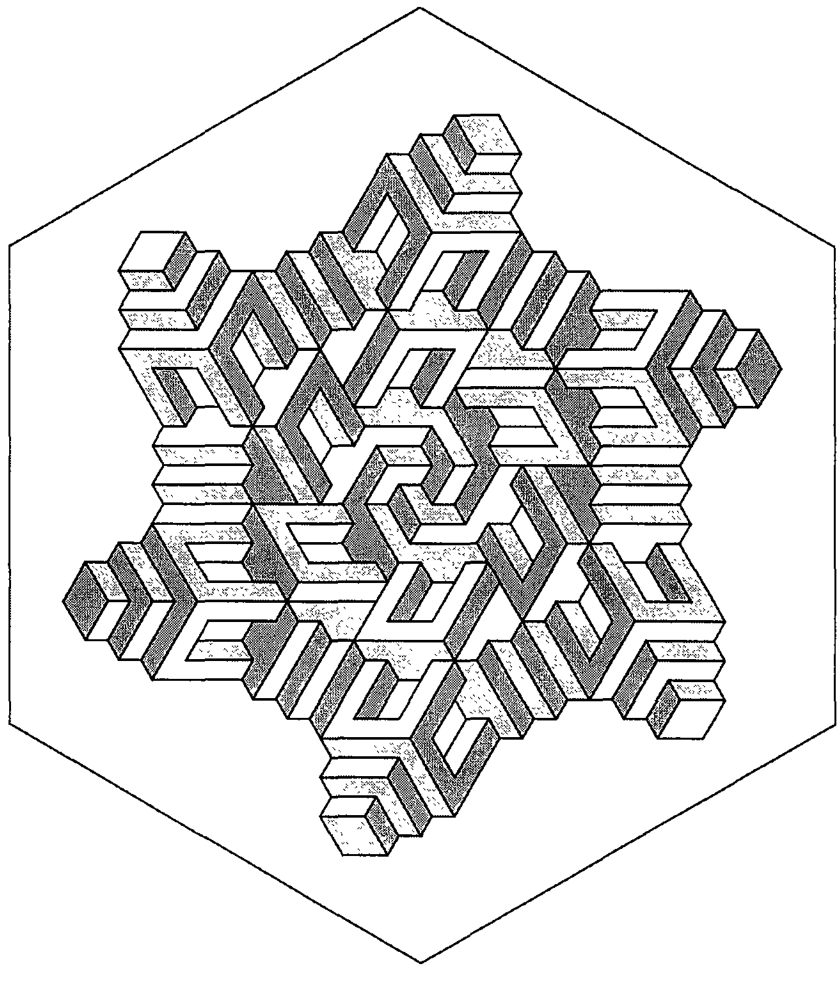
Gladischefski, H and Halmburger, K (1974) *Treppen in Stahl*, Bauverlag GmbH.

Hoffmann, K , and Griese, H. (1977) *Stahltreppen*, Julius Hoffmann Stuttgart.

Karni, E. (1997) *Stair Design*, Technion - Israel Institute of Technology, (In Hebrew).

Mannes W. (1981) *Technik des Treppenbaus*, Deutsche Verlags-Anstalt

Templer, J. (1992) *The Staircase*, Masachusetts Institute of Technology



Symmetry: Culture and Science
Vol. 9, No.2-4, 307-314, 1998

PHILOSOPHY AND AESTHETIC PREFERENCES: SYMMETRY VERSUS ASYMMETRY

Elzbieta T. Kazmierczak

Designer and a semiotician. (b. Poland)
Address: Art Department, 202 Center for the Arts, State University of New York at Buffalo, Buffalo, NY 14260-6010, USA. E-mail: elkakazm@acsu.buffalo.edu.
Fields of interest: the search of cognitive paradigms of visual representation, the origins and evolution of graphic forms and symbols, diagrammatic modeling as related to cognition.
Publication: Glowne choremy w semiografii: kolo, kwadrat, trojkat (Main Choremes in Semiography: Circle, Square and Triangle) (1986, rev. 1990) Warszawa, Poland.

Abstract: *In this paper, it is argued that the difference in aesthetic preferences between the Western culture and Japan is determined by the difference in their philosophical traditions, and not the unique features of the Japanese landscape itself, as is commonly believed.*

In support of this argument, the relationship between the ontological paradigms of the early Greek philosophy with its logical formalism is compared with the intuitionism of the Zen philosophy to show how the implications of these two philosophical orientations affect the aesthetic preferences in the Western and Japanese cultures and, consequently, the form and the content of their art and design works. The fundamental difference between aesthetic preferences in these two traditions narrows down to the opposition between symmetry and asymmetry. Therefore, the key questions asked in this paper are as follows: Why, in the Far East, is asymmetry preferred over symmetry? Why, in the West, is the approach opposite: asymmetry is considered to be a lack of order, and symmetry governs the popular aesthetics?

> "Everything exists according to its own nature. Our individual perceptions of worth, correctness, beauty, size, and value exist inside our heads, not outside them."
>
> 2^{nd} *Zen Tenet* (Holmes, 1992, p. 28)

1. WESTERN RATIONALISM AND EASTERN INTUITIONISM

Western culture has sprung from Greek origins and, as such, has inherited the Greek rational approach; the Greek Logos. Western rationalism is derived from logic, which, in turn, is founded upon a set of principles which are claimed to be the laws of thinking, thereby revealing the intrinsic structure of the cognitive mind. According to Leonard Russell (1914, p. 58), "They express the most elementary and fundamental characteristics of the nature of predication in its formal aspect." Let me recall these laws in their logical order first and then paraphrase them through their implications on the Greek metaphysics. There are three laws: (1) the Principle of the Identity of Terms, (2) the Law of Non-contradiction, and (3) the Law of Excluded Middle. The first law says that "every singular term is predictable of itself (A=A)" (ibid., p. 59). That tautology is expressed, for instance, in Yahveh's "I am that I am." The two remaining laws can be thought of simultaneously: "Of the two propositions, 'S is P,' 'S is not P;' one must be true and one false." (ibid., p. 62). In the realm of metaphysics, these laws require that the existing thing cannot fail to exist. In other words, the thing has to be identical with itself and cannot be its own negation, it cannot be self-contradictory. Thus, admitting that something is and is not at the same time is logically incongruous. The logical formalism draws a clear cut separation between affirmation and denial and thus enhances the conceptualization of the world in terms of binary oppositions, as if it were perceived in black or white. Thus, the mainstream of the Western culture, determined by bipolar oriented thinking, gives priority to the (bilaterally) symmetrical organization.

The Parmenidean theory of material monism (c.450 BC) embodies the principles of logic in the extreme form; "What is, is; what is not, is not." That simple tautology implies (1) the existence and the materiality of the world, thus, non-existence of the void, and (2) the static, self-identical, and uniform character of the world. According to that theory, empty space does not exist; it even cannot be thought of, for "one cannot think of nothing" (B. Russell, 1959, p. 28, 29). So, from that perspective, the term "empty space" is an empty name without designation, although its denotation is an empty set. On the contrary, in modern science, it is charged with designation, since the empty space is conceived as having a real existence.

Although the Parmenidean perspective — viewing the world as static and uniform — dominates in the Western tradition, the opposite concept of the world, that is, as perpetually changing, was also developed by the Greeks. Nonetheless, it never entered the mainstream of Western metaphysics. In regard to logical formalism, the Heraclitean "We step and do not step in the same river; we are and are not" (c.500 BC) implies a logical contradiction of "A" becoming its own negation, that is, "not A," during the process of change. Thus, the Heraclitean approach implies not only a conceptualization of the world beyond the dichotomy of being or not being, but also accepts the existence of void (Russell, 1959, pp. 25-29).

Although Heraclitus (c. 540-c. 470 BC) derives from rationalism, he goes beyond the logical formalism. His perspective, by assuming the existence of the void and the changeability of the world, relates to the Eastern tradition, which, on the other hand, ultimately embodies the intuitive approach of the Zen masters. "Zen has no philosophy of its own. Its teaching is concentrated on an intuitive experience... ." (D. Suzuki, 1959, p. 44). "In other words, it is a mode of activity which comes directly out of one's inmost self without being intercepted by the dichotomous intellect." (ibid., p. 140). Paradoxically, ideas which are logically incongruous from the Western perspective are perfectly valid in the Eastern approach. For instance, the concept of the Mind of No-Mind requires denial of self-awareness of the mind. According to Zen teaching "bad is good, ugly is beautiful, false is true, imperfect is perfect, and also conversely." (ibid., p. 33). In the West, these contradictions are rejected as logically incongruous. Zen aims beyond the conceptualizations of the intellect, to reach the feelng of, and identification with reality through the pre-intellectual and direct experience. Thus, the goal in Zen is "to restore the experience of original inseparability [...] to return to the original state of purity and transparency." (ibid., p. 359). "The body and the mind are not separated, as they are in the case of intellectualization. The mind and the body move in perfect unison, with no interference from the intellect or emotion. Even the distinction of subject and object is annihilated"(Suzuki, 1959, p. 146).

2. THE CONCEPT OF VOID AND ITS IMPLICATIONS

The conceptual denial of the existence of void by the mainstream Western tradition determines the connotations of the concept of space per se. Since empty space does not exist, humans in Western culture do not feel comfortable living in a non-existing void. Thus, empty space has to be filled by existing things to be validated. Hence, in the realm of the Western art and design, space is the element which does not count autonomously. In general, the importance of empty space is underestimated. Although the Gestalt theory of visual perception (1920-25) acknowledged empty space as a design element, it still charged it with pejorative connotations and assigned it a secondary function, naming it a "negative space" or "ground." Thus, the negation of the existence of empty space results in visual abundance, which characterizes Western culture, as if in quantity there was a quality.

Alternatively, the predominating Eastern tradition concept of the perpetually changing world does not imply the denial of the existence of void. On the contrary, it requires the existence of empty space, as without it, there is no room for change or progression. Consequently, it is given its own voice as a condition of the dynamic world. Since void is recognized as a value in itself, it does not have to be filled with existing things to be validated. Thus, it does not make people feel uncomfortable while realizing that they are living in the void. Moreover, space is utilized as an integral part of any art work and is given the most important role among other elements of visual language (Gunji, 1990, p. 1), leading toward minimalist art, with its principle of simplicity that is ingrained in the Eastern culture. The "thrifty brush" tradition, which demands "the least possible number of lines or strokes which go to represent forms on silk or paper" (Suzuki, 1959, p. 22), expresses the Zen spirit, which longs for primitive simplicity, close to the natural way of living, and "has no taste for complexities that lie on the surface of life." (ibid., p. 23).

3. WESTERN DUALISM AND JAPANESE TRIADIC HIERARCHY

Western logical formalism, with its demand of non-contradiction, determines the conceptualization of experience in terms of binary oppositions. The rule of binary oppositions is claimed to be an intrinsic structure of the cognitive mind. Functional anthropologist Bronislaw Malinowski (1884-1942) and others maintained that an opposition between life and death is fundamental and primary (p.1913), the one from which others have sprung (Brozi, 1983, pp. 68-82, 95-104). On the level of language, an absolute notion which does not have its counter notion has not been developed. Moreover, such an absolute is inconceivable. Every notion in the Western tradition is bipolar and relative. For example, is – is not, good – evil, subject – object, culture – nature. From the standpoint of rational philosophy, the basic binary opposition between being and not being requires only two elements to be satisfied. The two are enough to create the static world, which is in relation to itself, since nothing else exists. Therefore, the binary opposition is complete, as such, and reveals the tension between the two elements of the same importance. The Westerner — enveloped by rationality with its logical formalism and atomistic approach — takes for granted the discretness and stability of the world and demands it to be safe and predictable in its bilaterally symmetrical order.

On the other hand, the dominating factor in the Eastern tradition, the dynamic concept of the world, implies the third intermediate element. The two elements are not enough to describe the progression since there is something between the beginning and the end. Thus, the third element has to be introduced. The third intermediate element breaks the tension between the two and plays the role of mediator and opens the pair for development. The dichotomy between being and not being loses its contradictory character, since the changing nature of the world implies the concept of potentialities. That is, the process of change allows the potential qualities to become the actual ones. Consequently, the three nuclear elements need to be organized according to their importance. The order of the three categories of dominant, subdominant, and subordinant was developed in Japan to provide (1) a structure for every Japanese work of art and (2) a hierarchy of the Eastern image of the world. Interestingly, the Oriental hierarchy of dominant, subdominant, and subordinant has been incorporated into Western design for the sake of clarity of visual messages (Carter, 1985, pp. 52-56). In modern design, it is required that the visual message shows the degree of importance of its various parts. If everything is of the same importance and the image is complex, a viewer gets confused while not being given a direction for how to "read" the message encapsulated in the multiplicity of visual elements.

4. WESTERN ANTHROPOCENTRISM AND EASTERN SUBMISSIVENESS TO NATURE

Like anything else, the design elements have to be distributed in accordance to a certain order, so humans have to define their place in nature. Comparing the two traditions, it appears that there is a striking difference between attitudes towards nature. The Westerner's approach is arrogant when contrasted with the Eastern submissive one. Westerners have placed themselves in the center of the world and have given themselves the superior position over nature. The idea of "conquest of nature" comes from the Greek sources and, as illustrated by Protagoras' (i.480-421 BC) statement: "Man is the measure of all things." (Russell, 1959, p. 27), implies the hierarchy of two elements: conquering humans and the conquered rest of the world. The Westerner bends the forces of nature to his or her own will, regardless of the consequences. Westerners place themselves at the top of any hierarchy.

"In the East, however, this idea of subjecting Nature to the commands or service of man according to his selfish desires has never been cherished. [...] To look around for objects to conquer is not the Oriental attitude toward Nature." (Suzuki, 1959, p. 334). Easterners approache nature with a deep feeling of gratitude and appreciation. They submissively observe nature and learn how to coexist with the natural world harmoniously as its inherent part. The fifth tenet of Zen Buddhism expresses that attitude as follows: (p. 6th cent.) "Man rises from nature and gets along most effectively by collaborating with nature, rather than trying to master it."(ibid., p. 46). The Eastern intuitive and respectful experiencing of nature comes from the conviction that "...our Nature is one with objective Nature, not in the mathematical sense, but in the sense that

Nature lives in us and we in Nature." (Suzuki, 1959, p. 351). Consequently, "In Zen landscapes, the man-environment proportion shows a sound ecological relationship, in which no element dominates or damages any other." (Holmes, 1992, p. 49).

5. DEFINITIONS OF SYMMETRY AND HOLISTIC APPROACH

What is really meant by the term "symmetry"? The modern and commonly used notion of symmetry stands for "the identical disposition on either side of an axis or plane" (Ghyka, 1977), which is a drastically simplified and improperly reduced version, not only of the classical definition of symmetry, but also of formalistic mathematical classification, defining different types of symmetry. Therefore, what is commonly termed "symmetry" should properly be called "bilateral symmetry." The popular notion of symmetry applies only to the certain type of symmetry within the category of formal symmetry and is bilateral. On the other hand, according to the formalistic classification of symmetry, the notion of asymmetry is derived from the subdivision of symmetry and is named as "informal symmetry." The informality comes from the lack of any regular pattern of the distribution of elements that might define asymmetric arrangements. From the formalistic point of view, asymmetric composition is unpredictable and, as such, it introduces an element of uncertainty.

Let us look at the Vitruvius' (c. 1st century BC) definition to see how much it differs from the popular concept of symmetry: "Symmetry resides in the correlation by measurement between the various elements of the plan, and between each of these elements and the whole [...] As in the human body [...] it proceeds from proportion (called by Greeks "*analogia*"), and achieves consonance between every part and the whole. That symmetry is regulated by the modulus (ratio), the standard of the common measure" (Ghyka, 1977). The classical notion of symmetry is called dynamic symmetry because (1) it is related to the correlated proportions and (2) it implies the notion of the ratio, that is, the pattern of growth of the natural world (Korzeniowska, 1976, pp. 144-149). In natural numbers, the ratio is shown in the following proportion (13th cent.) called the Fibonacci series: 1:1, 1:2, 2:3, 3:5, 5:8, 8:13, etc. (Ghyka, 1977, p. xi, 7, 8, 13), The ratio is a number; it is an abstract concept created by the analytical, rational mind of the Western thinker, who imposes the ratio and the rational order onto the phenomena of the natural world.

The classical definition of dynamic symmetry is an expression of a holistic approach, viewing the world as a complex whole consisting of various interrelated elements in flux. That approach clearly has a long tradition in Western culture, although its impact is minor and, therefore, the classical definition was forgotten. Nowadays, there is a return toward holistic ideas, especially in aesthetics and in the formal training of art and design. For instance, according to the organicistic approach, "art is really a class of organic wholes consisting of distinguishable, albeit inseparable, elements in their causally efficacious relations which are presented in some sensuous medium" (Weitz, 1988, p. 27). By the same token, modern design is considered to be a matter of pure proportion, that is, how various elements can be made to work together to form a superior whole. Thus, the classical approach underlies the Western design, albeit only since the Gestalt theory of perception was conceived as the basis for modern design in the 1920s (Berryman, 1990, p. 9). The Gestalt theory is focused on the concept of the whole in relation to its elements and, as such, relates to the classical definition of symmetry. Nevertheless, it does so only partially, because it remains thoroughly unrelated to the other implication of that definition, that is, to the concept of growth. The ratio, as a Golden Mean of measure, is recognized as a valid principle in modern design, but without its relation to the growth pattern and, thus, without recognition of its derivation from the concept of the world as ever changing. Consequently, Western beauty and perfection is sought in the stability of perfectly balanced bilateral symmetry.

On the contrary, Eastern culture "ignores balance and inclines strongly towards imbalance [...] and embodies beauty in a form of imperfection or even of ugliness" (Suzuki, 1959, p. 24, 27). Consequently, the informality defining asymmetry is in favor, as is the perfection of reality in its imperfection. "All forms of evil must be said somehow to be embodying what is true and good and beautiful, and to be a contribution to the perfection of Reality" (ibid., p. 33). The search for perfect imperfections of the informally structured world comes from the holistic approach, being "inspired by the Zen way of looking at individual things as perfect in themselves and at the same time as embodying the nature of totality which belongs to the One" (ibid., p. 27). The Eastern culture "emanates from one central perception of the truth of Zen, which is 'the One in the Many and the Many in the One,' or better, 'the One remaining as one in the Many individually and collectively.'" (ibid., p. 28).

6. CONCLUSIONS

The search for the roots of the differences in aesthetic preferences between the West and Japan shows the deep influence of philosophical tradition upon our perception of reality itself and, thus, upon aesthetic preferences. To prevent rationality from limiting the spontaneity of the creative processes, I suggest to return not only to the origins of our tradition with its original definition of symmetry, but also to an implementation of the intuitive approach. Opening minds to different philosophical systems will broaden the aesthetic horizon and ensure creativity. The minds that merge rationality with intuition develop in the totality of their potentials. Finally, it is due to our individual and cultural conceptualizations that the world appears beautiful or ugly, and the world, per se, remains the same in its evolutionary movement and transcends any aesthetic categories.

REFERENCES

Berryman, G (1990) *Notes on Graphic Design and Visual Communication*, rev. ed., Los Altos, CA: Crisp Publications.

Brozi, K. (1983) *Antropologia Funkcjonalna Bronislawa Malinowskiego, (Bronislaw Malinowski's Functional Anthropology)*, Lublin, Poland: Wydawnictwo Lubelskie

Carter, R , Day, B. and Meggs, P. (1985) *Typographic Design: Form and Communication*, New York, NY: Van Nostrand Reinhold

Ghyka, M. (1977) *The Geometry of Art and Life*, 1st rev. ed., New York, NY. Dover Publications, Inc., orig publ., New York: Sheed and Ward, 1946.

Gunji, K (1990) *Through the Flower. Ikebana with Kimiko Gunji Book 1: Moribana*, Champaign, IL: School of Art and Design at the University of Illinois at Urbana - Champaign

Holmes, S. W and Horioka, Ch. (1992) *Zen Art for Meditation*, 13th ed , Rutland, Vermont and Tokyo, Japan: Charles E. Tuttle Company, of 1973 orig. publ.

Korzeniowska, M *et al.* (1976) *Przestrzen, Czas i Ruch, (Space, Time, and Movement)*, Warszawa. Panstwowe Wydawnictwo Naukowe

Russell, B. (1959) *Wisdom of the West: Historical Survey of Western Philosophy in its Social and Political Setting*, ed Paul Foulkes, s.l.: Crescent Books, Inc.

Russell, L. J (1914) *An Introduction to Logic From the Standpoint of Education*, London, UK: MacMillan and Co. Lmtd.

Suzuki, D. T (1959) *Zen and Japanese Culture,* New York: MJF Books, of 2nd rev ed , New York: Bollingen Foundation, distr. by Pantheon Books.

Weitz, M. (1988) "The Role of Theory in Aesthetics", In: Hardiman, G. and Zernich, T. eds., *Discerning Art: Concepts and Issues*, Champaign, IL: Stipes Publishing, 25-35. Originally published in. *Journal of Aesthetics and Art Criticism*, Vol.15 (1956), 27-55

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 315-333, 1998

"ORDO MUNDI" - ORDER OF THE WORLD AND STRUCTURE OF REASON IN RAFFAELLO'S "STANZA DELLA SEGNATURA"

Peter Klein

Address: University of Hamburg, Faculty of Education, Von-Melle-Park 8, D-20146 Hamburg. Tel. x49-40-4123-2151, Fax: .. - 2112

Abstract: *The "Stanze" are a flight of rooms in the Vatican Palace, ordered by Pope Julius II. to be ornated with frescoes by Raffaello Santi and his disciples. The second of them, called the "Stanza della Segnatura", was painted first, in 1508-12, by Raffaello himself. Housing the papal library for some time, the frescoes are usually understood as - though brilliant, yet rather conventional - representations of the common fields of study for a cultivated renaissance sovereign: science, Christian theology and the arts, all of which when working together are supposed to cause humane virtue in the studying person, as the images of the "Cardinal virtues" on the fourth wall indicate.*
Yet, the meaning of the images and their intentions reach much farther, thus making the room a climax of modern European self-understanding, its pride and failure: they are the program of a reasonable reigndom over the world, intending to do justice to Man and the world, proclaiming autonomy for the reasonable individual, who is guided, and protected, by the merciful self-revelation of God through Jesus Christ.
The threefold order of the world: the laws of nature, divine providence, and esthetic order, is reflected in the three components of human reason: cognition, morality, and esthetic judgment, as they were identified in the three "Critiques" of Immanuel Kant.
The complex esthetic qualities of the pictures, the interactions of their symmetries, thus are supposed to express both, the structure of the world and the structure of reason: symmetry gains a semantic meaning, a "spiritual dimension" transcending mere artistic beauty.

EUROPEAN HUMANISM: A CONCEPT OF SUSTAINABLE AND JUST DEVELOPMENT

In the current worldwide crisis of the environment, the term "sustainable development" figures as a key concept in environmental ethics and policy. It is meant to denote the final goal of human activities insofar they aim for ecological stability of nature and society. It intends a well-balanced long-term status of the world, that renders justice to Man and the world, for the benefit of both.

This idea is understood as a responsible reaction on the severe crisis and damage of nature - and in a more comprehensive sense: also of the crisis of the inner status of Man himself. This crisis is usually understood as the result of science-based industrialisation during the last two centuries, which, to its side, is supposed to be the result of the limitless (ab-)use of specifically European traditions of thinking and interpreting the world; in particular: by the technical application of Europe-born modern science, for the exploitation of the resources of the world, for the pretended benefit of Man - i.e.: esp. for the Europeans.

Since *European* thinking is supposed as the main source of this crisis, it is frequently expected to find salvation from that human corruption in turning to extra-European philosophical or religious traditions, which are supposed to have lived more in concordance with the laws of nature, take more responsibly into account the situation of Man as dependent from nature, are more peacable and humble, less imperialistic and exploitious.

This view, however, seems to be doubtful: research in the last years more and more has confirmed the suspicion that Man has always and in general been an essentially disastrous event in nature. We had to learn that all cultures had been on their, more or less effective, way to destroy their environment - if not European colonialisation had prevented them from finishing their cultural path into disaster on their own account.

On the other hand, I would like to defend the thesis that European philosophical and religious streams and traditions offer perspectives of cultural self-understanding, of responsibility and respect to nature. Though nowadays widely neglected, they teach us that the exploitive turn of reason is but the abuse of a *part* of full reason - usually called "instrumental reason" - and that by a sufficiently comprehensive concept of reason a system of incentives and criteria of human activities can be constituted that may be called "sustainable" in the above defined sense.

As usual in human culture, the ideas on which society is based are expressed by works of art, their symmetries giving symbolic reference of their structures. In a most comprehensive sense those ideas that incented European Renaissance and since characterize modern identity and self-understanding have been expressed by Raffaello Santi in the "Stanza della Segnatura".

THE "STANZA DELLA SEGNATURA"

The "Stanze" are a flight of rooms in the Vatican Palace which were inhabited by Pope Julius II, because he could not stand the idea of living in the old quarters below, now called the "Borgia appartments" that were so severely abused by his predecessor, Alexander VI "Borgia". So Julius, and his successor Leo X, ordered Raffaello Santi and his disciples to ornate the walls of the Stanze by frescoes. Of these, the frescoes in the second room, commonly called the "Stanza della Segnatura" because at times acts of justice were signed there, have been painted by Raffaello himself in 1508-12, most likely following a program of his friend and philosophical advisor, Fabius Calvus. As the room was also used as the papal library, the paintings usually are understood just as representations of the common fields of interest of a cultivated sovereign.

Instead – and this makes the room a climax of European identity and self-understanding – the images should be understood as the admonitory program of an ideal sacred reigndom, that intends to do justice to Man and the world. It is based on the Renaissance proclamation of individual autonomy and responsibility, and on an image of the structure of human reason identical with that later identified by Kant's "Critiques". It is the proud and self-confident spiritual program of modern (Renaissance) Europe, which in the future failed, however, caused by human weakness, arrogance and self-*mis*understanding - an ambiguity and failure of European culture of which this first room gives reference in advance.

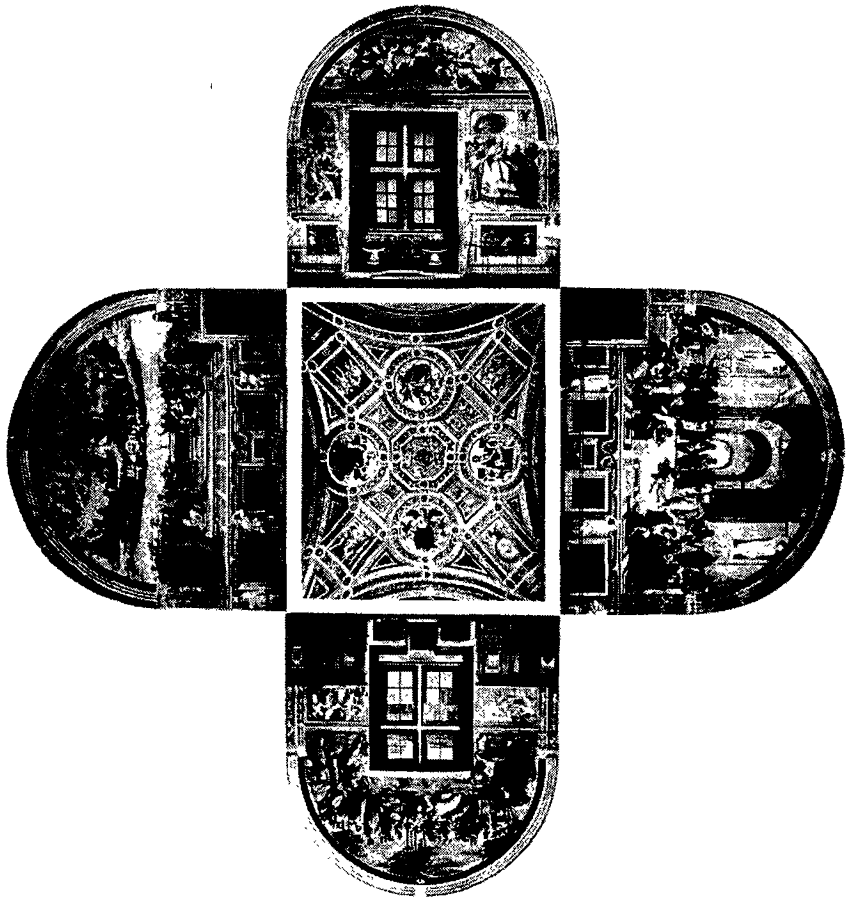


Figure 1

The *general structure* of the room's artistic ideas:

The pope is the ruler of the world insofar, as his mundane as well as his spiritual rulership are installed by divine order. As such, his judgementship over the things of the world should be based on the *Cardinal Virtues* (see IV). These, generated by profound eruditation of the three components of reason, must be grounded in:

– *Scientific Knowledge* about the laws of the natural world, and their comprehensive philosophical interpretation, including the image that Man has of himself (see I: "The School of Athens");

– *Religious Belief*, that embeds moral laws, as the guidelines of action, into a system of divine revelation and providence (see II: "The Disputá of the Holy Sacrament");

– *Artistic Beauty*, as the expression of a humane status of the world to be aimed for (see III, "The Parnassus"), and of the esthetically felt *Virtuous Welldoing* of Man (see IV again).

The seat of the pope is below the Cardinal virtues, as incentives of his reigndom. Science is to his left side, Religion to his right, and opposite are the Arts.

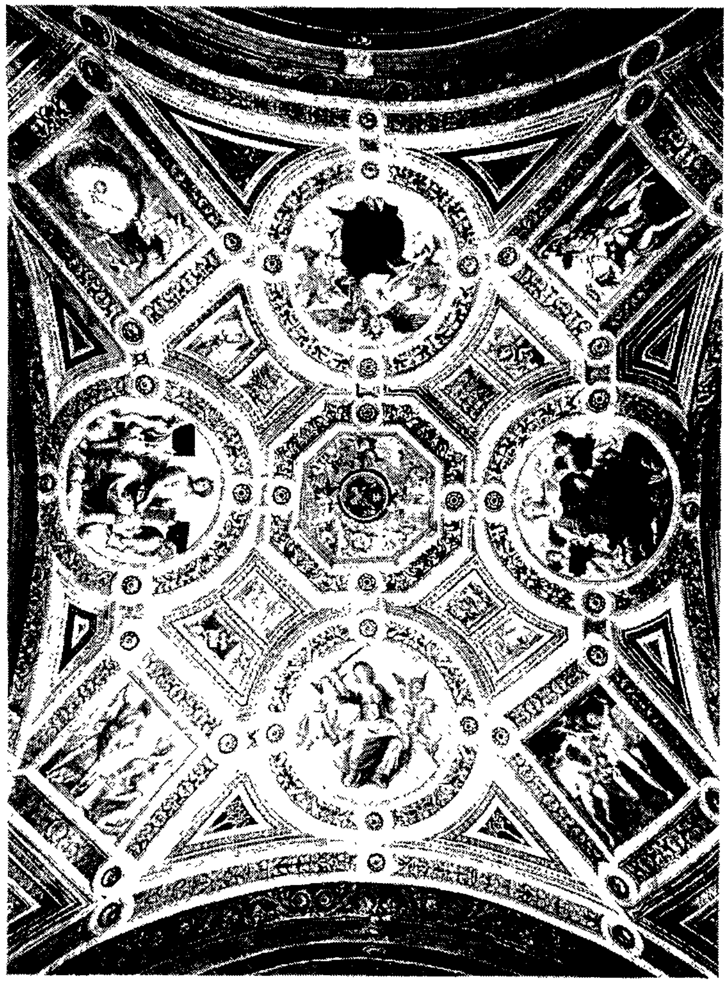


Figure 2

THE CEILING

Figure 2: The ceiling is a fourfold vault, exhibiting a central medallion, surrounded by four round symbolic medallions, right above the four wall paintings, and four rectangular pictures in the diagonal corners.

The *central medallion:* As became usual later in baroque vaults or cupolas, this central place is the location of transcendent or symbolic representation of the theme of the entire room. Thus, the central medallion, painted in the shape of the "lantern" of a cupola, shows the *papal arms* (= St.Peter's keys), thus referring to the pope's role as the mediator between heaven and earth, who is surrounded by angels, as divine messengers (the greek "*eu angelion*" meaning "good message").

The *round medallions:* Four symbolic *female figures*, also accompanied by angels, give emblematic reference of the spiritual gifts acquainted by the mental activities illustrated on the walls below:

– *Theology:* (motto:) "*divinarum rerum notitia*" = Notion of the divine objects

– *Philosophy (Science):* "*causarum cognitio*" = Cognition of the causes of things

– *Arts and Poetry*: "*numine afflatur*" = Touched by spiritual mysteries

These three, when acting together, render justice, enabled by

– *Jurisprudence:* "*jus suum unicuique tribuit*" = Law awards everybody his due

The rectangular *images in the corners* (see Figure 3): These fields of reasonable functions truly form a *network* of mental activities which is symbolized by connecting two of them, respectively, illustrated by appropriate mythological scenes (clockwise):

– *Adam and Eve (the Original Sin)* (relating theology with jurisprudence): *Justice is guarded and promoted by divine providence*

– *Salomo's Judgement* (relating jurisprudence with science): *Science-based human wisdom enables true justice*

– *A female genius studying the Zodiacus* (relating science with poetry): *Man reflecting himself by the universe in science and poetry (two books of nature)*

– *Apollo punishing Marsyas* (relating the arts (poetry) with theology): *Artistic arrogance without respect to God will be punished by cruel defeat*

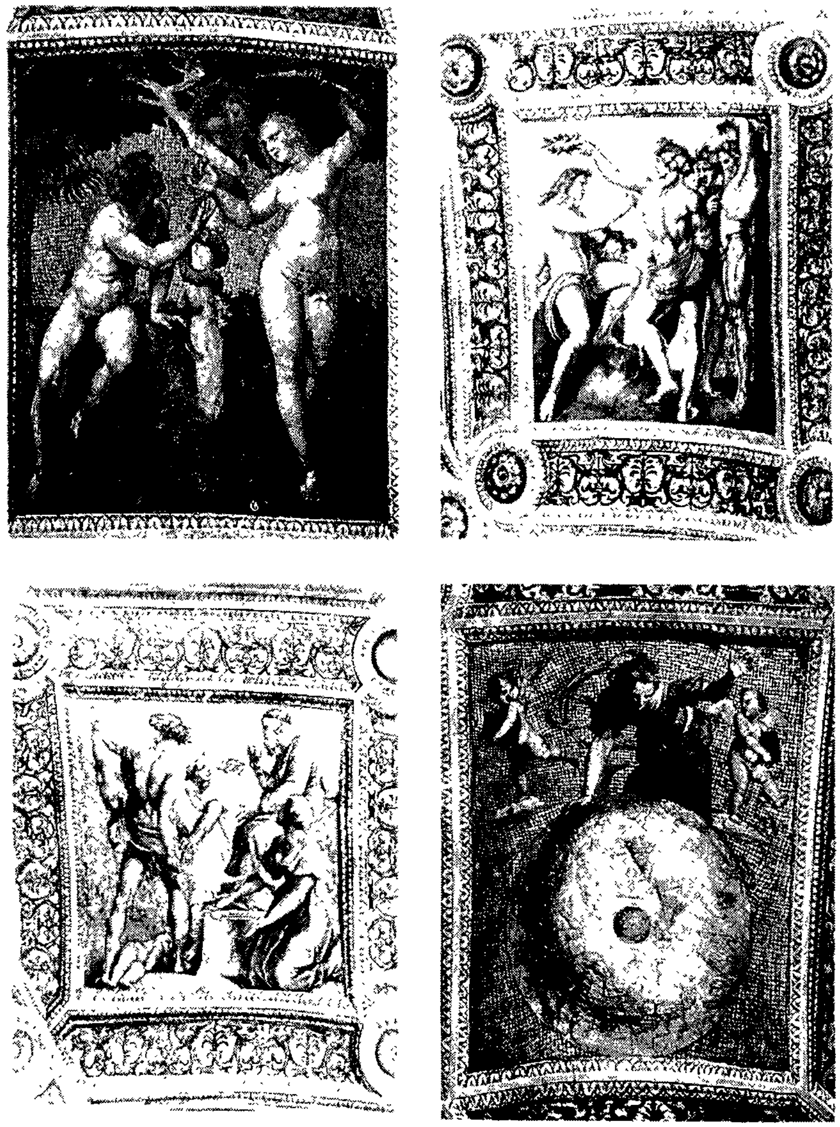


Figure 3

I THE SCHOOL OF ATHENS

Figure 4

Figure 4: In a vast hall, painted as a cassetted vault in the most modern manner by using central perspective (as Raffaello's friend, teacher and advisor Bramante had "built" just recently, as a mixture of architecture and painting, as the "*choir*" of San Satiro, Milano), we see an assembly of antique philosophers and scientists.

There was no strict cut between *science and philosophy* in those days; rather, science was supposed as rendering truth only if embedded into systematic philosophical contexts.

The motto above: "*causarum cognitio*", refers to both, modern *instrumental causal laws* as also to the four traditional Aristotelian *ontological "causae"*.

Two classic deities in the niches: *Apollon* (left) and *Athene* (right), supervise science by appealing to *good measure* and to *political relevance* of science, respectively, as crucial for true and useful science.

Few of the rich *personalities* have been identified, partly by tradition, partly by plausibility. Some at the same time also are portraits of contemporaries.

In the foreground, *to the left*: scientists representing esoteric paths in philosophy (Pythagoras, Heraklit, Averroes); *to the right*: applied sciences and education (Ptolemaios, Zoroaster, Euclid (an effigy of Bramante, demonstrating to his disciples, including Raffaello himself); *up the steps* (passing the demonstrative scepticism of Diogenes, as an obstacle against a simple ascent to truth; perhaps an effigy of Fabius Calvus), to the left once more the political relevance of science is indicated (Socrates and his disciples, esp. Alcibiades).

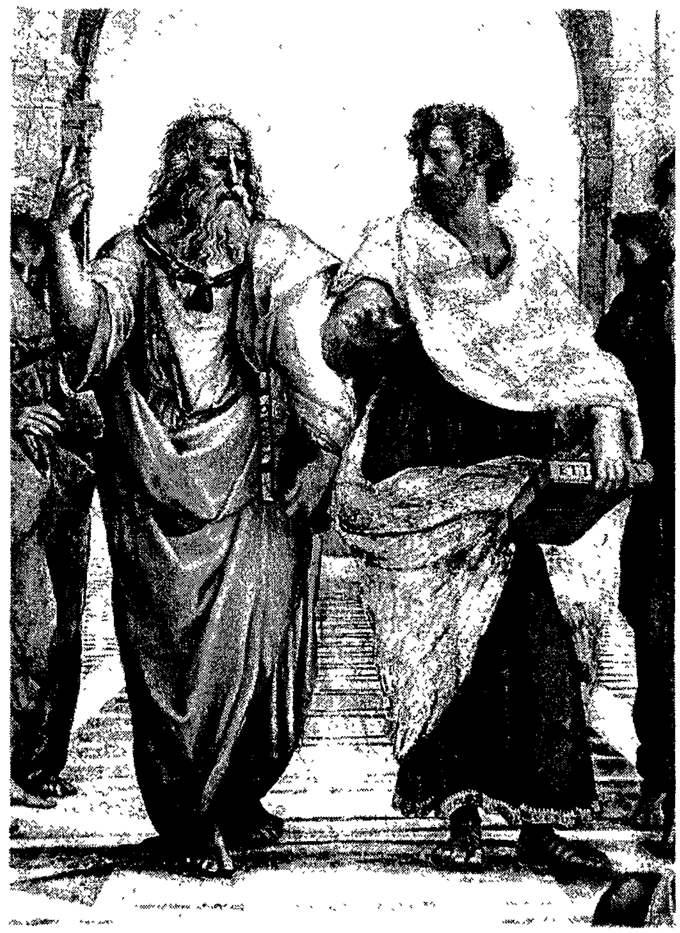


Figure 5

In the *center* (see Figure 5), the very representatives of European philosophy, *Plato* and *Aristotle.*

Plato is pointing upwards to heaven and eternity, in accordance with his philosophy of eternal ideas beyond the appearing world.

Aristotle, as the true founder of European science, by his demonstrative gesture is taking reign and command over the earth.

But *notice*: Plato does not bear in his hand one of his "idealistic" dialogues, but his "scientific", "*Timaios*", thus saying that these ideas are accessible only via science. Likewise Aristotle holds none of his scientific books, but his (Nicomachian) "*Ethica*", saying that science gains dignity only when aiming for goodness and virtue.

In general: Virtue must be the final aim of science, and there is no true virtue without scientific knowledge!

II THE DISPUTÁ OF THE HOLY SACRAMENT

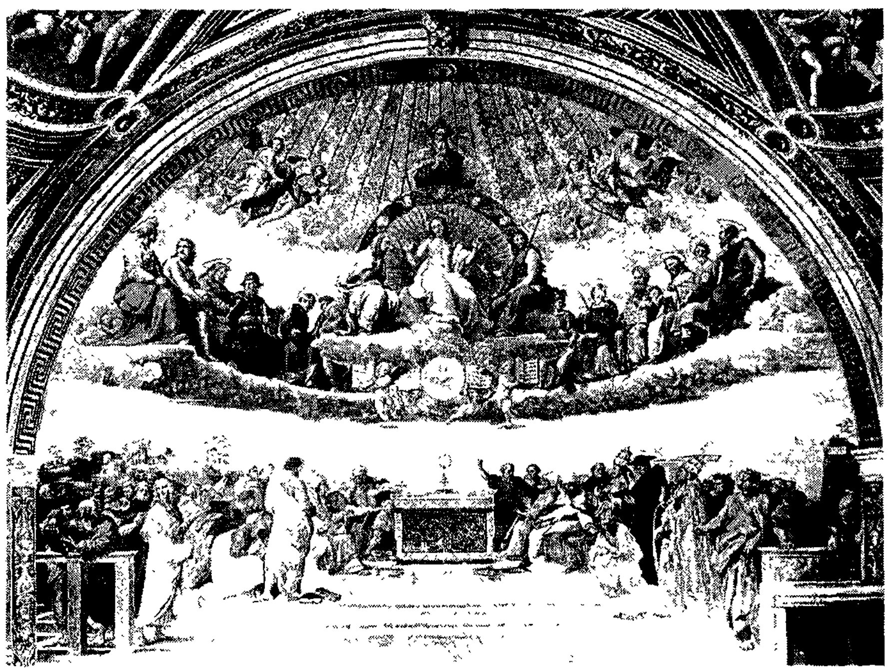


Figure 6

Figure 6: This picture is not so much a theological discussion, as the traditional title indicates, but a revelation of the *divine cosmic order and providence*, insofar it is guided and guarded by the three divine persons, respectively.

Consequently, the picture is divided into three spheres:

"Up in the highest" is the sphere of *God the Father*, who has *created* the world and preserves it "through the times". The structures and forces of cosmos, and the heavenly personage: the choirs of angels, whose hierarchic order had been elaborated by medieval theology, is supporting, appraising and revelling His creation, symbolized by a throughout symmetrical arrangement of human-like figures who, by their specific "sfumatic" peinture, clearly express their transcendent character.

The second sphere, still up in the clouds, and thus no subject of ordinary experience, yet participating in both spheres, the heavenly and the terrestrial, is the sphere of the revelation of divine providence with sinful mankind, as has been verified by salvation through *Jesus Christ.*

Closest to Him, St. Mary and St. John the Baptist, then persons from the Old (Adam, Moses, David) and from the New Testament (Apostles, Evangelists, with (right and left) St. Peter and St. Paul as "cornerstones" of the church). All these had been participants and direct witnesses of the historical event of salvation (therefore, the traditional identification of Nr. 7 as St. Lawrence certainly is false).

Divine salvation continues to penetrate our lives on earth by being revealed to us through witness of the four *gospels*, presented by angels, as attendants of the *Holy Spirit*, who acts as the mediator between the world "above" and "below", and will be with us "from now on till the end of the world".

Through all these times God is really present within the world in the shape of the *Holy Sacrament*, which thus becomes the focus point of all perspective lines that regulate and organize the earthly sphere below.

Now religious belief does not so much command us to "discuss" the presence of God, but to cope for understanding and for becoming agreeable to His eyes. The *Holy Fathers*, sitting around the sacrament, have given authoritative exegesis of this secret. But, according to catholic view of the importance of "tradition", also the *theologists* and *priests* (to the right) play their specific part in this never ending spiritual enterprise, as do also the *sceptics* (foreground, right) and the *heretics and gentiles* (foreground, left):

Jesus has dedicated His salvation to all mankind, and speaks to them "in different tongues".

Note how these three different spheres show a more and more "broken" symmetry, the more we descend from "heaven" to "earth", ending up with a nearly chaotic, yet vivid arrangement of the persons to the left: the clear and eternal "*logos*" of the world is hidden and gets confused by the human efforts to understand it by discussion, instead of trustful belief in His revelation.

III THE PARNASSUS

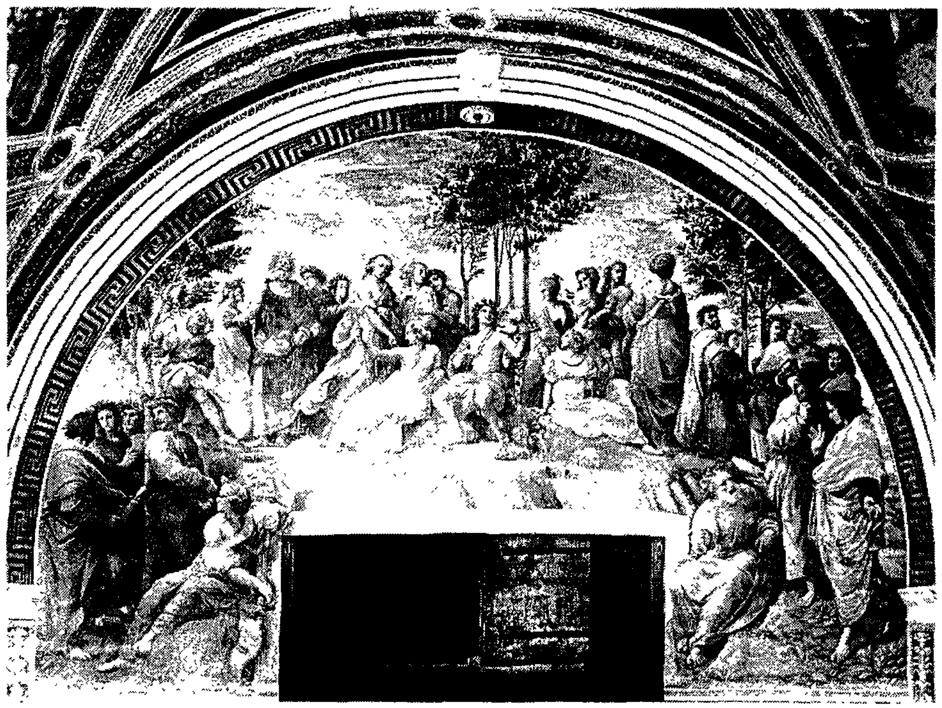


Figure 7

Figure 7: This picture perhaps may appear as the most conventional, though certainly it handles the complicated problems of the place on a remarkably high artistic level, in that typically "renaissance" style, relaxed and vividly "*humane*", yet "classical", which Raffaello mastered first and was famous for.

We see an assembly of *poets*, all characterized as "classical" by being coronated with a laurel-wreath (the left person is just a writer listening to the dictation of blind Homeros). Few of them can doubtless be identified (besides Homeros only Sappho and Dante). Some count as known by literary tradition; ancient authors are mixed with contemporaries.

The full range of the arts is present by their symbols, the nine *Muses*, gathering around *Apollon*, who sits beneath the Kastalian Well, creating the harmony of the spheres by playing an old-fashioned violone.

All this looks like a very standard illustration of the literary formative process of Renaissance humanism that focused so much on reading the classical authors. To understand the full programmatic importance *in our context*, we must refer to the book that explicitly elaborates this view, the book of Raffaello's friend, Baldessare Castiglione, "*Il Cortegiano*" (The Book of the Courtier) which exhibits the modern renaissance ideal of a truly cultivated and educated person, able to act in the world with scientific, moral and artistic responsibility, as an autonomous individual personality. Though printed only in 1528, its considerations go back to the conversations held about 1507 at the court of Guidobaldo da Montefeltro and his spouse, Elisabetta Gonzaga, at *Urbino*. This court, since the reign of Guidobaldo's father Federigo counted as the birthplace of renaissance humanism and as the ideal of a truly humane court, famous as singular in Italy for the absence of any crime, murder, conspiracy and immorality. Here, Raffaello, Castiglione and Bramante had met first, and also the later popes Julius and Leo at times had lived there.

These discussions at Urbino now in fact unfold the *ideal of a just reigndom*, based on true humanity, as was supposed to be found in the works of the classical antique authors, as "scriptures", equivalent and complementary to the Holy Bible. This ideal is characterized by continuous coping for moral virtue that comes into being - and is indicated - by the comprehensive beauty of conduct of life, that verifies the ideals of antique classics.

Thus, beauty is not an adversary of moral goodness, but rather expresses it; in fact, they are a unit of a whole: "Das Schöne ist das moralisch Gute in der Erscheinung" (Schiller).

IV THE CARDINAL VIRTUES

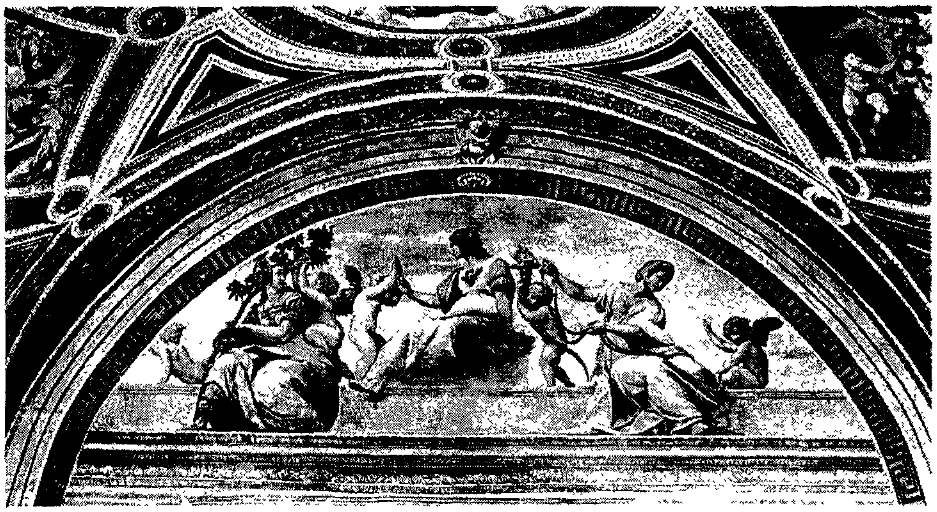


Figure 8

Figure 8: Having these three dimensions of reasonable mental activities in mind, their impact onto the human mind is supposed to cause the Cardinal Virtues, as continuous incentives of virtuous action in the world. So whenever the pope sits in the window bay, deliberating on political or spiritual problems, they are above him, acting as appealing guides.

The classical *Cardinal Virtues:*

– *Fortitudo* = courage or strength (with the attributes: oak-tree and lion)

– *Temperantia* = moderation or temperance, also self-control (curbs of passions)

– *Prudentia* = prudence or wisdom (torch of enlightenment, mirror of self-knowledge, also showing the "Janus"-face of knowledge: looking into the dark)

These three virtues, denoting *habits*, according to Aristotle are synthetically comprised by the fourth, that marks the resulting *manners* (hence is painted above them all, on the ceiling):

– *Justitia* = justice or righteousness

KANT: A COMPREHENSIVE THEORY OF THE STRUCTURE OF REASON

Up to now, we have interpreted Raffaello's pictures as an artistic expression of the renaissance image of the true order of the world, which in its threefold dimension can be understood by the human mind in a rational way, as science, as religion, and in the arts. But how is it possible to be convinced of the truth of this order, which in its core is based on metaphysical ideas about the structure of nature, of the transcendent world, and of artistic beauty. Obviously we must study the structure of our mind as well, hopefully expecting to find our mind well adapted to the world in which we are supposed to act.

The German philosopher Immanuel Kant has done this study in his three "Critiques", finally claiming that he has rendered an empirical description of the full and general structure of reason, independent from any cultural specifications. He discriminates three components of reason which in fact are coordinated to the three dimensions of the world mentioned above.

The *"Critique of pure reason"* describes the *cognitive* basis of instrumental (technical) action of Man in the world, stating that cognition is not a sufficient tool for responsibility in action because it principally cannot give us truth of the world, since it is always limited by the specific structures of cognition at all, namely to sensual data and their cognitive promotion by the logical laws. Especially, cognition is unable to give us knowledge about those metaphysical ideas "behind" the sensual data, which classically acted as foundations of cognitive truth.

Hence a second field of reason, specific *postulates of "practical reason"*, traditionally identified with moral laws and usually bound to a religious system, must be added to cognition, as the sources and guaranties of welfare and eternal salvation of Man: the postulates of *"freedom"* (as the condition of any responsibility in action), of *"eternal life"* (as the basis of aiming for moral perfection at all), and of *"God"* (as the guarantor of the compatibility of natural and moral laws).

Yet, Kant shows, that also these postulates are not sufficient for reasonable action because they in fact are just postulates, meant to fulfill the human aim for acting "free" in a philosophical sense, but grounded in a transcendent world independent from the world of sensual experience, in which we are supposed to act.

Hence, a third field of reasonable mental activity must be considered, called *"power of judgment"* that has to connect the transcendent with the empirical world. Judgment is guided by general ideas about the systematic, *"teleological"* structure of the world, and by the idea of an *"esthetically"* felt well-being of Man within the world. Thus, a just, responsible, and sustainable action of Man is a balanced system of cognition, moral laws, and esthetic judgment, enabling virtuous self-control and self-eruditation.

This Kantian system of the structure of reason is only a description of the *formal* structure of reason. In each "culture", as the concept of a manifold of human beings unified by their common image of the world, it is concretely verified by specific cultural ideas about the world as a whole, and of adequate human behaviour in it. In European Christian-based culture the laws of nature and of human morality both are supposed to have been created by God, who, as Christians believe, has created them as an image of His divine essence. Thus, a kind of isomorphic structure unifies them all, as the medieval "*doctor mirabilis*", Albertus Magnus, stated in a most condensed manner: *"Tribus ordinibus ordinatum est universum: scilicet in se, et ad hominem, et ad Deum creantem."* (i.e., Threefold is the order of the universe, namely: - in itself, - with respect to man, and - with respect to God, the creator.)

The "Stanza della Segnatura" perhaps is the most comprehensive, and also the most beautiful, illustration of these ideas insofar they are considered as mental guidelines for modern European, Christian Man and his culture.

AMBIGUITY AND CATASTROPHE

We had not yet finished our interpretation of the Stanza, but had left it at the point that the cardinal virtues promise the ability of justice in human action. Exactly at this point, however, the tragedy of the "Stanza della Segnatura" - and symbolically: the tragedy of Renaissance Europe - started by an interruption: by the corruption of Raffaello's program, - an event tragically figuring as a symbol of the failure of European spirit of enlightenment, and of the European concept of individual autonomy at all.

We do not know how Raffaello intended to complete this last wall (below the lunette of the "virtues"), and thus his whole program. But we know what happened instead: When the pope returned from a war campaign, he ordered that two historical scenes should express *his* interpretation of justice:

– Cesar Justinianus donating the *Roman Law* to Trebonianus (Fig. 9a)
– Pope Gregorius IX donating the *Canonic Law* to St. Raimund of Penaforte (Fig. 9b).

Figure 9a

Figure 9b

At first glance these themes quite innocently illustrate two representative historical events that have founded justice in the mundane and in the spiritual sphere. Yet, they are a severe mistake:

"Justice" originally meant the comprehensive destination of mind that overarches knowledge, morality and beauty, thus enabling individual responsibility, now has become a matter of *merely formal law.*

Even worse: The effigy of pope Gregorius was ordered as a portrait of pope Julius himself: thus *self-celebration, diplomacy, and politics take command* (and the following Stanze confirm this decline of spirituality: Raffaello left them to his disciples, disinterested in what would follow after his general idea had been corrupted.)

Up to this point, the program had intended images of pure (infinite) ideas, now they are interpreted as mere (finite) history.

The pope pretends to do already what originally he was appealed to do, as an ideal:

Humble endeavor for ideals is replaced by the arrogant self-confidence of modern Man - a symbol of the continuous treachery of reasonable Christian ideals, that we commit every day by our irresponsible conduct of life.

It is the "sin" of Marsyas: proclaiming ourselves as divine, instead of striving modestly after God's benevolence; and like Marsyas we will be punished by getting stripped our skin.

Hence we agree with Wilhelm Kelber who (in his book on Raffaello) summarizes the ambiguity of the Stanza, that *Raffaello "on these walls has documentated the crucial and tragical crisis in the history of human spirit",* from which we do suffer now, and perhaps will perish.

Symmetry: Culture and Science
Vol 9, Nos. 2-4, 334-345, 1998

HUMAN FLUCTUATING ASYMMETRY: WHAT DOES IT MEAN FOR PSYCHOLOGISTS?

Rotem Kowner

Address: Department of Multidisciplinary Studies, The University of Haifa, Haifa 31905, Israel; tel · (+972) 4-824-2559, fax: (+972) 4-824-9155; e-mail kowner@research haifa.ac.il.

INTRODUCTION

In recent years researchers from diverse disciplines in the life sciences such as biology, genetics, physical anthropology, medicine, behavioral ecology, as well as several domains in psychology, have turned their focus onto the role of developmental stability of the organism. Developmental stability, sometimes referred to as developmental homeostasis, can be defined as the ability of the organism to resist or buffer the disruption of precise development by environmental and genetic stresses. When instabilityoccurs, it may be marked by relatively conspicuous morphological or behavioral errors, but more often by only minor random deviations from bilaterally symmetrical traits called fluctuating asymmetry (FA). Although the concept of FA was introduced more than 60 years ago (Ludwig 1932), it has entered into the limelight of scientific research only in the last decade.

The general assumption regarding FA, whether implicit or explicit, is that it reflects developmental stability, which is in turn correlated with Darwinian or overall fitness. Fitness here refers to the individual's own survival as well as the relative number of offspring contributed by her or him to the next generation. Sources of instability, namely environmental and genetic stresses, appear to diminish fitness since they often increase the degree to which random growth perturbations are magnified, and may further result in disruption of the developmental check mechanisms. Genomes less sensitive to such disruption are said to *canalize* development; that is, to suppress phenotypic variation (Waddington 1957). Thus, individuals better resistant to such stress may show greater fitness as well.

There are several reasons for the growing fascination with FA. First, it measures minor genetically and environmentally induced departures from a perfectly designed developmental program, and thus it provides a gauge to the organism's developmental noise, and indirectly to its viability fitness. Compared with other indices of development and fitness, FA has two important advantages: Not only it is easy to assess it, but it measures aberrations from a known morphological ideal, that is, perfect symmetry (in most organs). In addition, the study of FA provides some insights for researchers of genetics and the environment, as well as those interested in the interaction between the two. Both the likelihood for the appearance of FA in an individual and the degree to which it appears has often some heritable basis, and interestingly, it is associated with two key issues in genetics: It often decreases with protein homozygosity but increases with hybridization. In addition, studies on the change in FA due to effects of the interaction between environmental and genomic stresses may prove useful as a monitor of the disruption of developmental stability at an array of integration levels from molecular to the epigenetic. Finally and more specifically for psychology, a number of recent studies have indicated that FA is correlated with various behavioral and psychological phenomena ranging from intelligence, depression, and mating behavior. Hence, some believe, this research direction may shed light on the biological basis of human behavior and foster an integration between several research domains.

All multicellular organisms exhibit a high degree of bilateral symmetry in many of their organs. Nevertheless, many organisms, or even parts of a single organism, may show deviations from perfect symmetry. These deviations can be divided into three types of asymmetry, each characterized by a different combination of mean and variance of the distribution of differences between the right-minus-left plane (R - L) of a bilaterally symmetrical trait (Van Valen 1962). The first type, *directional asymmetry*, refers to a normally greater development of a trait on one side of the plane of symmetry than on the other side (e.g. the human heart and brain). The second type, *antisymmetry*, also refers to a normally greater development of a trait on one side of the plane of symmetry, but in contrast to the previous type it is unpredictable which side of an organism shows greater development. Thus, antisymmetry is characterized by a bimodal distribution of R - L differences, often with a mean of zero (e.g., the claws of the fiddler crab; the right and left hands in humans when measured by preference scales).

The focus of this review is on the third type of asymmetry, called *fluctuating asymmetry* (FA). FA denotes randomly produced deviations from perfect symmetry of two sides of quantitative traits in an individual for which the population mean of R - L differences is zero and their variability is near-normally distributed. This type of asymmetry arises in the course of development, and is called 'fluctuating' because its direction is not under genetic control. In fact, the subtle departure from symmetry in the individual ought not to be stimulated by genetic factors (Palmer & Strobeck 1992).

FA is basically used to measure either the effect of asymmetry per se on a certain trait or performance, or to assess the variation between either populations or individuals. Typical performance studies seek to understand the effect of asymmetry itself, whereas typical variation studies usually examine the effect of genomic or environmental causes on a number of populations or a group of individuals, using independent variables such as time span or location.

The FA found in 'normal' population, regardless of the species, is usually about 1-2% of feature size, while exposure to stress may lead to a higher levels (about 3-5%) of FA (Parsons 1990). Although humans are bilaterally symmetrical in essence, researchers have noticed the presence of small morphological asymmetries in various bilateral traits located in the body and the face. The right side has more departures from the fairly symmetrical morphology found in normal people, and this is true even when antisymmetry (handedness) is controlled (Van Dusen 1939; Garn, Mayor & Shaw 1976; Kowner 1995). In its non-pathological form, asymmetry in humans can be defined as lower than the mean ± 2 SD of a specific bilateral trait (Livshits & Kobyliansky 1991).

There is growing evidence that FA, as a marker of developmental stability, may help researchers in various domains of psychology. It may be related to plethora of psychological phenomena, and may serve as a key for future theories attempting to link a wide range of psychological themes such as mental and physical maldevelopment, behavioral variance, as well as issues concerning the interplay between environment and heredity. In the next part the relevance of FA to several psychological domains is examined.

CLINICAL PSYCHOLOGY

There is growing evidence for the importance of FA as a marker of mental health, either as an indicator of general well-being or as a predictor of severe mental disorders. The most investigated link between FA and any mental pathology concerns psychotisicm, and particularly schizophrenia. Markow and Wandler (1986) revealed that schizophrenic patients have greater dermatoglyphic FA compared with controls and that the severity of the symptoms correlates with the magnitude of the FA. In fact, schizophrenics exhibit greater FA than controls in almost any dermatoglyphic traits (see also Mellor 1992) as well as high frequency of deviance in various forms of hand morphology (Shapiro 1965). The research on FA may shed some insight on the etiology of schizophrenia. Markow (1992) proposed that schizophrenia reflects abnormal balance of symmetry and asymmetry in the brain as a result of reduced developmental buffering associated with homozygosity. The increased homozygosity may also lead to increased developmental instability in the central nervous system, which results, in turn, in deviant behavior (Markow & Gottesman 1989).

A related phenomena which may elucidate the role of FA is the development of minor physical anomalies (MPAs). Found in less than 4% of the general population, these structural features of no functional significance (e.g., wide-spaced eyes, low or malformed ears) have been used widely to infer the degree of embryonic developmental instability of the individual. (Hoyme 1993) MPAs are usually attributed to injury or atypical ectodermal differentiation during the first or second trimester of fetal life (Murphy & Owen 1996), but some may stem from slowed development of specific traits in time-locked development. (Yeo, Gangestad & Daniels 1993) There is a significant body of evidence regarding the increased prevalence of MPAs in schizophrenia (e.g., O'Callaghan, Larkin & Kinsella 1991), particularly familial schizophrenia (Waddington, O'Callaghan & Larkin 1990), and in autism (e.g., Campbell, Geller, Small, Petty & Perris 1978). In fact, there is a direct indication for the relation between FA and MPAs. Green, Bracha, Satz & Christenson (1994), for example, showed that MPAs in schizophrenia are associated with increased dermatoglyphic FA.

PERSONALITY AND SOCIAL PSYCHOLOGY

Future studies on the role of developmental stability in humans may also open a new window to the interpretation of personality, ranging from behavior, emotions, to cognitive skills. Although only a few studies have examined the role of developmental stability in these domains in humans, the strong impact which developmental instability exerts on general development suggests it may also play a role in human behavior, communication, and emotional state. Despite its limited comparative value, there is extensive evidence from studies on these aspects in non-human animals which attest the importance of developmental stability in affecting and moderating a wide range of behaviors. Multivariate analyses of courtship behavior in wild type drosophila (Markow 1987) revealed that phenodeviant males (namely, individuals suffering from sporadic occurrence of abnormal morphological deviance measured by FA) show aberrations in courtship component sequences, level of performance within each sequence, and appropriateness of courtship delivery. As for humans, Waldrop and Halverson (1971) reported greater frequency of MPAs, another manifestation of developmental instability, in children exhibiting a range of behavioral abnormalities.

One of the mechanisms which links behavior and stress was sketched by Alados, Escos & Emlen (1996) who examined changes in behavioral patterns of Spanish ibex with parasitic infection or during pregnancy. Alados et al. contended that biological structures and behavior patterns have evolved to allow the organism to explore its environment and enhance its tolerance for changes. Natural selection ought to increase complexity in behavior to maximum levels consistent with energetic constraints. Stress, however, increases metabolic rate which entails energy consumption, and consequently a reduction in behavioral complexity. Hence, as a measure of stresses, increased FA is hypothesized to predict altered behavior and increased emotional state.

FA is also related to dominance, a behavioral aspect of fitness which is associated with access to resource and ultimately with survival and reproduction. Dominant individuals among male European starlings, but not among females, were found to exhibit invariably lower levels of FA than subdominants (Witter & Swaddle 1994). Likewise, the degree of FA of antlers of the fallow dear, as well as their height, was found to be an important predictor of dominance among males (Malyon & Healy 1994). The negative relation between dominance and FA is found also in gemsboks, but in this species dominant individuals of both sexes exhibit lower FA of their horns (Møller et al. 1996). Finally, FA appears to be associated with dominance also in humans. Recent studies suggests that symmetric men are more socially dominant according to both self-report and report by romantic partners (Gangestad & Thornhill 1997a), and similarly that boys with low FA (10-15 years old) are more aggressive than their high-FA counterparts (Manning & Wood 1998).

Facial FA may also be related to one's emotional states. Shackelford and Larsen (1997) who conducted the first study regarding this association suggest facial FA may signal not only physiological stress but also psychological and emotional distress. Using multiple evaluations (e.g., self-reports, observer ratings, daily diary reports, and psychophysiological measures), they found facially asymmetric men to be more depressed, more emotionally labile, and more impulsive than relatively symmetric men. Asymmetric women experienced more muscle soreness and were also more impulsive than more symmetric women. Although the samples Shackelford and Larsen used were small and homogeneous and their findings often inconsistent, this study raises important questions regarding the etiology of emotional states and psychological stress. Critically, it moves beyond social judgments of people displaying particular facial characteristics, since subjects with lower facial FA were rated as emotionally stable not only by observers, but also rated themselves as such.

DEVELOPMENTAL PSYCHOLOGY

The research on the effect of developmental stability on human development is probably the earliest within this emerging domain. The main contribution of this research is in indicating sources of genetic and environmental stress which affect the individual during development. In contrast to other morphological phenodeviance which often occurs during a limited period of development FA occurs throughout prenatal growth and continues to alter after birth and into adulthood.

Studies show that different levels of developmental stability start to affect human development in very early stages. Babler (1978), for example, found that aborted fetuses in the second quarter of pregnancy had greater frequency of dermatoglyphic abnormalities than either older fetal abortuses or normally newborns. There is an association between certain developmental delays and high FA. Preterm newborns, for example, exhibit greater FA than term newborns, and there is an inverse correlation between the FA of infants and their gestational age as well as their health status (Livshits et al. 1988). Likewise, pre-pubertal children with delayed development show greater dermatoglyphic FA than controls (Naugler & Ludman 1996), and greater degree of FA is also associated with spinal deformity (Goldberg, Dowling, Fogarty & Moore 1996) or miscellaneous multifactorial illnesses, such as cleft lip (Woolf & Gianas 1976).

There appears to be a correlation between parents' bodily asymmetry and the level of asymmetry in their offspring, and a similar correlation between infants' body symmetry and their parents' number of infectious diseases during pregnancy (Livshits & Kobyliansky 1991). The behavior of the pregnant mother may affect the development of her postnatal infant as well, and several studies identified smoking, consumption of alcohol and obesity as maternal liabilities which have detrimental effect on the developmental stability of the newborn (see in detail in the section on environmental sources of stress). These effects notwithstanding, the socioeconomic status of parents was not found to affect the FA of Israeli infants (Livshits et al. 1988), older children in Japan (Dibennardo & Bailit 1978), or American college students (Thornhill, Gangestad & Comer 1995).

FA appears to be a useful marker of development also because it shows constant transformation throughout life span, arguably with higher points of bodily and facial FA values to emerge in a very young age, during puberty, and notably during old age. Livshits & Kobyliansky (1989) found full-term newborns in Israel to show greater bodily FA than older children (age 5-18 years), whereas two non-longitudinal studies found various levels of increase in bodily and facial FA among adolescents as compared with younger children (Wilson & Manning 1996). The increased developmental instability in puberty is probably the result of increased overall body growth as well as the development of secondary sexual traits. The appearance or growth of these traits at this period occurs under the influence of androgens and estrogens and is accompanied by temporary reduction in the immunocompetence of the individual due to the hormonal surge (e.g., Grossman 1985). Evidently, increased energy demands for body growth and immunocompetence lead to less energy for maintaining developmental stability. Still, it is possible that normal growth results in temporary FA that is of no long-term physical, psychological or clinical significance.

As a measure of the general buffering capacity of an individual's ontogenetic development, high degree of developmental instability in humans is linked to developmental abnormality. The genetic component of various multifactorial congenital anomalies not only results in malformations but also reduces resistance to adverse environmental influences. It has been repeatedly demonstrated that individuals suffering from various developmental anomalies, congenital conditions, or illness also exhibit greater FA in various traits (see Thornhill & Møller 1997). Likewise, greater dermatoglyphic and tooth FAs often characterize various genetic syndromes such as Down syndrome (e.g., Barden 1980), trisomy 14 syndrome (Fujimoto, Allanson, Crowe, Lipson & Johnson 1992), fragile X (Martin-Bell) syndrome (Peretz et al. 1988), and Goltz syndrome (Landa, Oleaga, Raton, Gardeazabal & Diaz-Perz 1993), and even mentally retarded individuals show greater anthropometric FA than normal individuals (Malina and Buschang 1984). These studies indicate that health and pathology are related to phenotypic and genetic quality, and that developmental instability may serve as a measure for this quality.

EVOLUTIONARY PSYCHOLOGY

The psychological study of human mating patterns and sexual behavior has gained much ground in recent years (e.g., Buss & Schmitt 1993; Gangestad & Simpson 1990; Sprecher, Sullivan & Hatfield 1994). The new psychological perspective on human sexuality is based on a synthesis between two interdisciplinary domains: evolutionary psychology and social cognition (for review, see Kenrick 1994). The former domain, which is the scope of this section, assumes that human mating behavior has biological roots which are found in all the evolved psychological adaptations underlying behavior in general, such as fitness (as its ultimate cause) and parental investment (as its main proximate cause), and thus we may find its fairly consistent pattern across cultures.

Evolutionary psychologists, among others, refer to the reproductive success of an individual, compared with that of all other individuals in the population, as relative fitness. This reproductive success depends importantly on access to mates, which is determined by sexual selection. This important mechanism of evolutionary change occurs because individuals of a species choose to mate with certain other individuals due to some traits. Sexual selection has two main forms: The first form is intrasexual selection (i.e., competition between males) and the second intersexual selection (i.e., female choice). FA is probably linked to sexual selection through the development of secondary sexual traits, a process which is mediated by enzyme heterozygosity and its effect on general developmental stability (Mitton 1995). It has been argued that epigamic structures and weapons of an organism should show higher levels of FA than that found in non-sexual traits, because sexual selection is essentially directional. Further, since FA is moderately heritable, reflects general fitness, and occurs in numerous features which affect reproductive success, it seems reasonable to assume that sexual selection would favor low levels of FA, and that individuals would be able to assess the level of FA of their potential mates.

To examine these assumptions, many studies investigated the relation between the degree of FA of individuals and their success in acquiring mates (see meta-analysis in Møller & Thornhill 1998). This relation was first examined in regard to female choice. Few researchers doubt such choice of males occurs in nature for access to resources or parental care. Yet, is there female choice of males who exhibit lower level of FA? A number of studies with insects (e.g., McLachlan & Cant 1995; Thornhill), and birds (e.g., Møller & Pomiankowski 1993) showed that female individuals prefer males who show lower FA in natural conditions. Forewings that differ in length by just a fraction of a millimeter, Thornhill found, may keep a male scorpion fly from finding a mate. As for birds, Swaddle and Cuthill (1994) manipulated the degree of FA in secondary sexual plumage trait in male zebra finches in order to examine the preference of birds for symmetry in non-arbitrary traits. Indeed, they revealed that females of this species prefer to perform ritualized courtship display-jumps in front of males wearing symmetrically manipulated chest plumage. This description may fit mammals as well. FA in the horns of gemsboks appear to be negatively related to the number of offspring in females and mating success of males (Møller et al. 1996).

Although there are several potential explanations for this choice, the majority of researchers seem to prefer the "good genes" model of sexual selection which posits that females choose males with the least FA because they reflect a heritable quality that affects offspring viability (for review on models of sexual selection, see Gangestad & Thornhill 1997). Inherent in this model is a close association between the indicator (FA) and fitness and that both are heritable (for criticism of the proponents of this model, see Markow & Clarke 1997). Polak (1994) offered an elaborated version of this approach, suggesting that reduced FA in males may indicate greater resistance to parasite, a trait that may pass on to their offspring. In humans, Manning, Koukourakis & Brodie (1997) found positive relation between resting metabolic rate and FA in male college students, but not in females. Viewing sexual selection as any other stress, Manning et al. suggested that high-quality males are better able to withstand the pressure of sexual selection since they are able to allocate more energy to growth and reducing FA than low-quality males.

FUTURE PROSPECTS

In spite of the many flaws and imperfections that seem to prevail in the field, the current state of research suggests that developmental instability is related to plethora of psychological phenomena. The interest in FA as the main maker of developmental instability and the consequent surge of research on this topic has brought about much expectation as well as slight disappointment. By now, researchers possess much greater data than they had a decade ago, and thus we may reasonably assess the potential of this measure.

Developmental instability, as reflected in FA, shows overall negative correlation with the overall fitness of the organism as well as various indices of health and performance in humans. Although this relation is promising, there are still some hesitations as to the widespread application of FA as a major substitute for conventional fitness measures (cf. Clarke 1996). Within humans, FA in several traits shows a relation to various disorders of both developmental and genetic etiology. However, its role as a general risk marker for pathology needs to be further clarified. This may become possible with the elimination of several methodological problems (e.g., neglect of measurement errors, lack of standardization of FA measurements, no reports of odds ratios of various morphological traits for each pathology, measurement of many traits rather than one or a few) as well as further evidence about the occurrence of FA and its relation to effects of other types of symmetries in both sexes.

With the influx of research we may expect in the coming years to know more about the advantages and limitations of FA in general and for psychological research in particular . In more practical terms, perhaps, FA can serve as a gauge of genomic (e.g., heritability, inbreeding) and environmental (e.g., nutrition, exposure to pollution, living density) stresses in human populations, and even bridge conceptually between these two clusters factors (and too often unbridgable approaches.)

The concept of FA offers additional applied outlook for future research since its assessment may become an important monitoring tool and even early-warning method for detecting individuals and populations under stress (for seminal research on non-human populations, see Sarre, Dearn & Georges 1994; Wagner 1996). Ideally, measurements of FA may enable researchers to isolate and reduce these stresses, and thus may limit developmental problems and improve developmental resistance.

REFERENCES

Babler, W. J (1978) Prenatal selection and dermatoglyphic patterns, *American Journal of Physical Anthropology 48*, pp 21-28

Barden, H. S (1980) Fluctuating dental asymmetry. A measure of developmental instability in Down syndrome, *American Journal of Physical Anthropology 52*, pp. 169-173

Buss, D. M. and Schmitt, D. P (1993) Sexual strategies theory. An evolutionary perspective on human mating, *Psychological Review 100*, pp 204-232

Campbell, Geller, Small, Petty & Perris (1978)

Clarke, G M (1996) Relationships between fluctuating asymmetry and fitness: How good is the evidence?, *Pacific Conservation Biology 2*, pp. 146-149.

DiBennardo, R. and Bailit, H L (1978) Stress and dental asymmetry in a population of Japanese children, *American Journal of Physical Anthropology 48*, pp. 89-94

Fujimoto, A., Allanson, J., Crowe, C A, Lipson, M. H. and Johnson, V. P. (1992) Natural history of mosaic trisomy 14 syndrome, *American Journal of Medical Genetics 44*, pp. 189-196.

Gangestad, S. W and Simpson, J. A (1990) Toward an evolutionary history of female sociosexual variation, *Journal of Personality 58*, pp 69-96.

Gangestad, S W and Thornhill, R. (1997b) Human sexual selection and developmental stability, in Simpson, J. A. and Kenrick, DS T., eds *Evolutionary Social Psychology*, pp. 169-195, Mahwah (NJ) Lawrence Erlbaum.

Garn, S M, Mayor, G H. and Shaw, H. A. (1976) Paradoxical bilateral asymmetry in bone size and bone mass in the hand, *American Journal of Physical Anthropology, 45*, 209-210.

Goldberg, C. J., Dowling, F E., Fogarty, E E. and Moore, D P (1996) Adolescent idiopathic scoliosis as developmental instability, *Genetica 96*, pp. 247-255

Green, M. F, Bracha, H. S., Satz, P and Christenson, C. D (1994) Preliminary evidence for an association between minor physical anomalies and second trimester neurodevelopment in schizophrenia, *Psychiatry Research 53*, pp. 119-127

Grossman, C J (1985) Interactions between the gonadal steroids and the immune system, *Science 227*, pp 257-261.

Hoyme, H E. (1993) Minor anomalies: Diagnostic clues to aberrant human morphogenesis, *Genetica 89*, pp 307-315

Kenrick, D T (1994) Evolutionary social psychology From sexual selection to social cognition, in Zanna, M P, ed *Advances in Experimental Social Psychology 26*, pp. 75-121, San Diego. Academic Press

Kobyliansky, E. and Livshits, G (1989) Age-dependent changes in morphometric and biochemical traits, *Annual Human Biology 16*, pp 237-247.

Kowner, R (1995) Laterality in facial expressions and its effect on attributions of emotion and personality A reconsideration, *Neuropsychologia 33*, pp. 539-559

Landa, N., Oleaga, J. M., Raton, J A, Gardeazabal, J. and Diaz-Perz, J L. (1993) Focal dermal hypoplasia (Goltz syndrome) An adult case with multisystematic involvement, *Journal of American Academy of Dermatology 28*, pp 86-89

Livshits, G. and Kobyliansky, E. (1989) Study of genetic variance in the fluctuating asymmetry of anthropometrical traits, *Annals of Human Biology 16*, pp. 121-129.

Livshits, G and Kobyliansky, E (1991) Fluctuating asymmetry as a possible measure of developmental homeostasis in humans, *Human Biology 63*, pp 441-466.

Livshits, G, Davidi, L., Kobyliansky, E., Ben-Amitai, D., Levi, Y. and Merlob, P. (1988) Decreased developmental stability as assessed by fluctuating asymmetry of morphological traits in preterm infants, *American Journal of Medical Genetic 29*, pp 793-805

Ludwig, W (1932) *Das Rechts-Links Problem im Tierreich und beim Menschen [The Problem of Right-Left in the Animal Kingdom and in Humans]*, Berlin· Springer.

Malina, R M. and Buschang, P. H (1984) Anthropometric asymmetry in normal and mentally retarded males, *Annals of Human Biology 11*, pp. 515-531

Malyon, C and Healy, S (1994) Fluctuating asymmetry in antlers of fallow dear, Dama dama, indicates dominance, *Animal Behavior 48*, pp 248-250

Manning, J T and Wood, D (1998) Fluctuating asymmetry and aggression in boys, *Human Nature 18*, pp 53-65

Manning, J T., Koukourakis, K. and Brodie, D. A (1997) Fluctuating asymmetry, metabolic rate and sexual selection in human males, *Evolution and Human Behavior 18*, pp. 15-21.

Markow, T. A (1992) Genetics and developmental stability An integrative conjecture on aetiology and neurobiology of schizophrenia, *Psychological Medicine 22*, pp 295-305

Markow, T A and Clarke, G M (1997) Meta-analysis of the heritability of developmental stability A giant step backward, *Journal of Evolutionary Biology 10*, pp. 31-37.

Markow, T A. and Gottesman, I I (1989) Fluctuating dermatoglyphic asymmetry in psychotic twins, *Psychiatry Research 29*, pp 37-43.

Markow, T. A and Wandler, K (1986) Fluctuating dermatoglyphic asymmetry and the genetics of liability to schizophrenia, *Psychiatry Research 19*, pp 323-328

McLachlan, A and Cant, M (1995) Small males are more symmetrical: mating success in the midge Chironomus plumosus L (Diptera Chironomidae), *Animal Behaviour 50*, pp 841-846.

Mellor, C S (1992) Dermatoglyphic evidence of fluctuating asymmetry in schizophrenia, *British Journal of Psychiatry 160*, pp. 467-472.

Mitton, J. B (1995) Enzyme heterozygosity and developmental stability, *Acta Theriologica 40 (Suppl. 3)*, pp. 33-54

Møller, A P (1996) Sexual selection, viability selection, and developmental stability in the domestic fly, *Musca domestica Evolution 50*, pp 746-752

Møller, A P and Pomiankowski, A (1993) Fluctuating asymmetry and sexual selection, *Genetica 89*, pp 276-279

Møller, A P and Thornhill, R. (1998) *Bilateral symmetry and sexual selection. A meta-analysis*, *American Naturalist 151*, pp 174-192

Murphy, K C. and Owen, M J (1996) Minor physical anomalies and their relationship to the aetiology of schizophrenia, *British Journal of Psychiatry 168*, pp 139-142.

Naugler, C. T. and Ludman, M D. (1996) A case-control study of fluctuating dermatoglyphic asymmetry as a risk marker for developmental delay, *American Journal of Medical Genetics 66*, pp 11-14

O'Callaghan, E., Larkin, C. and Kinsella, A (1991) Familial, obstetric and other clinical correlates of minor physical anomalies in schizophrenia, *American Journal of Psychiatry 148*, pp 479-483

Palmer, A R and Strobeck, C (1992) Fluctuating asymmetry as a measure of developmental stability Implications of non-normal distributions and power of statistical tests, *Acta Zoologica Fennica 191*, pp 57-72

Parsons, P A. (1990) Fluctuating asymmetry An epigenetic measure of stress, *Biological Review 65*, pp. 131-145.

Peretz, B , Ever-Hadani, P , Casamassimo, P , Eidelman, E., Shellhart, C and Hagerman, R. (1988) Crown size asymmetry in males with far (X) or Martin-Bell syndrome *American Journal of Medical Genetics, 30*, 185-190.

Polak, M (1994) parasites increase fluctuating asymmetry of male *Drosophila nigrospiracula*. Implications for sexual selection, in Markow, T., ed. *Developmental Instability Its Origins and Evolutionary Implications*, pp. 257-267, Dordrecht Kluwer

Sarre, S., Dearn, J. M and Georges, A (1994) The application of fluctuating asymmetry in the monitoring of animals populations, *Pacific Conservation Biology 1*, pp. 118-122

Shackelford, T K. and Larsen, R. J. (1997) Facial asymmetry as an indicator of psychological, emotional, and physiological distress, *Journal of Personality and Social Psychology 72*, pp. 456-466

Shapiro, T. (1965) Hand morphology in some severely impaired schizophrenic children, *American Journal of Psychiatry 122*, pp 432-435

Sprecher, S , Sullivan, Q and Hatfield, E (1994) Mate selection preferences. Gender differences examined in a national sample, *Journal of Personality and Social Psychology 66*, pp 1074-1080

Swaddle, J. P. and Cuthill, I. C (1994) Preference for symmetric males by female zebra finches, *Nature 367*, pp 165-166

Thornhill, R (1992) Females prefer the pheromone of males with low fluctuating asymmetry in the Japanese scorpionfly (Panorpa Japonica), *Animal Behaviour 43*

Thornhill, R and Møller, A P (1997) Developmental stability, disease and medicine, *Biological Reviews 72*, pp. 497-548.

Van Dusen, C R (1939) An anthropometric study off the upper extremities of children, *Human Biology 11*, pp. 277-284.

Van Valen, L (1962) A study of fluctuating asymmetry, *Evolution 16*, pp. 125-142.

Waddington, C H (1957) *The Strategy of Genes*, London. Allen and Unwin.

Waddington, J. L., O'Callaghan, E and Larkin, C (1990) Physical anomalies and neurodevelopmental abnormality in schizophrenia, *Schizophrenia Research 3*, p. 90.

Wagner, E. J. (1996) History and fluctuating asymmetry of Utah salmonid broodstocks, *Progressive Fish-Culturist 58*, pp 92-103

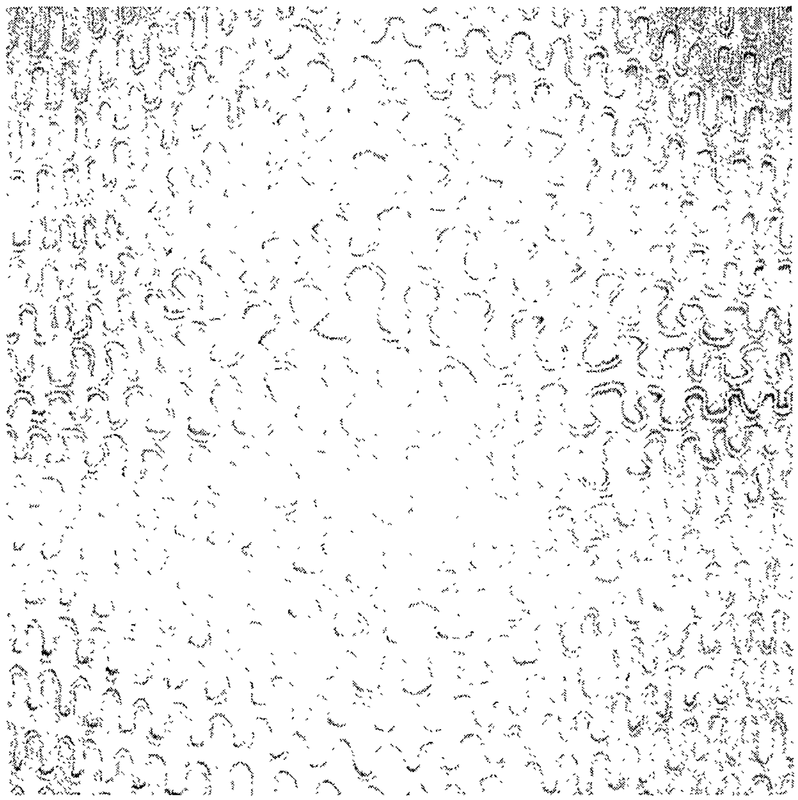
Waldrop, M. F. and Halverson, C. F (1971) Minor physical anomalies and hyperactive behavior in young children, in Hellmuth, I. J , ed *Exceptional Infant Studies in Abnormalities (Vol. 2)*, pp 343-380 , New York· Brunner, Mazel.

Wilson, J M and Manning, J T (1996) Fluctuating asymmetry and age in children· Evolutionary implications for the control of developmental stability, *Journal of Human Evolution 30*, pp. 529-537.

Witter, M. S. and Swaddle, J P (1994) *Fluctuating asymmetries, competition and dominance. Proceedings of the Royal Society of London* (B), pp. 256, 299-303.

Woolf, C M and Gianas, A. D. (1976) Congenital cleft lip and fluctuating dermatoglyphic asymmetry, *American Journal of Human Genetics 28*, pp 400-403

Yeo, R A., Gangestad, S W and Daniels, W. F. (1993) Hand preference and developmental instability, *Psychobiology 21*.



Władysław Strzemiński: *13. unistic composition*, 1934
50 × 50 cm; Museum of Art in Łódź

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 347-371, 1998

ON WŁADYSŁAW STRZEMIŃSKI'S 13. UNISTIC COMPOSITION. PHILOSOPHICAL, PHYSICAL AND MATHEMATICAL ASPECTS

Ryszard Krasnodębski

Mathematician, economist, philosopher, pensioner, (b. Łomża, Poland, 1925); assistant at the Dept. of Political Economy (1948–1949); lecturer in mathematics in the Wrocław University of Technology (1951–1968, 1981–1985); lectured on the symmetry theory (1961–1968) for students of the Dept. of Architecture; head of the Mathematical Models Laboratory in the Institute of Meteorology and Water Economy (1969–1981).
Address: ul M Kasprzaka 4, 51-676 Wrocław 9 skr p. 77, Poland
Fields of interest: philosophy (metaphysics, philosophy of science), physics, mathematics, economics, pictorial art.
Publications: Imbedding of a Space with an Affine Connection in the Affine Space, Ann. Pol. Math. (1964), *On Osculating Spheres to Rectifiable Curves*, Ann. Pol. Math. (1970); *On Probabilistic Weights of Biological Hypotheses*, ISEM Journal (1985); *Rozważania o demokracji (Considerations on Democracy)* (1994); translation into Polish (1967) H.S.M. Coxeter, *Introduction to Geometry.*

INTRODUCTION

More than 30 years ago when I saw for the first time the Władysław Strzemiński's 13. unistic composition, its intellectualism struck me. It would not be easy for me to say why, for many years, I had not lost my inward union with the painting even though my professional interests were then rather far removed from art problems. The possibility of delivering a lecture on Strzemiński's work, in a seminar organized by Janusz Rębielak, under the auspicies of the ISIS-Symmetry, renewed my attention towards it with great pleasure. Therefore, the painting is the subject of our present contemplation.

Symmetry, which is so mightily manifested in Strzemiński's composition, is a rather specialized notion; it is understood in general, not only in a confined sense. Nevertheless, it is a commonly accepted view that the idea of symmetry permeates deeply and comprehensively the reality which surrounds us. Interdisciplinary studies meet the notion half-way.

If we adduce the two notions and concentrate our attention on Strzemiński's composition – that is, if we confine our interest to the art of painting which is named commonly as an abstract one – we detect a number of gateways some which lead us to other areas of creation and at the same time to language regions other than those which are used when we speak about artistic work. If we go through such a gateway and enter other scenery, it is not easy to stop on the path which inevitably directs our thinking to unexpected points of dispute.

The painting under consideration reveals two primary aspects, one deterministic, the other probabilistic. The first aspect should be named structural, while the second aspect is inseparably connected with the production of the work. The design of the work, in any art activity carried out in the mind, is, one may suppose, almost always structural, whilst the production of it embraces, however, unavoidably, random deviations from the contemplated structure. I disclose in what follows the mathematical, or, more precisely, the algebraic structure of Strzemiński's composition. The execution of the work is however a more comprehensive problem, involving as it does the probabilistic aspect of the painting. I will merely touch on the point to be solved by the work as it presents itself to our eyes.

Now, on the theoretical background I will elaborate my probability doctrine, which I submit for serious consideration. Because I had the painting in front of me in all its concrete material attributes, it would be inappropriate to refer to probability abstractly defined. So the definition I give in Chapter III will contain only such notions as have their referents directly in the world perceptible by our senses. This definition provided in one of the gateways which led me to fundamental physical problems in particular Born's probabilistic interpretation of the solution to Schrödinger's equation. The consequences of methodological failure in this interpretation are of course philosophical ones. I point out the other failure in modern physics which has philosophical consequences, namely the departure from classical differentiation of two language regions: in one of them the terms have their referents in the physical world, in the other, in the world of notions created in mathematics.

The other gateway which directed me to metaphysics is the problem of time. The point in dispute raises some inconsequences, similarly to the previous case, in the fundamentals of physics.

I would like to insist – in spite of the prevalent opinion – that metaphysics deals with *the reality*, i.e. something which exists, *exists* independently of whether or not can we grasp it, either with our senses, or intellect. This attitude of mind could not be unrecognized as conformable with the common sense. We speak in this philosophy about *being as such*, i.e., about beings which can be grasped as a whole. And we assume that the reality is a whole of all beings, independently of whether we can grasp differences between them or not.

The production process of a work of art and its finished form belong to two worlds: to the one which surrounds us with concrete materialized matter and to the world which transcends matter. The aim of my paper is to throw some light both on the process, and on the finished work, in this case by examining is Strzemiński's 13. unistic composition. In Chapter I I renew one of the primary existential questions touching the world in which we are at the same time immersed and make ourselves its particular particles.

I. NINE PRIMARY QUESTIONS

The many-weftedness which I have outlined in the Introduction, does not relieve me from discussing some crucial notions with suitable accuracy. It occurs consequently as indispensable to allude to the fundamental problems investigated in contemporary physic, not only by these gateways which are commonly accepted, but by those not noticed, needful nonetheless I think, in order to better apprehend something that goes on in the microworld.

I formulate our questions in a general way.

Is the world
- 1° deterministic?
- 2° indeterministic? (probabilistic?)
- 3° a juncture of both these potentialities?

A positive answer to the first question would accept the view that all events are deterministic. It is unreasonable, I think, to be of the opinion that a positive answer to the second one would mean all events are random. It would rather imply that there exist random events. My deliberations, which I offer in this lecture, show, however, that the notion *random events* is, in a physical sense, vague. Not without good reasons, therefore, I have involved in the third question both notions mentioned in the two previous ones.

Let me repeat, however, that such a juncture is rather to point out the difficulties entangled in them than to penetrate their precise meaning.

I draw up in what follows the definition of the deterministic event; it will be useful, however, to precede the definition by some remarks. If we talk about *an instant* we appeal to the physically unrecognizable notion. Similarly, we can not avoid other physically unrecognizable notions too: *a point*, *a curve* (in special cases *a straight line*), a surface (in special cases *a plane*). Our considerations do not lose anything as to their generality if we use the term *plane-area*, being, if necessary, an approximation of *surface-area*.

When talking about an event, we have in mind however, such a distinguished and bounded space-area which "moves" in a well defined direction and inside of which something goes on. To be consistent we will use the term *process-event* occurring in a *time-interval*. Consequently, we have found ourselves in an unavoidable situation: when talking about physical process-events we are not able to omit four abstract (or, if you prefer, mathematical) notions mentioned above.

Their referents – an instant, a point, a curve, a surface – however, in the physical world are meant as: a small time-interval, a small spot, a pipe having a small cross-section, a flat and suitably thin part of space. When talking about *space* we have in mind at the same time three-dimensional space mathematically meant, and the space in which we are immersed; sometimes it will be more convenient to say, in the last case, *physical space*.

Let us consider:

1° a process–event, $\gamma(t, t')$, going on in a time-interval $[t, t']$, $t < t'$; by $\gamma(t)$ and $\gamma(t')$ let us denote the "ends" of it; they could be named plane-events; let us assume they are going on in a bounded space-area;

2° a process-area, $\Pi_\gamma(t, t')$ going on in the physical space: $\Pi_\gamma(t, t')$ "is swept" by the plane-event $\gamma(t, t')$;

3° the plane-area, $\Pi_\gamma(\tau)$, which is defined by $\Pi_\gamma(t, t')$ for any instant τ, $t \leq \tau \leq t'$; $\gamma(\tau)$ is the end of the process-event $\gamma(t, \tau)$.

Definition

A process-event, $\gamma(t, t')$, is said to be determined iff:

1° there exists in the physical space a limited plane-area $\Pi(\tau)$ which in a time-interval $[t, t']$ "sweeps" a process-area $\Pi(t, t')$ such that: (a) at every instant τ, $t \leq \tau < t'$, $\Pi_\gamma(\tau) \subset \Pi(\tau)$; (b) $\Pi(\tau) \div \Pi_\gamma(\tau)$ constitutesa surface-area which is homeomorphic with a ring;

2° there is not going on any other process-event in the process-area $\Pi(t, t')$;

3° when there occurs at the left end of the process-area $\Pi(t, t')$ a plane-event $\gamma(t)$, the process-event $\gamma(t, t')$ is the one and only.

The definition seems to be reasonable. If we agree that the notions *undetermined process-event* and *random process-event* are synonyms and then we would like to formulate the definition in such a way that the new condition instead of 3° would deny the existence of *the one and only* process-event, we would be, I think, in an unsolvable situation.

And what about *all* events? (The notion *all events* is equivalent to the notion *world* used in our triple fundamental question). The question puts us in limitless difficulties and at the same time in limitless possibilities. I therefore omit it and I go to further considerations.

As the next step in our deliberations let us begin with the following doubt: should we assert that intellect (I have in mind the intellect of the human being) and its function are components of the physical world or at least the function is something *beyond* it? I leave out epistemic questions yielded by this dichotomous query. A good deal of effort has been devoted by many thinkers to this problem.

I assume that intellect performs its functions when engaging its physical part and its other part being beyond the physical world. Consequently, I should reform the questions which have been formulated above.

The full version of our inquiring questions is therefore the following:

$$\text{Is the reality}\left\{\begin{array}{ll} A = & \text{physical reality,} \\ & \text{i.e. the world} \\ & \text{which we observe} \\ & \text{and investigate} \\ & \text{by physical methods} \\ & \\ B = & \text{the rest} \end{array}\right\} \begin{array}{ll} 1^\circ & \text{deterministic?} \\ & \\ 2^\circ & \text{indeterministic?} \\ & \text{(probabilistic?)} \\ 3^\circ & \text{a juncture of both} \\ & \text{these potentialities?} \end{array}$$

II. TWO REALITIES

Let us denote the whole of two realities, A and B, by R. The reality R contains all that exists. It is groundless to think that beyond the reality A there exists nothing. So, let us assume the reality B is not empty. As to the reality A let us assert that it is uncreative in the following sense. New beings arise in the reality A by transformation (one should say: it takes a time-interval) of other beings or by compounding (one should say too: it takes a time-interval) of same other ones being components of A previously.

The distinction of these two realities, A and B, is essential from the physical point of view. It seems that it was Einstein who first used, in the Special Theory of Relativity, the notion *observer*. In spite of the suggestive term, however, observer belongs to the reality B not to A. Let us say more exactly: the term *observer* is a synonym for the *observer's senses*. The observer is, therefore, identified with instruments recording signals which are dispatched or received. There is, however, involved the distinction in the Theory of Relativity, though unspoken, which we have made in our scheme of questions. It is easily to see that both Theories of Relativity, Special and General, could not be presented in any other way as only by a *narrator*; "he" looks on what is taking place in the reality A, i.e., what had been observed by instruments. The narrator's point of "theoretized observation" belongs to the reality B. Let us imagine, for example, that an observer directs his eyesight – using a telescope – to a star; the sun's position is at the same time such that the starbeam bends because of the influence of the sun's gravity field. The observer locates the star in a position which is imaginary. The notion *imaginary position* could be understood, however, only by the narrator and not by the observer.

The above necessity, resulting from the point of view of physics sciences, *forces us* – that is, I think, the right word in this place – to direct our attention to metaphysics. The remark leads to the following generalization. I do not see any reasons to uphold that the reality B contains the narrator's intellect only. The gateway opens our thinking, of course, to the metaphysics.

I consider in this paper a materialized object: Strzemiński's 13. unistic composition. It is a part of the reality A. One could not agree that the painting is not a fruit of the author's spiritual powers. Therefore, the scientific attitude of mind – what I have stressed above – and, as well, the attitude of an artist's genius when creating his work implement our conviction that the transcendental reality does exist. The ninefold question formulated at the end of Chapter I is, therefore, justified.

III. EXPECTATIVE PROBABILITY

The judge casts a coin upwords on a football field. All football players and all spectators know that the coin will fall down some seconds later on the field. They know too that the probabilities of the two expected self-evident possibilities are equal. It is for them quite needless to attach to any of the two expected occurrences the number 1/2. It is meaningless too if anybody ever anywhere, in a similar or unsimilar situation, had been casting a coin upwards and estimated chances of the two interested parties. Consequently, there is *no set of events* connected with the notion *expected event's probability*. In other words: to speak of a space of events is needless. Actually, at this very moment, we are awaiting for one and only one event which will occur at the end of a time–interval. As a principle, as in the above example, we are interested in some two versions, *a priori* marked out, of the expected event.

This is another example. Let us consider an arrangement similar to a continuous roulette wheel. The rod turns round freely on its axis. The rod's end moves over the ring having the same axis as the center. There are marked on the ring n unequal sections. One of them, e, is indicated by a distinguishing mark; the section e covers the p-th ($0 < p < 1$) part of the ring. Let us give the rod an impetus, sufficiently strong, so that it turns round tens or more times. At the beginning of the experiment the expected event, E, is the stop of the rod. The event has in respect to our interest two versions: stopping over the section e, and the other, e', stopping beyond the section e. When the rod still turns round but sufficiently slowly we are able to attach to the expected version e the probability nearer and nearer to zero or to one. Before the experiment is being undertaken, or when the rod moves sufficiently quickly, we attach to the version e the probability equal to p and to the expected event E the probability equal to 1. I repeat the comment made on the occasion of the previous example: it is of no significance whether anybody, ever or anywhere, had performed such an experiment, or if the impetus to the rod more had been given than once.

Definition

Let us assume that we expect at an instant t, $t < 0$, the event E having to be realized at $t = 0$. To the expected event E let us attach the integer 1 called the expected probability of it. Let two versions, e and e', be distinguished in E, $e \cup e' = E$: if it realizes the version e, then it does not realize the version e' and vice versa; let us attach the number p,
$0 < p < 1$, to the version e; consequently, the number $1 - p$ to the version e'; they are called respective expected probabilities of the version e or of the version e'. (Compare [2]).

I make an obvious remark, but not without implications. The use of symbols in the definition does not break our essential goal: to formulate definition which terms have their referents in the physical world. Let me say this another way: semantic meaning of the terms used in the definition refer *a priori* to the world A and not to the area of mathematical language. If one talks, however, about *a probability of an event* and not about *a probability of an expected event* the referent of the term *event* belongs to the area of mathematical language and it denotes an element of an abstract set and nothing more.

The notion *respective expected probability* used at the end of the definition connects it with its applications. There are two sources of knowledge about expected events probabilities. We have in mind the knowledge to be possessed of by the intellect evaluating the probability at an instant t, which anticipates the expected event. The two sources are the following: (a) the understanding of the process' mechanism yielding to the expected event; (b) information in general necessary on the same – or rather sufficiently similar – events which had taken place in the past. The first one let us name *technological*, the second one is usually named *statistical*. To evaluate the probabilities mentioned in our two examples statistical methods are needless. In the case when it is indispensable an adequate statistical model is built; suitable parameters of it and parameters of suitable tests of them are not probabilities.

Let us say in the philosophical language: expected probability is a property of the relation rooted in the intellect and in the future (expected) event. The definition formulated above attaches to the relation – let us name it *uncertainty relation* – a measure. It seems that the expected probability is the only measure of this relation.

Is it possible to define probability – of a future event, obviously – if it is not expected by any intellect? In other words: is there a reasonable thing to say about an objective probability? Let us consider this questions.

Let $K(t)$, $t \in [\tau, 0)$, be a set of all phenomena which anticipate an event E including two possibilities, e and e'. Both $K(t)$ and E belong to the same bounded region. Let us consider in a *mental experiment* two situations. In one of them the version e of E had been realized. In the other one all phenomena $K(t)$ run exactly the same way as they were previously run but the version e' is realized. Realization of two such situations would affirm the existence of a random event in a physical sense in the physical world A. Our understanding of physical laws causes, however, that the above mental experiment grows to be internally contradictory. In every time-interval there *come to light* exactly the same physical laws relevant to the phenomenon, separated from all other ones, which is the same as in any other time-interval. Using the words *come to light* I would like to emphasize that physical laws have another ontic status than the observed phenomenon or the phenomenon which is not observed by any intellect. Therefore, random events do not exist in the reality A.

Another reasoning leads to the same conclusion. Let us assume that in a time-interval $[t_1, t_2]$ ($[t_1, t_2] \subset [t, 0]$) the run of the phenomena $K(t)$, considered in the above mental experiment, had been disturbed by an interference coming from the reality B. If the version e' would be realized at the instant $t = 0$, we would not have any reason to maintain the hypothesis announcing that the same running of the phenomena could result in two different versions. The disturbance caused, of course, that there occurred an another set of phenomena anticipating the event E, $K'(t)$ instead of $K(t)$. The reasoning belongs to the borderland of science and metaphysics languages.

Against the background of the above deliberations let us make two historical remarks. The well-known saying of Einstein: "God does not play dice with the universe" more exactly means: God does not play these dice which belong to the reality A. And at the same time He does not play till the end, i.e., till the end of time interval $[t, 0]$. The second remark is the following: if one assumes that in the reality A there do exist random events (whatever this means), then denying Einstein's saying one accepts at the same time the existence of the reality *non-A*.

Consequently, let us state positively: anything you like which is happening in the reality A has a determined character until there occurs the inference of a force which come from the reality B.

IV. ON THE INDETERMINACY PRINCIPLE. PRELIMINARY REMARKS

Most physicists are convinced that the view that the world has a probabilistic character is well-founded on the Heisenberg Indeterminacy Principle. Conclusions, however, supposedly arising from the Principle's formalization have been involved in such reasoning which is not deprived, at least, of inconsequences.

I develop an analogy; it will be quite perspicuous and shows, I think, an error – other ones will be mentioned in the further chapters – in the argumentation on Indeterminacy Principle. Let us consider a moving point, p, on a plane. We observe it, but we cannot come to an agreement as to its position at an instant, t, until we choose three points which do not move, and the positions of which are well-defined. In other words, until we fix a coordinate system, let us say, the rectangular Cartesian one, XOY. Let us place a theodolite in a suitable position with its object-glass directed perpendicularly to the OX-axis. We take, of course, the reading of the abscissa of p. When taking this reading we cannot say anything about its ordinate and anything about its velocity. A point is an integral notion; it is something by itself. Fixing of a coordinate system is something arbitrary. Giving the point two coordinates is something arbitrary too. The surveying of them has two limitations: the measurement instrument (it is such as it is, nothing else) and the theory (it is such as it is, not any other). It does not result therefore that, if we are not able to define both coordinates at the same time, and define one of them only, the other one does not exist. The transparency of the experiment and the simplicity of the theory assure us that the theory and practice are complete. It would be worth to add: not only because of the practice, e.g., location of rockets, we do not meet failure.

The simplicity of the experiment has other aspect too. Notions referring to the objects under consideration belong to such two language regions, separation of which would be unnecessary. The notions of one of them have their referents in the reality A, the notion of the other one in the mathematical terms-region. The mutual attaching of these referents – even when they have been expressed by the same words – is well-known.

Similar correspondence – in the case of the physical phenomenon called *electron* with its theory – is more complicated. Let us say without hesitation that because of unimaginability it is rather mysterious. The above developed analogy convinces us, however, I think, that Heisenberg's interpretation of the theory referring to a moving electron is not persuasive. In the further considerations I show that it is methodologically faulty. I concentrate the attention on some notions treated metaphysically: *existence*, *being*, *relation*, *movement* and *time*.

V. EXISTENCE, BEING AND SOME REMARKS ON METHAPHYSICS

Two notions, *existence* and *being*, belong to the fundamentals of every thinking. It is not the question if the being, which is the subject of a talk, or meditation, does exist, but how it exists. When we use a language to be considered as scientific it is rather needless to refer to these notions in an overt way. When a discussion touches, however, phenomena in which physics of the micro-world is interested, the notion *existence* is allowed to be spoken of.

In methaphysics understood as the Aristotelean *philosophia perennis* a being is called *something* or also a *thing*. If we say *this thing does exist* – and we use a known name for it – we have in mind something with which we are previously not unacquainted at least in some (maybe one) aspects. How does the thing we have named exist is not an ontic problem but an epistemic one. The statement that *the thing does not exist* is, therefore, senseless; it means nothing.

Every branch of science searches into some occurrences of a phenomenon distinguished from other ones and applies an appropriate methodology; reflection on observation is a part of the methodology, the other one is speculative reasoning. The results of scientific searches are *ex definitione* local. To make the last notion plain I characterize it by referring to mathematics. In differential geometry there are considered local properties of topological spaces provided locally with a metric form. In algebraic topology there are searched global properties of them. Methods used in one of them are quite different than in another one. Indirect inference from local properties of a metric topological space to its global ones would be a failure and even impossible. This analogy exemplifies in some respect the difference between science and methaphysics and at the same time validates metaphysics; the last one is such a field of human thought which comprehends the reality and its fragments as a whole. One of the methods of metaphysics is speculative reasoning aiming directly – this means: without intervention of measurement instruments and without building any mathematical model – to grasp the reality *A* and, consequently, the reality beyond *A*. The speculative thinking rests upon rigorous, coherent, i.e. intrinsically close, reasoning, upon a characteristic way of generalization yielding insight into the reality *B* and upon intuition.

VI. THE COPENHAGEN INTERPRETATION OF THE QUANTUM MECHANICS. LANQUAGE AND NOTION PROBLEMS

Friedrich von Weizsäcker [4, p. 10] comments briefly the essence of the Copenhagen interpretation of Heisenberg's discovery. He writes: "(...) die Messung zerstört den stabilen Zustand [des Elektrons; R.K.] und schafft einen neuen. Das sie [Ort und Impuls des Elektrons; R.K.] nicht zugleich existieren können, spiegelt sich darin, dass sie nicht zugleich messbar sind. Aus diesem Bruch mit der klassischen Ontologie folgt der Bruch mit dem klassischen Determinismus. Die Messergebnisse sind *prinzipiell* nur mit Wahrscheinlichkeit vorhersagbar".

The last sentence – though the word *prinzipiell* is rather vague in it – is correct; it does not result from the idea that the classical determinism should be rejected. It is not so. Weizsäcker says more categorically: "Die »Kopenhagener Deutung« der Quantentheorie, Bohrs und Heisenbergs gemeinsames Werk, gilt heute im Unterricht als orthodox und wird, wie alle grosse Orthodoxien, von der Mehrzahl ihrer Bekenner kaum verstanden [...]."

There are physicists who are not believers of the orthodoxy. R. Eisenberg and R. Resnick [1, p. 88] write:

"Among the critics of the Bohr–Heisenberg view of a fundamental indeterminacy in physics is Louis de Broglie. In a forward to a book by David Bohm [...] de Broglie writes: »[...] The construction of a purely probabilistic formulae that all theoreticians use today was completely justified. However, the majority of them, often under the influence of preconceived ideas derived from positivist doctrine, have thought that they could go further and assert that the uncertain and incomplete character of the knowledge that experiment at its present stage gives us about what really happens in microphysics is the result of a real indeterminacy of the physical states and of their evolution. Such an extrapolation does not appear in any way to be justified [...]«. (From *Causality and Chance in Modern Physics* by David Bohm, © 1957 D. Bohm; [...] D. Van Nostrand Co.)."

The inference, emphasized by von Weizsäcker, let us express in its full reading: from the fact that the position and the momentum are not measurable at the same time one should conclude that both aspects of the electron do not exist at the same time.

Apart from the opposite opinions mentioned above I would like to present a reasoning more convincing, I think, than those. It will occur that fundamentals of physics are deeply penetrated by metaphysical notions.

We should speak about the electron whose property is *position-momentum*, like the property of a coin is two-sideness: *obvers-revers*. The method of measurement is something arbitrary and is restricted by possibilities of our devices and, therefore, the result includes either *momentum* or *position*; similarly, we see either the *obvers* or the *revers* of a coin. A general remark is the following: any measurement made by man is ontologically somewhat different from what happens in the reality A.

The view that the structure of the reality A is indeterministic is believed to have its warrant in the conformability of experiment results with the probabilistic interpretation of the Schrödinger equations solution. Two methodological misunderstandings are implied in this feeling of certainty. One of them is known. The conformability mentioned above justifies the following inference only: experiment results do not negate the probabilistic interpretation of Schrödinger's theory. No more. The second misapprehension has its source in the language groundwork. I mentioned in the Introduction that, thanks to the genius of classical physics founders, the process in which man's intellect assigns the notions, used in the physics theories, to their referents in the reality A, grows to the unnoticeably routine. In modern physics such a clear distintion has been lost.

Let us consider Newton's equation

(*) $$F = m\frac{d^2x}{dt^2}$$

and the Schrödinger one

(**) $$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} + V(x,t)\Psi(x,t) = i\hbar\frac{\partial\Psi(x,t)}{\partial t}.$$

Let us divide symbols which occur in these equations into five groups:

1° Constants: m and h. Their physical meaning is exactly determined – in other words: it is the determined assignment of these symbols to the suitable »phenomena« occurring in the reality A – and at the same time they represent numbers.

2° Constant: $i=\sqrt{-1}$. The occurring of i in Schrödinger's equation and its lack in Newton's equation is one of the reasons which make the difference between them not only a formal one. The equation (*) is »comprehensible« for the observer and all the more for the narrator: all factors of it could be measured. The equation (**) is on the other hand, for the sake of i, »incomprehensible« for the observer but comprehensible for the narrator. The constant i is a link, in this case, between the realities A and B. No wonder, therefore, that any solution of Schrödinger's equation could not be applied to the reality A indirectly. The repeated interference of the narrator is indispensable; for the first time it took place during the construction of the equation.

The plausibility arguments, presented in a particularly clear way, leading to Schrödinger's equation are to be found in R. Eisberg's and R. Resnick's *Quantum Physics* [1]. The narrator's arbitrary, though necessary, interference went together with the calling into being of the equation. In other words: it went together with the translation of language directed to the reality into the mathematical language. This procedure brought out the arbitrary inference in the suitable translation of the mathematical solution into the language »comprehensible« by the observer. This is the contraposition which throws a light on the calling into being not only Born's interpretation, but the philosophical and methodological judgement about the interpretation too.

3° Parameters: t and x. Their referents in the world A are: a small time-interval and a small spot, respectively. The size of them depends on the question which is under investigation and on the implements being at our disposal. It is quite an obvious matter how they are understood in mathematical language.

4° Functions: $F(x, t)$, $\frac{d^2x}{dt^2}$ and $V(x, t)$. Mathematical comprehension of them and referring them to the world A are classical problems.

5° Functions: $\Psi(x, t)$, $\frac{\partial\Psi(x,t)}{\partial t}$ and $\frac{\partial^2\Psi(x,t)}{\partial x^2}$. From the point of view of »the comprehensive faculty« of all factors occurring in Schrödinger's equation those mentioned in this item have the same ontic status as constant i; I do not repeat, therefore, the commentary made in that case.

Here is Born's interpretation of the wave function $\Psi(x, t)$ being the solution of Schrödinger's equation [1, p.147]:

If, at the instant t, *measurement is made to locate the particle associated with the wave function* Ψ(x, t), *then the probability* P(x, t)dx *that the particle will be found at a coordinate between* x *and* x + dx *is equal to* $\Psi^*(x, t)\Psi(x, t)dx$.

The sentence, or to use a specific term *the postulate*, we could treat as such which reveals a physical law independent of any experiment: formulation of the postulate leaves out the momentum's measurement of the particle; accordingly, the observer's and the narrator's points of view come together. Let us concentrate our attention on some aspects of the sentence.

If we agree that the postulate refers to the physical reality, all notions involved in it should be referred to this reality too. Probability should be understood physically as well. After the definition of the expectative probability (Chapter III), i.e. of the one which refers to the physical world, there should be taken into account a time-interval, let us say $[\tau, t]$. The assumptions which precede the construction of Schrödinger's equation do not justify talking about any well-defined time-interval which would be involved in the going on of the phenomenon called electron or other particle. In the light of this aspect the future tense, used in the postulate, from the point of view of the particle's behaviour, i.e., from the point of view of an observer, is »incomprehensible«.

Let us assume that the particle leaves its initial spot (x_0, y_0, z_0) at the instant τ. Let E denote the expected event that it will reach the axis X at the instant t; let e be the version of E which denotes that the particle will be found inside an interval $[x, x + dx]$. Then and only then the following statement could be comprehensible: the expected "probability $P(x, t)dx$ that the particle will be found at a coordinate between x and $x + dx$ is equal to $\Psi^*(x, t)\Psi(x, t)dx$". It should be, however, asked: what does the trajectory of the particle depend on? Let us consider two aspects of this question. The impacts, which the particle is submitted to on its way, are responsible for the random run of it. Let us leave aside this aspect and concentrate our attention on the other one. There are no reasons to think that the particle is spherically shaped. The possible trajectory of it, therefore, depends on what position, in respect to the X-axis, it assumes at (x_0, y_0, z_0). A pencil, of course, of possible trajectories of the particle is determined, therefore, at (x_0, y_0, z_0) at the instant τ. Thus, there are determined parameters of the probability distribution assigned to the particle having a well-defined situation at the instant τ.

In the experiments which preceded Born's postulate and then in the ones which supposedly confirm it there had been observed many particles. We say, from the point of view of the expected probability *many events*. Let me make an obvious remark. There are observed in these experiments event frequencies in various intervals of X. The experiments and then the histograms justify the following interpretation of them. If we could mark at random (if it were possible) a particle which just begins its trajectory directed to X we could say: the probability that it will be found at the instant t in the interval $[x, x + dx]$ is equal to $\Psi^*(x, t)\Psi(x, t)dx$. The inference is, of course, worthless.

Experiments in which there would be used one and only one particle were impossible till now. Inference, however, which refers to the behaviour of a single particle, bases on, let us say, *a global experiment*. It is the methodological fault.

There should be brought out into relief other methodological fault implied by Born's statement. Time's duration, or time-interval, of an experiment – this interval determines frequencies – has other ontic status than the one in which a particle runs its trajectory. The first one is settled by the narrator and the other one is independent of him.

It is rather a common belief that the physical world is *a probabilistic* one. This philosophical inference resulting supposedly from Born's postulate has precarious premises. On the other hand, our speculative reasoning reveals that solids and particles run their trajectories – if they are not subjected to the impelling force having its source in the world B – in a determined way.

VII. POET'S INTUITION

As a particle and time are conceived by someone, who does not submit to the orthodoxy, so a grain of sand is conceived by Wisława Szymborska. The poetic language of her poem *View with a Grain of Sand* describes the world which surrounds us in a way which is not so precise as is made by physics but, I am ready to admit, more fundamentally from the point of view of our world's perception. I think, that not without reason, I quote this poem below [3, p. 247].

We call it a grain of sand,
but it calls itself neither grain nor sand.
It does just fine without a name,
whether general, particular,
permanent, passing,
incorrect, or apt.

Our glance, our touch mean nothing to it.
It doesn't feel itself seen and touched.
And that it fell on the windowsill
is only our experience, not its.
For it, is no different from falling on anything else
with no assurance that it has finished falling
or that it is falling still.

The window has a wonderful view of a lake,
but the view doesn't view itself.
It exists in this world
colorless, shapeless,
soundless, odorless, and painless.

The lake's floor exists floorlessly,
and its shore exists shorelessly.
Its water feels itself neither wet nor dry
and its waves to themselves are neither singular nor plural.
They splash deaf to their own noise
on pebbles neither large nor small.

And all this beneath a sky by nature skyless
in which the sun sets without setting at all
and hides without hiding behind an unminding cloud.
The wind ruffles it, its only reason being
that it blows.

A second passes.
A second second.
A third.
But they're three seconds only for us.

Time has passed like a courier with urgent news.
But that's just our simile.
The character is invented, his haste is make-believe,
his news inhuman.

VIII. TIME – PHILOSOPHICALLY COMPREHENDED

Time is "imagined" – when the word is used in physics – as a number axis and as such it does not belong to any reality: time is in physics an element of a suitable mathematical model. Man's intellect assigns – at the same time – to the notion *time* used in such a model, a position or position change of physical objects. (In Born's postulate such an assignment fails; it is a consequence of the methodological fault I have mentioned in Chapter VI). Thanks to the simplicity of this assignment the question *what is time* escapes the physicist's attention. I present in what follows a metaphysical comprehension of time. I am convinced that the conception I develop, is not without weight for physics fundamentals and for our understanding of the world.

Let us take the *motion's* notion as a starting point. Motion exists at least in the reality *A*; motion is, therefore, a being. We will distinguish motion as a being and motion's measure. It is rational to distinguish motion assigned to two or more beings. In the first case we say that the motion is a property of the being under consideration, in the second one that motion is a relation.

Now, let me sketch briefly the theory of relation. In order to be consequent in this display of my metaphysical outlook on what we name *reality*, I assume that every two or more beings are ontologically related: they are connected by a relation which is *a substantial being*. I say more expressively: the relation takes root, or is rooted, in the beings which the relation constitutes. It should be accentuated that every relation – i.e. beings involved in it and the relation as such – is subjected to a process or, we can say, to a metamorphosis. The most expressive example of a relation is marriage. Marriage is not two individuals only, but an institution as well, having its characteristic properties; they appear as something new, something which did not exist earlier. Maybe the most unexpressive relation is the distance between two beings. Let us emphasize that motion in which we are interested in should be treated, in our theory, as a relation taking root in two or more beings. Such an understanding of motion is compatible with the understanding of relative motion in the Special Theory of Relativity. Distance is a statical notion, while motion is a dynamic one. Every relation which connects organic beings, e.g. of some human beings, has some, maybe many, components. The relation of two beings which belong to the reality *A*, has at least one component: the distance. Every component of a relation is to be treated as a substantial being. I think the following hypothesis should be assumed: the only one common component, connecting no matter what beings under consideration, is motion. Motion exists independently of human intellect and, let us say, it does not recognize itself. That is just what Szymborska's poem proclaims. Reality being left to itself is what it is and has neither past nor future. The two notions, *past* and *future*, are, of course, differentiable from the point of view of the intellect. We are interested in this paper in human intellect only. We have revealed, therefore, the essential source of *time's* notion: time meant in the human sense, in which we are interested, does exist as much as there exists the human intellect. We will enter into the idea of time if we define it as *the relation which takes root in the human intellect and in the motion*; just that is its ontic status. I say again: the measure of time is an attribute of it and nothing more. Here is the reason why we are inconsistent when we talk about time: the instant has no referent which would be something real in the reality A.

IX. STRUCTURAL SYMMETRY OF STRZEMIŃSKI'S 13. UNISTIC COMPOSITION

Let us fix our attention on Strzemiński's 13. unistic composition as it offers a view to our eyes of. Two aspects of it are in a remarkable way brought into prominence. I express the first one algebraically when creating an idealized schema and defining a transformation group of it. The second aspect, an analytical one, has undeniably a probabilistic character. One should call the first a deterministic one, not only to stay in the conventionality of symmetry, but in order to accentuate the creative inspiration of the Symmetry Principle.

We take into account the shaping of curves and pay no heed to the groundwork's colour. Let us devide the drawing into four parts (Fig. 1). The axes, X and Y, are perpendicular, parallel to the painting's borders and intersect in its symmetry centre. Let us draw the rays, having their initial points on Y, in such a way that they would approximate the wave lines. Let us denote them with numbers: $n = 1, 2, \ldots$ Let us acknowledge that the lower part of the drawing is a mirror reflexion of the higher one. The points marked on the Y-axis have – in suitable choosen units – the following coordinates:

$$y_n = 17\sqrt{n} + 13n + \frac{2}{3}[9 - (n-3)^2], \quad n = 1, 2, \ldots$$

whereas the points on the borders have the coordinates:

$$y_n = 7\sqrt{n} + 13n + \frac{2}{5}[34 - (n-4)^2], \quad n = 1, 2, \ldots$$

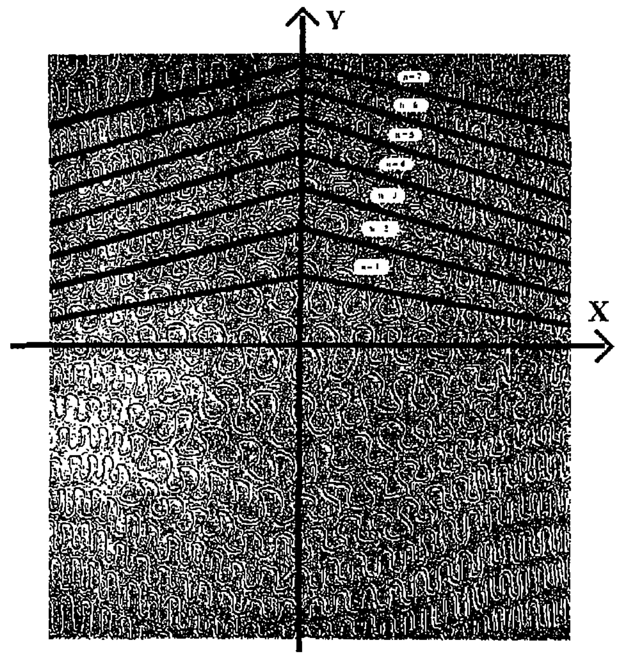


Figure 1

The first two components form the first approach of the approximation, the third one forms the first correction. The detection of further components would make our approximation more subtle. If someone does attach an importance to the number's magic, here, I think, he finds its confirmation.

I define now a transformation group and next we ponder on what part the group played in the creative act. Let there be given on the Euclidean plane three straight lines: horizontal axis X; axis OY vertical to OX intersecting the previous at the point O and directed upward; axis O^*Y^* parallel to OY intersecting the previous at the point O^* and directed upward (Fig. 2).

There is given on the positive half-axis OY: a well-ordered sequence of different points upward: $C_0 = O$, C_1, C_2, ...; well-ordered sequences of different points downward, $C_0^* = O$, C^*_{-1}, C^*_{-2},... The distances between neighbouring points are arbitrary. There is given on the positive half-axis O^*Y^*: a well-ordered sequence of different points upward, $C_0^* = O^*$, C_1^*, C_2^*, ...; a well-ordered sequence of different points downward, $C_0^* = O^*$, C_{-1}^*, C_{-2}^*, ... The distances fulfil the following conditions: $C^*_0C^*_1 < C_0C_1$, $C^*_1C^*_2 < C_1C_2$, ..., $C^*_0C^*_{-1} < C_0C_{-1}$, $C_{-}{}^{*}{}_{1}C_{-}{}^{*}{}_{2} < C_{-1}C_{-2}$, ...

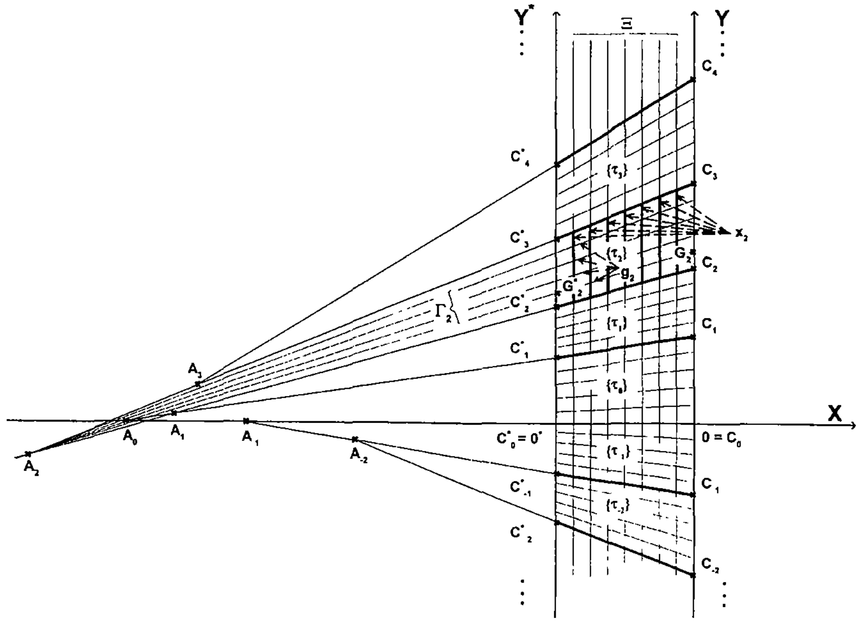


Figure 2

Let us denote in succession by A_0, A_1, A_2, ..., A_{-1}, A_{-2}, ..., the points: $X \cap C_1C^*_1 = A_0$, $C_1C^*_1 \cap C_2C^*_2 = A_1$, ..., $X \cap C_{-1}\, C^*_{-1} = A_{-1}$, $C_{-1}C^*_{-1} \cap C_{-2}C^*_{-2} = A_{-2}$, ... Let us denote the quadrangle $C_nC_{n+1}C^*_{n+1}C^*_n$, $n = 0, \pm 1, \pm 2, ...,$ by τ_n. Let τ_n be covered by suitable segments of two pensils of straight lines: one of them, Ξ, is the family of straight lines parallel to OY which lie between OY and O^*Y^*; the other one, Γ_n is the pencil of straight lines having the common point A_n. Then, every τ_n is covered by a net; let us denote it by $\{\tau_n\}$. Let us denote suitable segments of Ξ by x_n and suitable segments of Γ_n by g_n. End points of g_n let us denote by G_n and G^*_n respectively (see Fig. 2; all these symbols are marked for $n = 2$).

Definition

The set of transformations

$$\sigma^k(\{\tau_n\}) = \{\tau_{n+k}\},\ k = 0, \pm 1, \pm 2, ...$$
$$n = 0, \pm 1, \pm 2, ...$$

such that for arbitrary n and k the net $\{\tau_n\}$ is transformed into the net $\{\tau_{n+k}\}$ is said to be *Strzemiński's translation* iff:

1° segments x_n and x_{n+k} are on the same straight line;

2° $|C_nG_n| / |G_nC_{n+1}| = |C_{n+k}G_{n+k}| / |G_{n+k}C_{n+k+1}|$, where the symbol $|\cdot|$ denotes the length of the suitable segment.

From the Tales theorem it follows that $|C^*_nG^*_n| / |G^*_nC^*_{n+1}| = |C^*_{n+k}G^*_{n+k}| / |G^*_{n+k}C^*_{n+k+1}|$; the segment g_n is transformed, therefore, into the segment g_{n+k}.

The set of Strzemiński's translations forms a group of transformations. Indeed:

1° σ^0 is the unit element of the set;

2° there exists for any element, σ^k, of the set the reciprocal element, σ^k, which belongs to the set;

3° the composition – defined as it is usually defined in the theory of symmetry – of two elements of the set forms a transformation which belongs to the set;

4° the composition of three elements, σ^k, σ^l and σ^m, of the set is associative:

$$\sigma^k(\sigma^l\sigma^m) = (\sigma^k\sigma^l)\sigma^m.$$

Similarly as any bounded pattern – in the space or in a plane – we talk about, that it realizes a group of symmetry, so we say that Strzemiński's 13. unistic composition "realizes" *in respect of its structure* the above defined group. Quotation marks used here stress that the realization touches the idea not the ultimate shape.

X. BETWEEN CERTAINTY AND UNCERTAINTY

The painting under consideration, apart from the group structure, has a structure which may be called the probabilistic one. Its intentional manifestation is, if I may say so, exceptional. I would like to throw some light – to some limited extent only – on the matter.

When accomplishing the intention to describe Strzemiński's 13. unistic composition from, so to say, the performance point of view, we stand face to face with a paradox. On one side one may not not perceive the certainty with which the painter drives the brush – as it is the master does it always – on the other side one may not avoid notions referring to random phenomena or, it is rather better to say, notions among which the uncertainty belongs to the essence of the matter. It is not, however, an apparent paradox. It is a piece of evidence of the deep difference which we disclose between the artist's certainty and the uncertainty of the language which his works describe. One could look in various ways on the "ungeometriness" – let us use the naive word – of the curves which fill the painting. I have chosen the simplest one. More refined methods would give more interesting results.

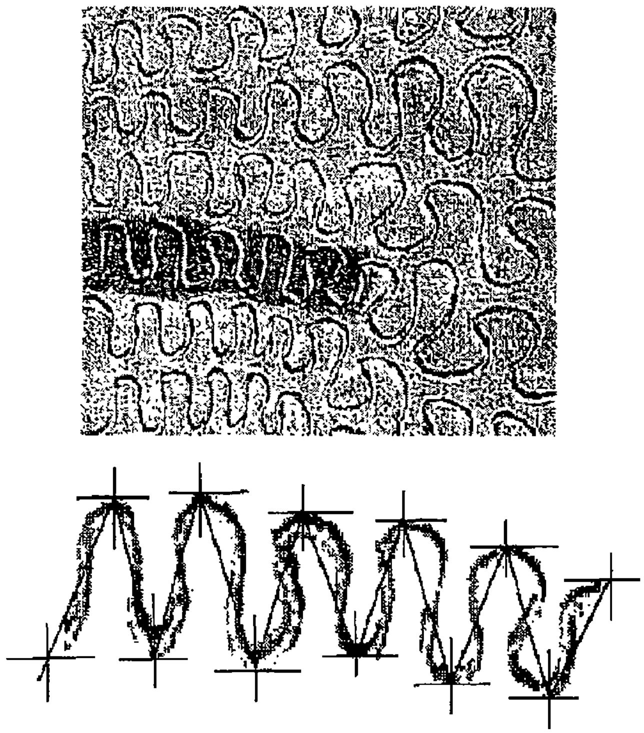


Figure 3

Let us assume that the amplitudes of all the curves of the painting have random sizes. In order to draw a suitable histogram (Fig. 4) amplitudes have been measured as is shown in Fig. 3. It is easy to put together the analytic outlook at Strzemiński's composition with the outlook at nature subjects when there are investigated some of their homogeneous subjects. Such a bringing of these two phenomena face to face, the artist's handwork and nature's "work", yields two questions which I present below. I will not attempt to answer the first one, while the other question not more than partly, i.e., on such a scale as it justifies what I have said in this paper.

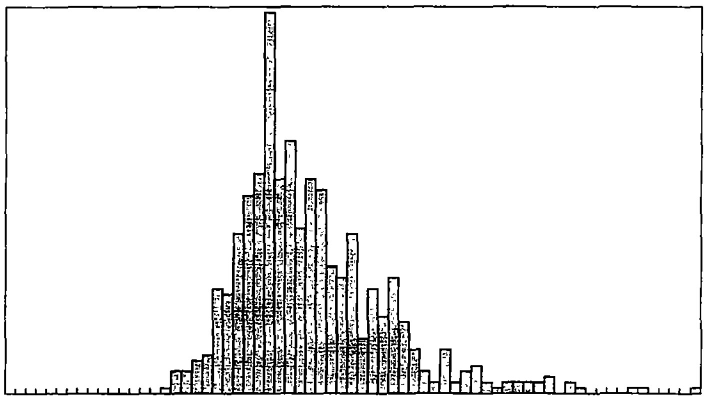


Figure 4

1° In what law strata, in both cases, i.e., in human and nature creativeness, there are hidden the same roots?

2° Has the paradox, we have talked about, its analogy in nature phenomena?

We ask in the second question, if the nature laws "work" at random. If the reality *B* does not exist in a real way, then – as it results from our previous considerations – we should assert that nature "works" with the entire certainty, without any hesitation and without any randomness. Our considerations, however, induce us to be more exact and we should complete the last statement as follows: nature "works locally" with... There arises then the next question: does nature "work" as a whole with unhesitating certainty?

EPILOGUE

Our considerations have been subordinate to the nine questions put in Chapter I. Their arrangement in the table given below seems to be natural and transparent:

A1°	A2°	A3°	B1°	B2°	B3°	AB1°	AB2°	AB3°

All these questions contain the fundamental and vexed points in dispute about the structure of the world which surrounds us, treated either locally or as a whole.

The attitude of my mind to them is the following:

+	–	–	?	?	?	–	–	+

Maybe the symmetry of this table is in some respect, or let me rather say, in a slight respect, a confirmation of the answers' correctness »yes« and »no«.

Our considerations have run parallel, either evidently or in their deeper stratum, along two currents. In one of them we reveal creative, persistently living and persistently reviving energy of the sources having their origins in the reality *B* and – we can state – being realized in the reality *A*. Along the second current we revealed the creative human genius. We contemplated it having in front of us Strzemiński's 13. unistic composition. It had taken its origin in the reality *B* and had been realized in a way approachable to our senses in the reality *A*. Crossing over the borderland – which connects the two realities, visible and unvisible – is the evidence of the human being's spiritual power.

ACKNOWLEDGEMENTS

I would like to express my grateful appreciation to the Wrocław University of Technology for sponsoring the paper, to Mr Romuald Tarczewski for preparing the drawings, to my friends Spencer and Graham, for language correction of the Introduction, to Mrs Marzena Łuczkiewicz for language correction of the rest of the paper and to Mr Stanisław Gancarz for the typographic layout

REFERENCES

[1] EISENBERG, R. and RESNICK, R. (1974) *Quantum Physics of Atom, Molecules, Solids and Particles*, John Wiley & Sons.

[2] KRASNODĘBSKI, R. (1985) On Probabilistic Weights of Biological Hypotheses, *International Society for Ecological Modelling Journal*, pp. 17–50.

[3] SZYMBORSKA, W. (1997) *Nic dwa razy, Wybór wierszy – Nothing Twice, Selected Poems*, Selected and translated by Stanisław Barańczak and Clare Cavanagh, Kraków: Wydawnictwo Literackie.

[4] WEIZSÄCKER, C.F von (1977) *Werner Heisenberg in memoriam*, [in:] C.F. von Weizsäcker und B.L. van der Waerden, *Werner Heisenberg*, München–Wien: Carl Hauser Verlag.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 373-374, 1998

INSTALLATION AND PERFORMANCE FOR REFLECTION OF SOUND AND LIGHT

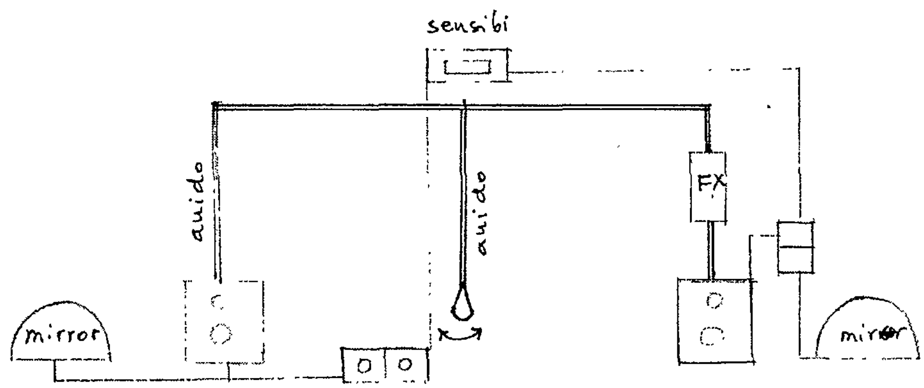
"CONCERTINO FOR ELECTRIC CORDS AND SALT"

Klara Kuchta and Shlomo Dubnov

Addresses: Klara Kuchta Création, 7 rue de la scie, CH-1207 Genéve, Suisse, klara.kuchta@freesurf.ch
Shlomo Dubnov, Hebrew University, Jeursalem, Givat-Ram, dubnov@cs.huji.ac.il

ORDER - Left mirror half sphere moves.
- Pendium swings: produces a regular sound.

DISORDER - Right mirror half sphere moves towards the audience
- Sound delay and echo appear.

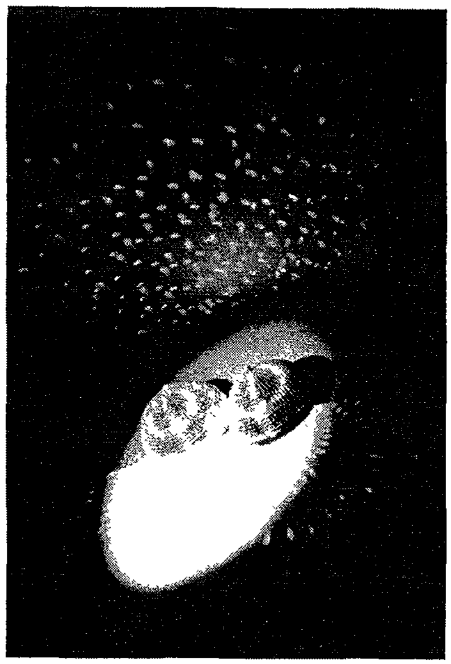


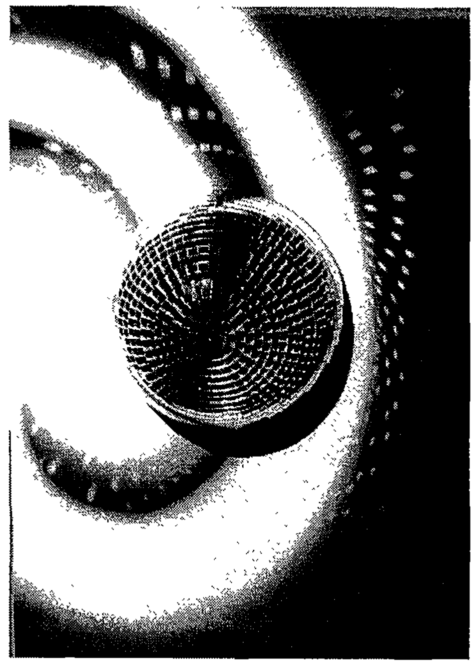
The sounds are generated by a *feedback* effect, which occurs when a microphone is put in front of a loudspeaker.

The microphone is hanged from the ceiling and it swings as a pendulum.

Every time it passes in front of the loudspeaker, feedback sound is added, with many echoes and delays that produce *harmony*, *rhythm*, and light effect.

The direction of movement of the light reflection is produced by the two mirror hemispheres. The sound direction also corresponds, i.e., order is on the left (no effect) and disorder comes from the right (sound with echo and delay).





Symmetry: Culture and Science
Vol. 9, Nos. 2-4., 375-382, 1998

ON THE DEFINITION OF SYMMETRY

Roza Leikin (1) – Abraham Berman (2) – Orit Zaslavsky (3)

Address: (1) Faculty of Education, University of Haifa, Haifa, 31905, Israel (to whom correspondence should be addressed.)
(2) Technion - Israel Institute of Technology
(3) Technion - Israel Institute of Technology

DEFINITION OF SYMMETRY

Many mathematicians and mathematics educators emphasize the role of definition of a mathematical concept for developing conceptual understanding. The analysis of mathematical and educational literature indicates that the meaning of symmetry is not precisely defined. In addition, there are several different approaches to the definition of symmetry, depending on the content, the perspective taken, and consequently on the types of symmetry. Thus, symmetry is often viewed as a collection of disconnected concepts. In this paper we propose a formal definition of symmetry reflecting a unifying approach to the concept of symmetry in mathematics. This definition lends itself to considering a hierarchy of the different types of symmetry.

BACKGROUND

Symmetry is a very important scientific concept. It appears in different branches of mathematics and connects them. Symmetry often helps in making mathematical proofs elegant and can be considered as one of the important heuristics in mathematical problem solving. Nevertheless, the notion of symmetry is not precisely defined (Lowrey 1989). There are several different approaches to the definition of symmetry, depending on the perspective taken. For example, in the Russian dictionary of foreign words (Pchelkina 1988) four different meanings for the Greek word $\Sigma\upsilon\mu\mu\varepsilon\tau\rho\iota\alpha$ are given:

- corresponding proportions between parts of a whole or of a body;
- a property of a geometric figure for which the figure can coincide with itself, in a way that not all its points remain in the same place;
- a global property of nature connecting the laws of conservation of energy, of movement, of atom and molecular structure, and of crystals' structure;
- a mutual relationship between the parts of the body with respect to an axis, a point or a plane.

Mathematicians treat symmetry in various ways: as a property of an object, as a special relation between objects, or as a special kind of transformation. In teaching mathematics, teachers and textbooks usually distinguish between symmetry in geometry and symmetry in other branches of mathematics. Even in geometry they deal separately with symmetrical geometric figures and with different transformations (reflection, rotation, and translation), neglecting to point out the underlying common feature of all these figures, transformations, and relationships (Eccles 1972; Ellis-Davies 1986; Fehr, Fey & Hill 1973; Lowrey 1989; Marcus 1989; Seneschal 1989; Skopets 1990; Yaglom 1962). In algebra and *calculus symmetry* is defined differently for different kinds of objects (e.g., functions, systems of equations, matrices, groups, mathematical problems) (Daintith & Nelson 1989; Dreyfus & Eisenberg 1990; Polya 1973; Waterhouse 1983). For example, according to Polya (1973) the expression $xy+yz+zx$ is symmetrical because an exchange of variables does not change its value. Thus, the variables have symmetrical roles. Polya used the notion *role symmetry* to express this type of symmetry. Another type of role symmetry, i.e., logical symmetry, can be found in the mathematical and educational literature. For example, Dhombres (1993) claims that symmetry exists between any two different proofs for one mathematical statement. These proofs have symmetrical roles with respect to the statement they prove. This logical symmetry of proofs has many implications. For example, Silver, Mamona-Downs, Leung & Kenney (1996) refer to symmetrical change as one of the strategies in posing mathematical problems.

As a result of this variety of types of symmetry and the differences between them symmetry is often viewed as a collection of disconnected concepts. An analysis of mathematical dictionaries supports this claim (Borowski & Borwein 1991; Clapman 1990; Daintith & Nelson 1989; Schwartzman 1994). For example, Daintith & Nelson (1989) define separately a symmetric function, a symmetric matrix, and a symmetric relation and only then define symmetry as follows:

"*Symmetry:* In general, a figure or expression is said to be symmetric if parts of it may be interchanged without changing the whole. For example $x^2+2xy+y^2$ is symmetric in x and y. *A symmetric operation (symmetry)* is an operation on a figure or expression that produces an identical figure or expression...." (Daintith & Nelson 1989, p. 313)

Many mathematics educators emphasize the important role of the definition of a mathematical concept for building mathematical understanding (De Villers 1994; Fischbein 1987; Moore 1994; Vinner 1991; Wilson 1989). Only a small number of authors discuss symmetry in a more general sense (Leikin, Berman and Zaslavsky 1995; Marcus 1989; Rosen 1989; Rosen 1995; Sonin 1987; Stewart & Golubitsky 1992, Weyl 1952). In these cases they refer to symmetry as proportion, harmony, order, or repetition. A general definition of symmetry in science is given by Rosen (1995) as follows: *Symmetry is immunity to a possible change* (p. 2, ibid.).

In this paper we suggest a unifying approach to the definition of symmetry in mathematics that is similar to what Rosen does for symmetry in science. We look at the immunity of a *property* of a mathematical *object* with respect to a possible change. This possible change corresponds to a *transformation* that can be applied to the object. Thus, symmetry has to do with a triplet — an *object*, a *property* and a *transformation* — as proposed in the following definition.

THE PROPOSED DEFINITION OF SYMMETRY

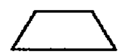
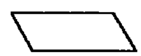
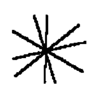
Definition. Symmetry is a triplet (S, Y, M) consisting of an *object* (S), a specific *property* (Y) of the object, and a *transformation* (M) satisfying the following two conditions:

i) The object belongs to the domain of the transformation;

ii) Application of the transformation to the object does not change the property of the object.

The *object* S is said to be *symmetrical* and the *transformation* M is called a *symmetry transformation.*

We refer to *two main types of symmetry:* (i) *Geometric Symmetry* where the object S is a geometric figure, and (ii) *Role Symmetry* where the transformation M is a permutation. We distinguish between different types of geometric symmetry according to the different types of geometric transformations: isometries (reflection, rotation, translation) and non-isometries (homothety). We consider different types of role symmetry according to the type of an object (algebraic, logical, and geometric). Figures 1 and 2 present examples of the different types of geometric and role symmetry.

TYPE OF SYMMETRY	OBJECT (S)	PROPERTY (Y)	TRANSFORMATION* (M)
GEOMETRIC SYMMETRY			
Isometric Symmetry	•An equilateral trapezoid 	• The location of the figure	• Reflection over a line
	• A quadratic function • An even function	• The location of the graph	
	•A parallelogram 	• The location of the figure	• Reflection over a point
	• An odd function	• The location of the graph	
	•A periodic function	• The location of the graph	• Translation
Non-Isometric Symmetry	•A family of straight lines passing through one point 	• The location of the figure	• Homotety

* Note that the given definition of symmetry applies also to three-dimensional geometrical objects where any plane symmetry transformation can be extended to a solid symmetry transformation

Figure 1: Examples of different types of Geometric Symmetry

Observe that our definition of symmetry includes some trivial cases. We will say that a symmetry (*S, Y, M*) is a *trivial symmetry* if property *Y* is immune to transformation *M* for *any* object *S* having this property. Examples of trivial symmetries are: (a) Any triplet (*S, Y, I*) where *I* is the identity transformation; (b) Any triplet (*G, D, IS*) where *G* is a geometric figure, *D* denotes distance, and *IS* is an isometric transformation.

TYPE OF SYMMETRY	OBJECT (S)	PROPERTY (Y)	TRANSFORMATION* (M)
ROLE SYMMETRY			
Algebraic Role Symmetry	▪ An algebraic expression, e.g. $a+b$,	▪ The numerical value of the expression	▪ Permutation of variables
	▪ A function, e.g. $xy=8$;	▪ The graph of the function	
	▪ A systems of equation, e.g.. $\begin{cases} 3x+y+2z=30 \\ 2x+3y+z=30 \\ x+2y+3z=30 \end{cases}$ (Polya, 1981)	▪ The solution of the system	
Geometric Role symmetry	▪ An isosceles triangle $\triangle$	▪ The location of the triangle ▪ The equality of the sides	▪ Permutation of the equal sides
Logical Role symmetry	▪ A symmetrical relation, e.g.: $a \parallel b$, $G \cong F$	▪ The correctness of the statement	▪ Permutation of the related objects
	A Symmetrical solution, e.g. *Problem:* Of all the triangles inscribed in a given circle which one has the maximal area? *Solution:* If side b and side c of the triangle are not equal, then the area of the triangle can be increased. Hence $b=c$. In the same way, $a=b$. By analytical considerations there exists triangle of maximal area, and by the preceding steps it must be equilateral.	▪ The correctness of the solution	▪ Permutation of steps in the solution

Figure 2: Examples of different types of Role Symmetry

RELATIONSHIP BETWEEN ALGEBRAIC AND GEOMETRIC SYMMETRIES

Relationships between different types of geometrical symmetry transformations are widely discussed in the mathematical literature (for example, see Yaglom 1962). In this section we consider relationships between algebraic and geometric symmetry of mathematical objects.

A mathematical object may have different representations. For example, a quadratic function can be represented algebraically as $y = ax^2+bx+c$, or geometrically, as a parabola. We consider an object S to be symmetrical if it is symmetrical at least in one of its representations. If an object is symmetrical in its algebraic (geometric) representation we will say that it is algebraically (geometrically) symmetrical. The relations between these two types of symmetry are given in the following two statements (Leikin 1997).

Statement 1. If a mathematical object has both a geometric and an algebraic representation and if it is algebraically symmetrical, then it is geometrically symmetrical.

Statement 2. Any mathematical object having both a geometric and an algebraic representation is algebraically symmetrical if and only if there exists an isometric transformation (M) satisfying the following conditions: (1) M does not change the location of the object in its geometric representation; (2) M maps each coordinate axis to one of the coordinate axes without change of the axis's direction.

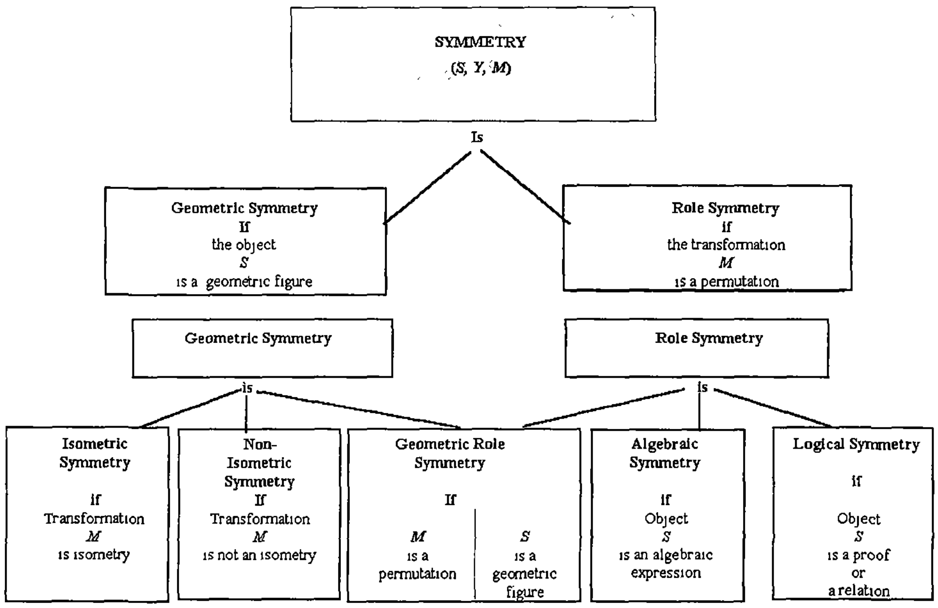


Figure 3: Hierarchy of the Concept of Symmetry in School Mathematics

HIERARCHY OF TYPES OF SYMMETRY

As shown above, the proposed definition of symmetry, which reflects a unifying approach to the concept of symmetry, lends itself to considering its different types as particular cases of the general notion of symmetry. This definition highlights the hierarchical connections between the different types of symmetry. Figure 3 depicts a hierarchy of the concept of symmetry, which depends on the types of symmetrical mathematical objects and on the types of symmetry transformations.

REFERENCES

Borowski, E. J. & Borwein, J M. (1991) *The Harper Collins Dictionary of Mathematics,* Harper Perennial: A Davidson of Harper Collins Publishers

Clapman, C. (1990) *A Concise Oxford Dictionary of Mathematics*, New York· Oxford University Press

Daintith, J. & Nelson, R. D , eds. (1989) *Dictionary of Mathematics*, New York· Penguin Books

De Villers, M. (1994) The Role and Function of Hierarchical Classification of Quadrilaterals, *Learning of Mathematics* 14(1), pp. 11-18.

Dhombres, J. (1993) Is One Proof Enough? Travels With a Mathematician of the Baroque Period, *Educational Studies in Mathematics 24*, pp. 401-419

Dreyfus, T. & Eisenberg, T (1990) Symmetry in Mathematics Learning, *ZDM — International Reviews on Mathematical Education* 2, pp. 53-59

Eccles, F. M. (1972) Transformations in High School Geometry, *Mathematics Teacher 65*, pp. 103, 165-169

Ellis-Davies, A (1986) Symmetry in the Mathematics Curriculum, *Mathematics in School 15*(3), pp. 27-30

Fehr, H. F., Fey, J T & Hill, T. J. (1973) *Unified Mathematics. Course I. Chapter 11. Transformations in a Plane*, California: Addison-Wesley Publishing Company.

Fischbein, E. (1987) *Intuition in Science and Mathematics. An Educational Approach,* Kluwer Academic Publishers, USA

Leikin, R (1997) *Symmetry as a Way of Thought - a Tool for Professional Development of Mathematics Teachers,* Doctorate Dissertation, Technion - Israel Institute of Technology, Haifa, (In Hebrew)

Leikin, R., Berman, A. & Zaslavsky, O. (1995) The Role of Symmetry in Mathematical Problem Solving. An Interdisciplinary Approach, *Symmetry· Culture and Science, Special Issue. Symmetry Natural and Artificial* 6(2), pp 332-335.

Lowrey, A. H. (1989) Mind's Eye, in Hargittai, I., ed *Symmetry 2, Unifying Human Understanding,* pp. 484-503, UK: Pergamon Press.

Marcus, S. (1989) Symmetry in the Simplest Case: The Real Line, in Hargittai, I , ed *Symmetry 2, Unifying Human Understanding,* pp 103-115, UK: Pergamon Press.

Moore, R. C. (1994) Making the Transition to Formal Proof, *Educational Studies in Mathematics* 27, pp 249-266.

Pchelkina, V , ed. (1988) *Dictionary of Foreign Words*, Moscow: Russky Yazik (Rus.).

Polya, G. (1973) *How To Solve It. A New Aspect of Mathematical Method*, New Jersey. Princeton University Press

Rosen, J. (1989) Symmetry at the Foundations of Science, in Hargittai, I., ed. *Symmetry 2, Unifying Human Understanding,* pp. 13-15, UK. Pergamon Press.

Rosen, J. (1995) *Symmetry in Science An Introduction to the General Theory*, New York· Springer-Verlag.

Schwartzman, S. (1994) *The Words of Mathematics. An Ethymological Dictionary of Mathematical Terms Used in English*, USA. MAA.

Seneschal, M. (1989) Symmetry Revisited, in Hargittai, I., ed. *Symmetry 2, Unifying Human Understanding*, pp 1-12, UK: Pergamon Press.

Silver, E. A., Mamona-Downs, J , Leung, S. S & Kenney, P. A. (1996) Posing Mathematical Problems: An Exploratory Study, *Journal for Research in Mathematics Education* 27, pp. 293-309.

Skopets, Z. (1990) *Geometric Miniature*, Moscow: Prosveshenie (Rus)

Sonin, A. S. (1987) *Comprehension of Perfection*, Znanie: Moscow (Rus.).

Stewart, I & Golubitsky, M. (1992) *Fearful Symmetry. Is God A Geometer?*, New York: Penguin Books.

Vinner S. (1991) The Role of Definitions in the Teaching and Learning of Mathematics, in Tall, D. O , ed *Advanced Mathematical Thinking*, Dordrecht: Kluwer.

Waterhouse, W C. (1983) Do Symmetric Problems Have Symmetric Solutions?, *The American Mathematical Monthly 90*, pp. 378-387

Weyl, H (1952) *Symmetry*, New Jersey: Princeton University Press.

Wilson, P. S. (1990) Inconsistent Ideas Related to Definitions and Examples, *Focus on Learning Problems in Mathematics 12*(3 & 4), pp. 31-47

Yaglom, I M. (1962) *Geometric Transformations. Vol 1. Displacements and Symmetry*, New York: Random House.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 383-392, 1998

THE LANGUAGE OF CELLS: A PARTITIONAL APPROACH TO CELL-SIGNALING

Pedro C. Marijuán – José Pastor – Morris Villarroel

Address. Department of Electronic and Communication Engineering, CPS, Universidad de Zaragoza, 50015 Zaragoza, Spain E-mail: marijuan@posta.unizar es

Abstract: *Cell signaling is the fastest growing subject in biochemistry yet no general mathematical principles have been found to decipher a "language" of cells using the discrete positional approach (Shannonian Information Theory does not appear to provide a workable framework). Instead, cellular information may be compositional; messages are exchanged as the presence or absence of symbols and few meaningful positional relationships are involved. Here we introduce a new informational grammar using uni- and multi-dimensional partitions that may help us to better understand signal processing in the eukaryotic cell.*

1. INTRODUCTION: POSITIONAL VERSUS COMPOSITIONAL INFORMATION

The notion of a "language" of cells does not seem consistent with the standard views of Information Theory applied to biology. Although Shannon (1949) distinguished between discrete, continuous, and mixed information sources, the standard application (and possibly overextension) of his ideas to cell biology have been heavily influenced by the sequential structure of DNA and RNA and, traditionally, only the discrete-positional case has been considered (e.g., Gatlin 1972, Schneider 1995). As a consequence, the lack of distinction between "positional" and "compositional" forms of information and the subsequent neglect of the latter have implied an analytical dead-end concerning the possibilities of elucidating formal mechanisms of cellular languages.

The assumed preconditions for information transmission, and particularly for any workable language, refer to sequences of messages containing combinations of symbols which are deciphered or transmitted always following a positional order (only broken for crytographic purposes – Pastor and Sarasa 1998). Shannon's formula appears to be the natural way of measuring the average combinatory content of these positional messages and of establishing their relative index of surprise in order to design appropriate channels, codes, etc. Subsequently, a workable language can be created by following a set of grammatical (Markovian) rules to connect successive positional messages comprised within the dictionary scope of the language.

However, one can point to a number of instances in natural and social communication where symbols are used in a rather different way. Instead of a "positional" context (which also generally implies the assumptions of sequence, stability and hierarchy – see Marijuán and Villarroel 1998) symbols may be used in a "compositional" way. In this alternative context, messages are exchanged as presences or absences of symbols which have been accumulated upon predetermined sets of objects. No meaningful positional relationships are assumed among the objects within the set or among the symbols accumulated on these objects. For example, several glasses on a tray may contain a variable number of different symbolic items (ice cubes, soda, vermouth, olives, cherries). We may consider the set of glasses on the tray as the message, each glass being an individual object that accumulates several symbols which make it distinguishable. Then two subjects could communicate by exchanging trays with a variable number of glasses and contents (Marijuán and Pastor 1998). That messages can be reliably distinguished and transmitted by the "concurrent processing of discrete states of media", has already been postulated by Karl Javorszky (1995). A whole body of partitional calculus (or granularity algebra) has been envisioned by this author (Javorszky 1995b, Steidl and Javorszky 1996). Interestingly, partitional reasoning has also been applied to problems in pattern recognition (Frigui and Krishnapuram 1997), logic (Mosterín 1987, Modica and Rustichini 1994) and even economics (Caianiello 1985).

Biological examples of compositional information exchanges may be found in the communicational use of colors, odors and tastes. We may also consider pheromones in social insects and, anecdotally, the etiquette "language of flowers", and perhaps even musical compositions and the formative frequencies of vowels and consonants of our own spoken languages. The "language of cells" we shall discuss here may be one of the most interesting instances of communication by means of such compositional tools; and it has been the forerunner of all further means of biological communication. Marshall McLuhan's famous dictum "the medium is the message" and the particular disdain this author showed about Shannon's information theory (McLuhan 1962) are worth recalling when considering this fundamental distinction between positional and compositional forms of information exchange.

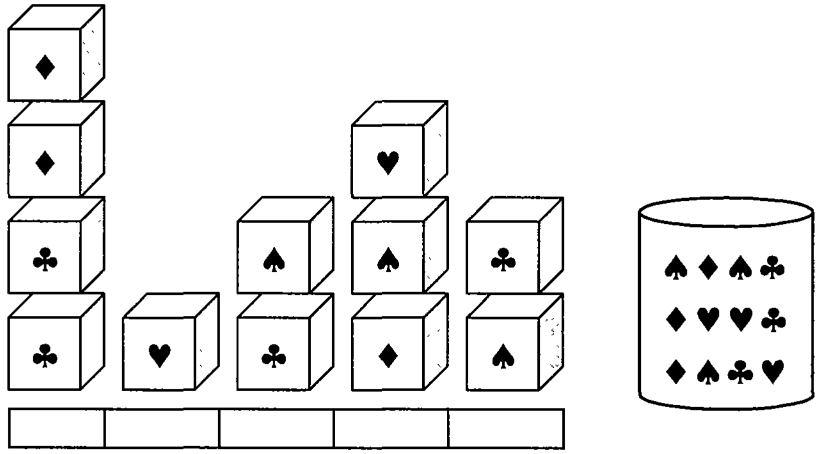
2. ANALYZING A COMPOSITIONAL MESSAGE

2.1 Unidimensional Partitions

The theory of compositional messages is formally based on the partitional-additive properties of natural numbers. Following this theory, messages are distinguished and analyzed by measuring the relative frequency of partitions in the overall structure of the message. Mathematically speaking, partitions are a very straight forward concept, i.e., the additive decompositions of natural numbers. For instance, the set { (5), (4,1), (3,2), (3,1,1), (2,2,1), (2,1,1,1), (1,1,1,1,1) } represents all the unidimensional partitions of the number five. By adhering to this mathematical treatment, one can use the well-known partitional properties of numbers to discuss the most probable logical states of a compositional message encrusted upon the elements of the set N.

Compositional messages have to be mapped onto a different kind of state space, where what counts (what generates dynamics) is the absence or presence of specific symbols on a set of N elements. When receiving a compositional message, the presence of the different symbols on each element of the set has to be counted and grouped in homogeneous classes of overlapping or non-overlapping nature. For instance, the message of Figure 1 generates the following partitions of 5: hearts (3, 2), spades (3, 2), clubs (2, 2, 1), diamonds (3, 1, 1).

Each class is defined by the presence of a specific symbol, and this symbol effectively creates a partition of the set of N elements. After the classes are defined by single symbols, the more complex coincidences of combinations of symbols (class overlaps) among the elements can also be considered. It can be easily proved that, in the first case of linear of unidimensional partitions for single symbols, all the possible countings of symbolic presences among the N elements of the set lead to the whole set of partitions of N, called $E(N)$. The successive consideration of two, three, four symbols, etc. can then be considered multidimensional partitions and their mutual coincidences would generate families of unidimensional second, third order partitions, etc.



Message set:	the five slots
Elements:	each one of the slots
Symbol:	the distinctions of each cube
Object:	one slot with the symbols on it
Sign:	the group of symbols on the element

Figure 1: A Compositional Message - An indefinite number of symbols (taken here from a "jar") are placed upon the discrete elements that make up the compositional message No positional relationships of order are involved.

It is worth noting that, whereas Shannonian entropy increases with the total number of symbols, partitional entropy reaches a limit. Thus, not only do partitions convey the abstract "form" of messages but they establish boundaries on the state space of possible messages using three important logical principles that characterize this approach:

The Principle of Parsimony precludes the addition of a symbol *different* from those already present if that symbol does not introduce further distinctions. It follows that a maximum of N-1 *different* symbols may accumulate on a single element,

e.g., the message [AXYZH, A, 0, 0] = [AB, A, 0, 0]

The Principle of Economy precludes the addition of a symbol the *same* as those present if this symbol does not introduce further distinctions (redundant symbols on all objects will not be perceived by the receiver). It follows that, a maximum of N-1 *equal* symbols may accumulate on a single element,

e.g., the message [AAA, AA, 0, 0] = [AA, A, 0, 0]

The Principle of Symmetry precludes the distinctions derived from the mutual exchange among symbols. So, commutative relations apply among the set of symbols,

e.g., the message [AB, AA, B, 0] = [AB, BB, A, 0]

2.2 Kmax- the most probable partitional state

After the above principles have been applied, the set of partitions $E(N)$ can be immediately transformed into a probability body (for the unidimensional case). The probability of any state of the set to exist as described by a specific partition is given by the relative frequency of this partition among all partitions. For instance, on $E(5)$ the probability is 1/7 - for states (5), (2,1,1,1) and (1,1,1,1,1), 2/7 - for states with either 2 or 3 summands each, 15/20 - for any summand to be an odd number, etc.

$E(N)$ is obtained by Ramanuyan:

$$E(N) = 1\frac{1}{\pi\sqrt{2}}\sum_{k=1}^{\infty}\sqrt{k}\,A_k(n)\frac{d}{dn}\frac{\sin h\left\{\frac{\pi}{k}\sqrt{\frac{2}{3}}\sqrt{n-\frac{1}{24}}\right\}}{\sqrt{n-\frac{1}{24}}}$$

(see Javorszky 1995)

The partition with the highest relative frequency is called the Kmax. In this most probable partitional state, the set shows Kmax distinct summands with respect to a one-describing dimension. In the case of $E(5)$, there is a Kmax shared both by 2 { (4,1) (3,2) } and 3 { (3,1,1) (2,2,1) }.

Heuristically, it appears that a compositional message can be univocally described by its corresponding "trace" of unidimensional partitions (Steidl and Javorszky 1996), if a few additional statistical measures that act as a sort of context or shared background in the communicational process have been previously established: most probable message length, ratio of symbols/elements, structural depth, shallowness, etc. Then the possible use of partitions of further combinations of symbols becomes redundant – and its inclusion would notably complicate the mathematical description of the message. Only the unidimensional-multidimensional problem may be pertinent.

The Kmax of every property or symbolic presence may be used as the origin or natural cannon to which the respective deviations of successive messages can refer. Kmax is the main message. Therefore, the information in a complex compositional message is represented by a comparatively small collection of distinguishing maxima instead of an overwhelming collection of symbols (and their permutations and combinations) that make up the message. The Kmax of a message would represent a natural property to which successive messages can refer. This simplifies the description of a specific message in the context of a continuous communication process and may speed up the process while diminishing errors.

2.3 Multidimensional partitions

Karl Javorszky (1995) has argued that an efficient massively parallel communication procedure – using multidimensional partitions – can be built around minimized partitional traces of the above Kmax. It seems to work particularly well with data sets of moderate size, which are preferably prestructured and come in a quasi continuous stream, so that the number of possible symbols is always kept rather finite. Although symbols might come from an infinite multitude, there should be a relatively small collection of distinguishing items employed at the communicational session, and their group relations should not generate a cardinality overstatement symbols/elements above a certain limit.

To the extent that Javorszky's estimates are correct, the overall capacity of a multidimensional compositional channel making use of discrete states of media can be generically expressed as:

$$T(N) = E(N)\ exp\ ln\ E(N),$$

where $T(N)$ is the number of different logical states which can be distinguished by means of collections of symbols put on the elements of the set N. Only non-redundant states are counted (see the principle of economy), because redundant symbol groups can always be substituted by single symbols, coalescing into a unique logical state. $E(N)$ is the already mentioned number of unidimensional partitions of the set N.

It is also interesting to compare $T(N)$ and the strictly positional use of the same elements of the set N in a combinatory way. According to the positional approach, a total of $N!$ different messages or logical states can arise using the same elements.

Surprisingly, $T(N)$ yields a larger number of logical states than $N!$ for values of N in between 31 and 95, with a maximum around 63-64. However, for N=12, the number of combinations $N!$ reaches a maximum with respect to $T(N)$. Apparently, several parameters of the genetic code would correspond with such max./min. extremes that characterize the compositional-positional interrelationship (see Javorszky 1995, for a detailed expression of all these formulae and calculations). Even a cursory analysis of the multidimensional partitions for N=3 and N=4 shows the emergence of an intricate "geometrical" (compositional) realm where symmetry patterns can be replicated by means of partitional operations (Villarroel, Pastor and Marijuán, in prep.).

2.4 The Emergence of Power Laws

Numerical partitions are characterised by exponential growth, and, heuristically, compositional messages have elements which are, by their very nature, contingent. The set seems to follow power laws and thus be vulnerable to small changes (Bak 1996). It is worth noting that a very simple way to obtain a power law is by the superposition of the whole partitional summands of a given number. We have graphed numerical partitions up to 20 (see Figure 2) and, except for some interference in the frontiers of the numeric interval, every summand's relative presence cleanly depends on a power law (Marijuán and Villarroel 1998). We have yet to solve for the general expression of the exponent of the power law which may describe partitional growth.

The power law theme leads towards a physical paradigm that apparently shares basic formal properties with the above compositional dynamics: self-organized criticality. The generalization built upon the well-known sand-pile paradigm that seems to apply to numerous natural phenomena (geologic, chemical, physical ones), leaving the characteristic signature of "power laws" in the involved structures and processes (Bak 1996), could also apply to the critical exchange of compositional messages that biosocial informational entities are collectively orchestrating by means of their communicational activities superimposed upon the structural ones... The fact is that power laws are omnipresent in cellular, organismic, economic, and social realms too (Scarrott 1996; Bak 1996; Marijuán 1998). Inescapably, this biologically-inspired train of thought on compositional messages has to be linked not only with self-organized criticality but also with the engineering-inspired views of G. C. Scarrott on "recurrence" (and related power laws) postulated as one of the basic tenets of natural information systems.

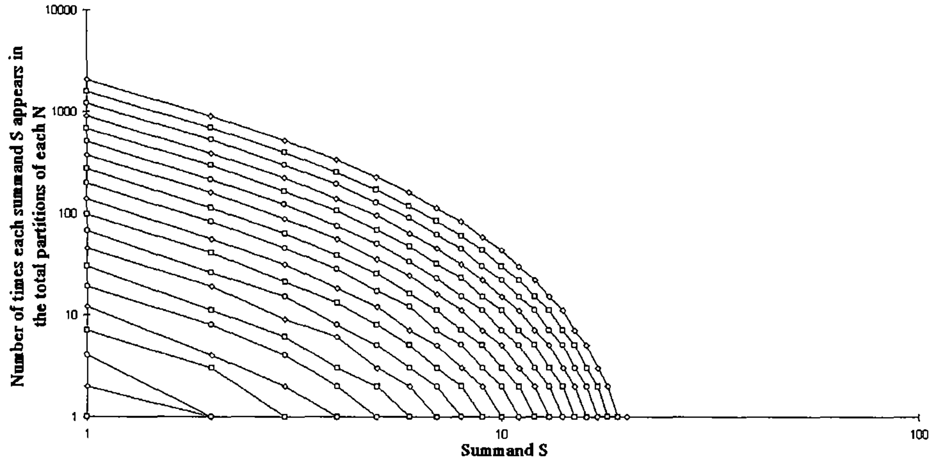


Figure 2: Power laws in partitions. Each curve stands for a particular value of N starting from 1 until 20. The ordinate represents the number of times each summand appears in the total partitions of each N. The x-axis represents the summand S

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
S																				
1	1																			
2	2	1																		
3	4	1	1																	
4	7	3	1	1																
5	12	4	2	1	1															
6	19	8	4	2	1	1														
7	30	11	6	3	2	1	1													
8	45	19	9	6	3	2	1	1												
9	67	26	15	8	5	3	2	1	1											
10	97	41	51	13	8	5	3	2	1	1										
11	139	56	31	18	12	7	5	3	2	1	1									
12	195	83	45	28	17	12	7	5	3	2	1	1								
13	272	112	63	38	25	16	11	7	5	3	2	1	1							
14	373	160	87	55	35	24	16	11	7	5	3	2	1	1						
15	508	213	122	74	50	33	23	15	11	7	5	3	2	1	1					
16	684	295	164	105	68	47	32	23	15	11	7	5	3	2	1	1				
17	915	389	222	139	94	63	45	31	22	15	11	7	5	3	2	1	1			
18	1212	526	298	190	126	89	61	44	30	22	15	11	7	5	3	2	1	1		
19	1597	686	395	250	170	117	84	59	43	30	22	15	11	7	5	3	2	1	1	
20	2087	911	519	336	226	159	112	82	58	43	30	22	15	11	7	5	3	2	1	1

3. A PARTITIONAL APPROACH TO CELLULAR COMMUNICATION

How can cells reliably communicate without any consideration about positional order in the "letters" of the chemical "words" they exchange? The experimental evidence is that every organismic cell, and every tissue, has sculpted its own coding and decoding apparatus, the Cellular Signaling System, basicly devoted to the analysis of communicational concentrations found in the extracellular-intracellular milieu, i.e., compositional messages. It implies the combined workings of thousands of receptors, hundreds of related protein kinases and phosphatases, and less than ten second messengers, all of them interconnected in order to make sense of the incoming stream of diluted messages. A very complex array of internal states (control of functionality, growth, cell-cycle stages, migration, apoptosis...) is regularly communicated among cells, tissues, and organs by means of such a peculiar molecular-processing apparatus.

In the Shannonian sense, one could conceive of a superimposed state-space built out from the whole variable concentrations that participate in the communication games, so that message patterns would map onto cellular states or onto molecular actuators leading to such cellular states (an automata table, or a grammar could be built). But there appear troubling evidences. The molecular adaptation of receptors (for instance by methyl or phosphate groups), the abundance of sigmoid curves and saturating effects, the vertical organization of signaling pathway components in "transducisomes", and the generalized cross talking among such pathways imply quasi-unsurmountable barriers for handling a regular information-thermodynamic state space. The tools used to describe physical states may not necessarily be meaningful for the description of a communicational space in the cellular milieu (Marijuán and Villarroel 1998).

Instead of the classical analysis of DNA sequences, it seems that the natural target to explore the possibilities of the partitional approach should be the "mysterious" processing operations performed by the cellular signaling system. In this sense, the system of receptors, membrane-bound enzyme and protein complexes, second messengers, and the dedicated kinase and phosphatase chains, could be understood as an abstract partitional processing-system capable of extracting the relative information differences within the stream of incoming compositional messages and physically transport these differences down to final effectors at the nucleus, cytoplasm, or membrane. That's the basic hypothesis that we are presently trying to explore.

If (and what a big "if") cells would make use of formal tools of logico-partitional nature in their management by means of the cellular signaling system of the compositional messages they receive, then the notion of a genuine cellular language, with specific dialects for every organismic tissue, could be seriously argued. And perhaps more interesting than that, quite a few other bizarre aspects of the signaling system could receive some more formal (and simpler) treatment: the cross-talk between signaling pathways, the checkpoints relating signaling operations with cell-cycle stages, the chaotic fluctuations of second messengers, and even the widespread formation of aggregates and complexes (transducisomes) among signaling components.

The studies by Caianiello (1985) on the partitional dynamics inherent in monetary systems and the suggestion by one of us (Marijuán 1998) about the "currency" role played by the set of second messengers in the internal measurement of cellular function might finally be stepping stones pointing out in the same direction: the foundations of information processing in nature and society.

REFERENCES

Bak, P. (1996) *How Nature Works· The Science of Self-organized Criticality*, New York: Copernicus.

Caianiello, E R. (1985) A Thermodynamic Approach to Self-Organizing Systems, in Yates, F E, Garfinkel, A. , Walter, D.O., Yates, G.B., eds. *Self-Organizing Systems, theEmergence of Order*, New York. Plenum Press.

Frigui, H. and Krishnapuram, R (1997) Clustering by competitive agglomeration, *Pattern Recognition* 30, pp 1109-1119.

Gatlin, L L. (1972) *Information Theory and the Living System*, Columbia: New York.

Javorszky, K. (1995) *Granularity Algebra*, Vienna: Mackinger-Verlag

Javorszky, K (1995b) The Logic of Self-Sustaning Sampling Systems, *IPCAT95 Proceedings*, pp. 386-398, University of Liverpool, UK.

Marijuán, P. (1996) Gloom in the Society of Enzymes: on the Nature of Biological Information, *Biosystems* 38, pp 163-171

Marijuán, P. (1996b) Information and Symmetry in the Biological and Social Realm: New Avenues of Inquiry, *Symmetry Culture and Science* 7(3), pp. 281-294.

Marijuán, P and Pastor, J. (1998) *The Language of Cells Foundation of Information Science*, http //fis.iguw.tuwein ac.at/igw/fis/fis htm.

Marijuán, P and Villarroel, M (1998) On Information Theory Stumbling Blocks. Some Biological Considerations about the Concepts of "Sequence", "Stability" and "Hierarchy", *Cybernetics*, in press.

Marijuán, P C. (1998) Information and the Unfolding of Social Life· Molecular Biological Resonances Reaching up to the Economy, *Biosystems* 46, pp. 145-151.

McLuhan, M (1962) *The Gutenberg Galaxy*, Toronto· University of Toronto Press

Modica, S. and Rustichini, A (1994) Awareness and Partitional Information Structures, *Theory and Decision* 37(1), pp 107-124.

Mosterín, J (1987) *Conceptos y teorías en la ciencia*, Madrid: Alianza Editorial.

Pastor, J. and Sarasa, M A. (1998) *Criptografía Digital Fundamentos y aplicaciones*, Zaragoza, Spain Prensas Universitarias de Zaragoza.

Scarrott, G (1996) *The purpose and nature of information*, (comm. workshop: Information, Instrument for the Survival of Society, UK Institution of Electrical Engineers, London, October 1996).

Schneider, T.D (1995) *New Approaches in Mathematical Biology Information Theory and Molecular Machines. Physics and the Origin and Evolution of Life*, Trieste. Italy.

Shannon, C E (1949) *A Mathematical Theory of Communication*, Chicago. University of Illinois Press

Steidl, R. and Javorszky, K (1996) *Message transmission by means of counting states of sets*. Unpublished manuscript.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 393-399, 1998

EXPERIMENTS WITH SOAP BUBBLES

Göran Rämme

Address: Department of Physical Chemistry, University of Uppsala, Box 532, S-751 21, Uppsala, Sweden; fax: 46(0)18-508542. E-mail: goran.ramme@tki.uu.se.

Abstract: *The lecture comprises some selected experiments with soap bubbles demonstrating:*

a) a simple experimental technique to illuminate a hemispherical soap bubble in such a way as to create brilliant rainbow colors in order to study surface phenomena,

b) a new experimental method to determine the lifetime of soap bubbles statistically,

c) bursting of soap bubbles.

INTRODUCTION

A soap bubble is actually an example of a supramolecular spontaneously self-organized system. Soap films and bubbles, single or in combination, often exhibit an inherent symmetry of remarkable beauty. They can attract young and old as well as scientist and layman due to their diverse geometrical shapes, brilliant rainbow colors and fragility. Experiments with soap films and bubbles have the capacity to bridge Art and Science and can be used for entertaining, educational or scientific purposes. Moreover, as Caspar Schwabe pointed out to me:

The surface of a soap bubble is a direct visual manifestation of the three space coordinates and that of time, i. e., reflects the fourth dimension!

For an entrance to the literature of soap films and bubbles see refs.[1-8].

EXPERIMENTS

a) Experimental method to study events on the surface of a soap bubble

The history of a soap bubble as reflected by processes in the soap film from its birth to its ultimate collapse presents an ever changing scenario where many interesting observations can be made.

An experimental method to study these phenomena is to inflate a hemispherical bubble blown directly on a milk glass sheet of an ordinary light table. In this way the bubble becomes evenly illuminated by diffuse light and the soap film dome will after a while display marvelous iridescent rainbow color bands due to interference effects. The colors appear when the bubble has become sufficiently thin, of the order of the wavelength of visible light, i. e., 400 - 800 *nm*.

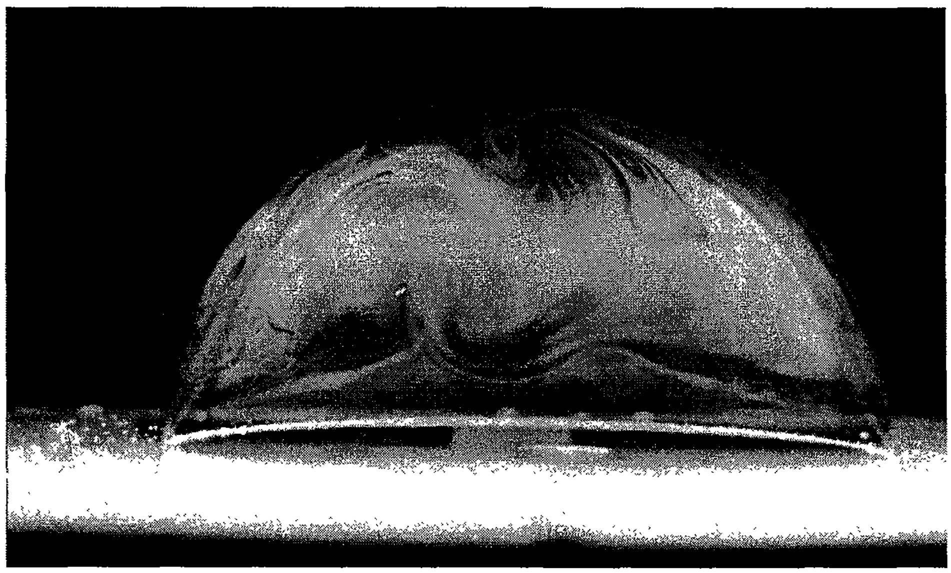


Figure 1

When the bubble is formed the thickness of the soap film is of the order of 10^{-5} *m* and no interference colors can be seen. Draining processes due to viscous flow and gravity convection rapidly brings excess liquid to the base of the bubble collecting at the so-called Gibbs ring. As a consequence the bubble becomes thinner and interference color bands appear. They reflect the gradual and steady increase in thickness of the bubble surface towards the base.

Light waves that are reflected simultaneously by the front and back surface of a soap film interfere with each other. This means that they either reinforce or cancel depending on their wavelengths relative to the thickness of the film. Because the different colors of white light illuminating the film correspond to different wavelengths, the colors are not all reflected in the same way. This is what gives the reflected light its color. Thus, if white light illuminates a film of thickness corresponding to the color yellow having a wavelength of 600 *nm*, this color is canceled by destructive interference, then only blue and red remain in the reflected light and the film appears purple. It was Newton who was the first man to understand these phenomena.

As the thinning process proceeds by evaporation, distortion of the concentric bands eventually takes place. The effect is enhanced by thermal heat from the illuminating lamp and more or less chaotic patterns may arise. Sometimes there will be a spectacular outburst of liquid streams from the Gibbs border due to pressure gradients, spreading over the soap film surface like a mushroom plume eventually dividing into two separate streams finally returning to the border. Sometimes swirls and spiral-formed structures can evolve.

Videotaping the bubble during its lifetime makes it possible to study more in detail the continuously changing thickness and topography of the soap film surface as reflected in the dynamics of the different colors caused by static and dynamic processes. Under favorable conditions, that is if the bubble survives long enough, the bubble can develop an ultrathin so-called "black film" starting at the top of the hemisphere, the thickness of the film being about 30 *nm* when formed. This occurs when the bubble survives long enough to drain and reach a thickness that is considerably less than one-quarter of the wavelength of the incident light, the reflected light from the outer and inner surfaces suffer from destructive interference. All colors are weakened to about the same extent and the film reflects little light, then it looks black against a dark background.

b) Experiment designed to measure lifetime of Soap Bubbles statistically

The lifetime of an individual soap bubble can easily and trivially be determined simply using a stopwatch. More informative is to use methods for a quantitative estimation of the lifetime of an ensemble of soap bubbles. The most obvious and easiest way to do this is to produce a large number of bubbles in order to get some statistics and then count the number of remaining bubbles after suitably selected time intervals. We face two problems: 1) the bubbles have to be of equal size, thickness, and 2) we want to produce all the bubbles at the same time.

A simple experiment turned out to satisfy these requirements. When touching the surface of a soap solution with an ordinary drinking straw, a thin soap film is produced across the open end of the tube. Next, the straw was turned upside down and dipped once more into a liquid that can be simply water. Obviously, in this way a bubble is produced at the upper end of the tube, the size of the bubble being dependent on the extent the straw is depressed into the liquid.

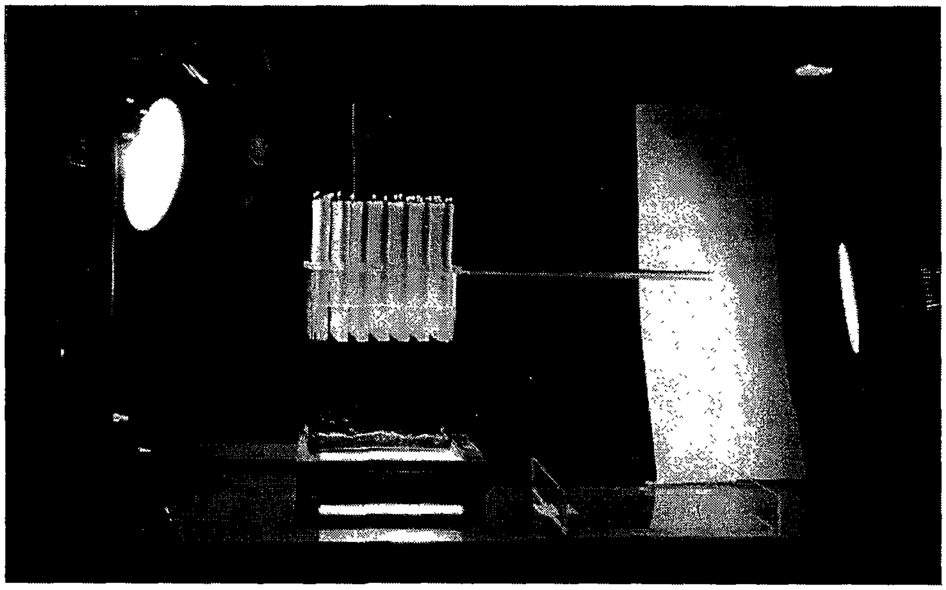


Figure 2

This experiment was the foundation for constructing a so-called "bubble-organ". In Figure 2 the experimental arrangement is depicted. This bubble plate holds 42 symmetrical placed plastic tubes of the same diameter and length depressed into the liquid to the same depth, producing instantaneously a corresponding number of bubbles, all of the same size. The length of the tubes was 20.0 *cm*, the diameter 15.0 *mm* and the spacing between the centers 30.0 *mm*. One must make sure that the distance between the tubes is such that a bursting bubble does not trigger the collapse of other bubbles in the vicinity. In fact, this is an effect that must be interesting to look at in more detail!

The number of remaining bubbles can be counted at selected time intervals and the half-life, $t_{1/2}$, determined. In order to facilitate the counting of bubbles it was found convenient to videotape the bubble field for later analysis. This method is convenient to use especially when the lifetime of the bubbles is rather short.

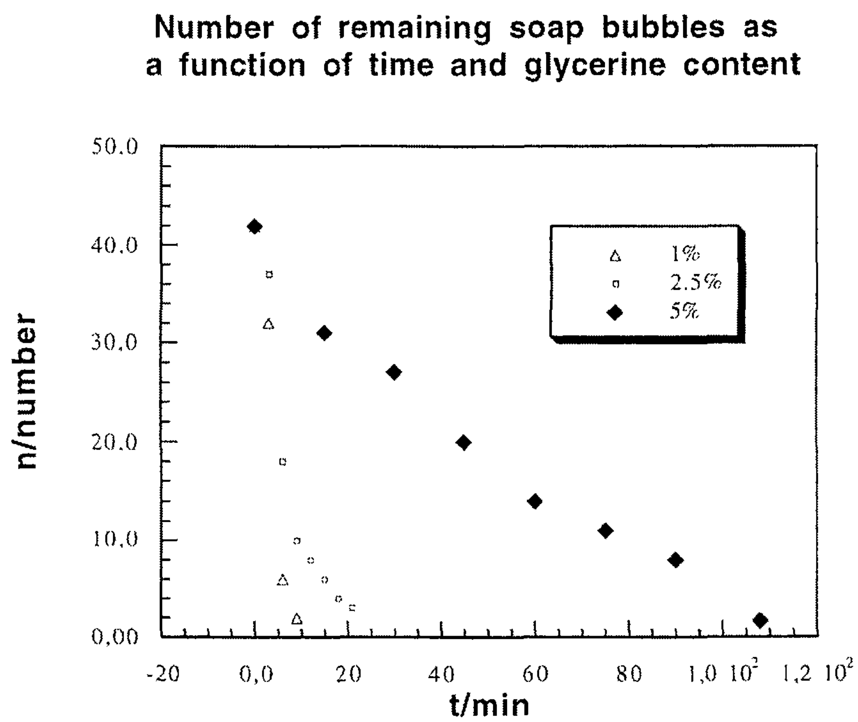


Figure 3

Typical results from experiments using the bubble plate with 42 tubes is shown in Figure 3. The average value of three identical experiments is presented. The number of remaining bubbles at selected points of time has been plotted as a function of time. Three different soap solutions were used containing 1.0, 2.5 and 5.0 % of glycerine, all of them having 1 % by weight of a commercial concentrated liquid dishwasher solution.

c) Experiments with bursting bubbles

Blowing a soap bubble and allowing it to settle on a piece of paper became the starting point for a series of interesting experiments with the aim to study the resulting pattern from bursting bubbles. A soap bubble on a piece of paper rests in the form of a half-sphere, a hemisphere. Of course it bursts sooner or later and the fragmented film gives a more or less symmetrical reminiscent pattern on the paper. In order to reveal the pattern powdered carbon was sprinkled over the paper surface. The carbon particles selectively stick to the soap patches on the paper and the excess carbon can be wiped off the surface by shaking the paper slightly. An aesthetically very attractive pattern appears.

A more systematic analysis of the distribution of these droplets was performed under different experimental conditions. For instance, bubbles derived from different soap solutions were deliberately punctured at specific spots. Moreover, these experiments were found to be more conveniently conducted by producing the bubble from a colored soap solution using the dye Congo red. In this way an excellent colored fingerprint of the reminiscent of the bursting bubble was obtained. Figures 4a and b give examples of droplet patterns that can be obtained from these experiments.

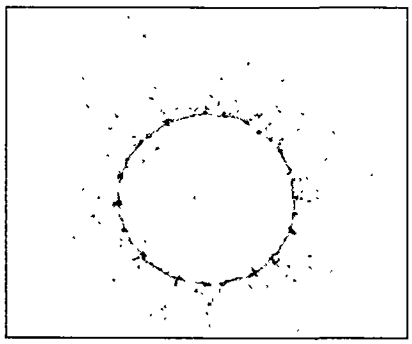


Figure 4a

Figure 4b

Figure 4a shows a bubble that has been punctured centrally at its highest point, giving a highly symmetrical droplet pattern on the paper. When a circular hole is generated in the film, it expands uniformly and radically due to action of the surface tension and the release of the excess pressure in the bubble defined by the equation $\Delta p = 4\gamma/R$, γ being surface tension and R radius of the bubble. The liquid is collected in the growing rim forming a cylinder that becomes unstable, collapses and creates droplets of different sizes which are thrown out. Depending on the composition of the soap solution as many as 1000 droplets can be produced. Figure 4b shows a bubble that has been punctured close to where the bubble is supported by the paper. In this case the droplets from the fragmented soap film have been thrown out within a very narrow angle on the opposite side of the point of rupture. A more detailed description of these experiments on exploding bubbles of various kinds is given in ref. [8].

REFERENCES

Boys, C. V. (1959) *Soap Bubbles and the forces which mould them*, Dover, New York.

Isenberg, C. (1978) *The Science of Soap Films and Soap Bubbles*, Tieto Ltd.

A S. C. (1929) *Soap Films, a study of molecular individuality*, Bell and Sons Ltd.

Lovett, D. (1994) *Demonstrating Science with Soap Films*, IOP Publishing Ltd.

Mysels, K J et al. (1959) *Soap Films*, Pergamon Press

Seifenblasen (1988) Forming Bubbles, in Frei, O., ed. *Information of the Institute for Lightweight Structures (IL)*, no. 18., University of Stuttgart.

Wolf, K. L. (1968) *Tropfen Blasen und Lamellen*, Springer-Verlag.

Rämme, G (1998) *Soap Bubbles in Art and Science*, Singapore: SCT Publishing.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 400-407, 1998

SOME ASPECTS OF SPACE STRUCTURES SHAPING

Janusz Rębielak

Architect (b. Bierutów, Poland, 1955)
Address: Division of Building Structures, Department of Architecture of the Technical University of Wrocław, Poland
Fields of interest: the morphology of space structures together with their applications as roof constructions of large spans and as structures for high-rise buildings.
Publications: Regular spherical grids obtained with the method of deformation of a secondary network, in: *Symmetry of Structure*, eds. : G Darvas, D. Nagy, August 13-19, 1989, Budapest, Hungary, pp. 467-470., The application examples of the method of the secondary grid deformation, *Symmetry: Culture and Science*, 3 (1992), 3, 293-303, Symmetry in spherical grid shaping. Pattern examples of some geodesic domes, in: *Symmetry of Patterns,* Second Interdisciplinary Symmetry Symposium and Exhibition, Hiroshima, Japan, August 17-24, 1992 , *Symmetry: Culture and Science*, 3 (1992), 1, 62-63., Elementy symetrii w kształtowaniu struktur przestrzennych, (Elements of symmetry in space structures shaping - in Polish), in: *Symetrie w przyrodzie, sztuce i naukach humanistycznych*, ed. A. Nobis, Wrocław: Leopoldinum, Studium Generale Universitatis Wratislaviensis, IV (1994), 119-133.; Bar space structures - Rules of shaping, in: *Symmetry: Natural and Artificial, Proceedings of the Third Interdisciplinary Symmetry Symposium and Exhibition*, Washington, USA, August 14-20, 1995, *Symmetry: Culture and Science,* 6 (1995), 3, 442-445.

The subject of the paper refers to the basic principles of the process of space structures shaping. In architecture and civil engineering the notion of *space structure* usually concerns the structures composed of straight bars which are appropriately connected together in articulated nodes.

At present the meaning of this notion is larger than it was in the past. It can be used to describe different kinds of engineering works and it may refer to many various types of structures being the nature creations. In both cases the tendency to apply the same simple rules in order to obtain the efficient, lightweight and strong structure is noticeable. The same basic principles, which are taken into consideration during the creation processes in both cases, have to give in the result the solutions of similar features. Some nature creations are seemingly of very complex shape, for instance the inner structure of bones, but after the comprehensive analysis it becomes clear that they are of very economic shape.

Many types of space structures, designed by an engineer, may consist of optional bars having various lengths and sizes of their cross sections. In that case it can be expected that the obtained structure will be of irregular shape and in its pattern it will be almost impossible to find any regularity. These forms of structures are not considered as good solutions also because of architectonic reasons. The accidental usage of any type of bar will give in the result a space structure, which will have small practice usability or which can not be considered as a building structure. The aim of the engineer's activity is to apply rules of exact sciences in order to obtain desirable product by means of minimum usage of energy, material and the human manual labour. The economic reasons force designers to use simple forms of structures having legible regularities, sometimes called also as the symmetry.

The endeavor to apply structural solutions of visible symmetry may be noticed in many important buildings of great dimensions designed in the historic periods, at present and proposed for the future. The high-rise buildings and roof covers of great span are very spectacular members of that group of architectonic objects.

INITIAL REMARKS

The notion *space structure* refers today usually to the modern type of construction which is characterized by the high degree of the inner ordered state. Then it seems that space structures are sometimes very distant to the nature creations in patterns of which one could very often observe the considerable quantity of accidental connections between their component parts. There could be given many examples of very close similarity between shapes of engineering creativity and forms founded in the nature (Rębielak, 1995).

The arrangement of basic component parts in the inner space of some creations of the nature appears very often like a kind of chaotic connections but after the careful analyses it is taken as a clear and economic solution in this particular case. On the other hand some very regular engineering structures may be considered as having very irregular patterns which depends on many factors, and among others, on the point of their observation. In both cases the component parts are displaced in the planned way in the whole structure which causes the obtained construction system to have an economic form and its pattern can be of many good esthetic features.

TYPICAL FORMS OF SPACE STRUCTURES

The very often used definition of modern type of space structures is as follows: it is a construction composed of straight bars arranged in a uniform way in its space. These bars are connected together in theoretically articulated nodes. The meaning of this definition is presently much larger than it was in the past.

Shell structures belong to the surface space structures. They are usually made of reinforce concrete that is why they are expensive and they were mainly designed and built in the 50s and 60s of the 20^{th} century. Because of economic reasons structures having mainly the tensile members are nowadays very often preferred in the practice.

Space structures are initially applied in civil engineering and in architecture in the second half of the 20^{th} century. They are lightweight and at the same time they are suitable great rigidity, they are flexible in the process of their shaping and they could obtain interesting architectonic views (Makowski, 1992). The spectacular achievements in the development of space structures were done by M. Mengeringhausen, M. Nowicki, R. Le Ricolai, R.B. Fuller, Z.S. Makowski, F. Otto, S. Du Chateau, W. Zalewski and M. Kawaguchi.

EXAMPLES OF STRUCTURAL SHAPES

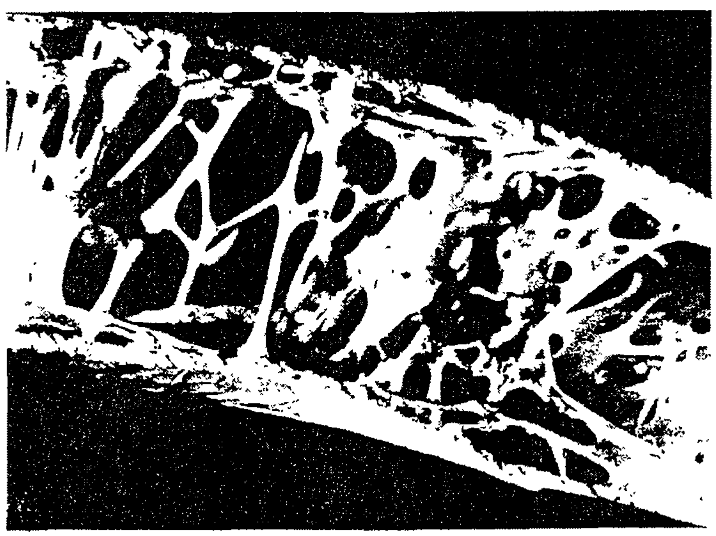


Figure 1

The short review of structural forms may be started from e.g., the nature creation in shape of the inner construction of bones. The main part of the human skull has the shell form which it is made possible to use the minimum amount of material in order to build very efficient type of construction which is able to take sometimes great loading applied from various directions. By the further investigation one could notice that in the cross section this structure has more complex and sophisticated shape. It consists of two surfaces of shell constructions and these two external shells are connected by means of seemingly chaotic system of little pieces of the osseous tissue. Some similar regularities can be particularly good observed in the example of the cross section of the humerus of the eagle (Otto, 1985), see Figure 1.

This sophisticated shape of natural support structure has to be lightweight and it can be able to take sometimes really great load. The structural form is the result of many experiences made in a very long way of evolution of the species. The human works made by engineers can obtain patterns similar to the presented above. Figure 2 presents results of the computer aided topological optimization of a simple beam-truss (Reichhart, 1997). They can be considered as a certain kind of the outcomes of the conceptual design of structures. Space structures obtained in this way are efficient construction solutions and they have interesting forms of individual esthetic features.

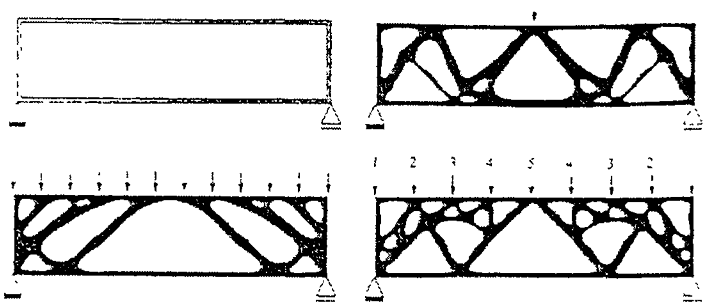


Figure 2

Very effective forms of structures, applied in architecture and in civil engineering, were obtained by means of special procedures of the model testing of chosen materials having the initial simple shapes. For example the forming process which uses the moistened threads with "limited excess length" may be applied in designing of many types of support structures. In this case one should notice the compromise between the rules of the so-called direction path system and the minimal path system (Kołodziejczyk, 1997). Certain types of structural systems obtained in this way have patterns very similar to patterns of systems of tree branches.

SOME PROPOSALS OF FORMS OF SPACE STRUCTURES

This chapter presents examples of systems of space structures. These examples are representative for many groups of structural systems proposed by the author for large span roofs and for high-rise buildings (Rębielak: 1992, 1993, 1996, 1997, 1998).

Various types of space structures are used as construction systems in designing of many forms of large span roofs. The economic clear span of a flat shape of a space structure is estimated as about $20h$, where h means their construction depth defined as the distance between external layers of bars. The interdependence causes that structure designed as a construction of very large span cover has to be made of bars of considerable great lengths. The load carrying capacity of the whole structure is in great part determined by the load carrying capacity of a few members subjected to acting the greatest compress forces. Their load capacity depends, in an inversely proportional way, on the square of their reduced buckling lengths. Therefore in areas of acting of these forces should be applied the shortest bars of a space structure.

Figure 3 shows the bar arrangement in the space of one of the structures designed specially for large span roofs (Rębielak: 1996, 1997). This kind of structure is called B3{T – T}A. It was shaped by means of reduced parts of "big" tetrahedral modules separated by means of additional octahedral bar sets.

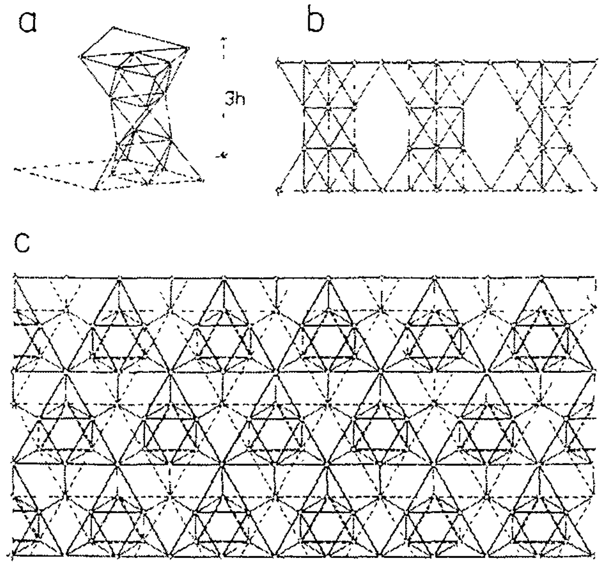


Figure 3

This space structure has considerably great construction depth and it is composed of relatively short bars of the same length. It is of somewhat complex form and that is why one could notice some similarities between its sample and the pattern of the nature creation shown in Figure 1. Because of great number of bars and nodes falling on the unit of the covered surface it can be expected that this form of space structure will have only the potential possibility to use in the real design of large span roofs.

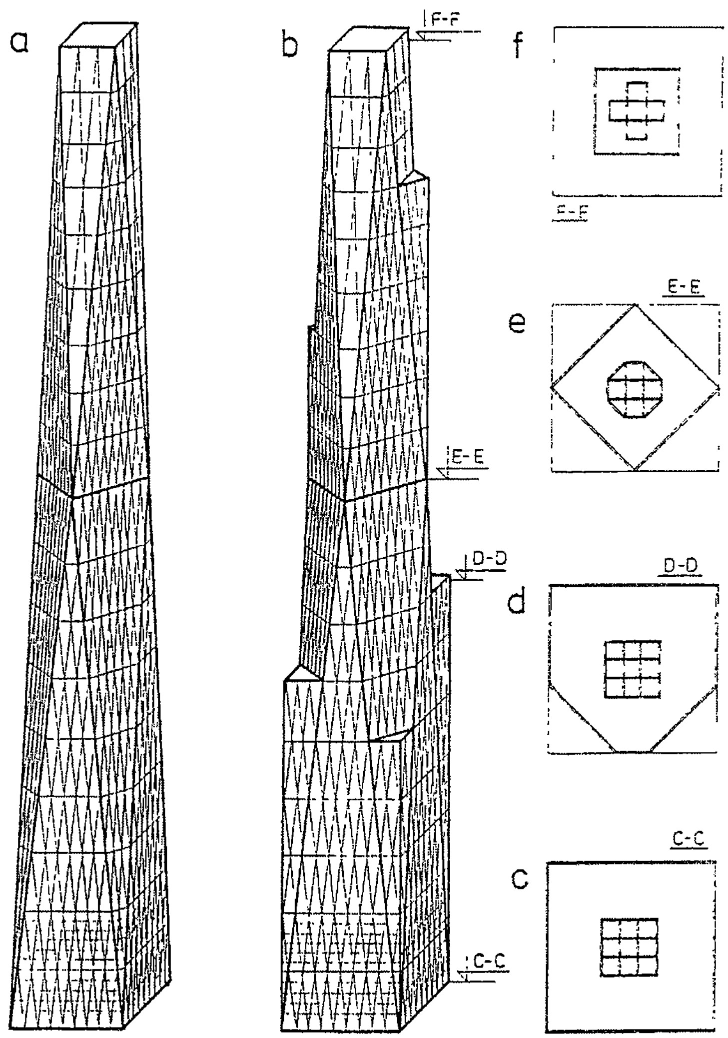


Figure 4

The basic rules of space structures shaping are also used in designing of structural system for high-rise buildings. The highest buildings are presently of the height equals about 450 meters. They are usually built in steel or reinforced concrete structural systems called as follows: exterior frame tube, tube in tube or bundled tube. The construction system should ensure a suitable stiffness of the tall building under acting of forces caused by many types of loading. Among a few systems there should be chosen that one which does not require considerable enlarging of dimensions of component parts above those which are necessary to carry the vertical load.

Figure 4 presents schemes of one of the structural systems proposed for designing of high-rise buildings (Rębielak, 1998). The structural system is proposed to call framed polyhedron. The basic component parts of such a building have forms of elongated halves of cubooctahedrons which are vertically located each on other, see Figure 4a.

The upper component part is suitably smaller than the lower one. The main support system consists of skew columns and horizontal beams creating the triangular grid onto each face of the chosen forms of big polyhedrons. The degree of subdivision of appropriate faces may be different.

A more complex form of a tall building is presented in Figure 4b. It was formed by means of arrangement of additional parts, in form of suitable tetrahedrons, in chosen corners of each component part. It can be expected that this shape of structural system could be a very efficient solution for designing of support structures for the highest buildings. Owing application of the proposed system these buildings could obtain individual and interesting architectonic forms.

CLOSING REMARKS

Certain accidental connections between component parts inside spaces of some nature creations can be considered as a kind of disorder but by the further investigation they become as intended way of their economic building. In the engineering design accidental joints between component parts are not accepted. It is the basic formula which requires the order in every planned economic activity of human beings.

REFERENCES

Kołodziejczyk, M. (1997) Certain aspects of the use of the thread model in the search for self-forming processes, Proceedings of Local Seminar of IASS Polish Chapter, *Lightweight Structures in Civil Engineering*, pp. 9-16.

Makowski, Z. S. (1992) Space frames and trusses, In· *Constructional Steel Design An International Guide*, London & New York, pp. 971-843.

Otto, F. (1985) Architecture et bionique, *Constructions naturelles*, Editions Delta & SPES.

Reichhart, A. (1997) O koncepcyjnym projektowaniu konstrukcji, *Inżynieria i Budownictwo*, 9, 439-444.

Rębielak, J. (1992) Space structures – proposals for shaping, *International Journal of Space Structures*, 3, 293-303.

Rębielak, J. (1993) Space structures used in the construction of large span roofs and tall buildings, in. *Space Structures* 4, Proceedings of the Fourth International Conference on Space Structures, eds. G.A.R. Parke, C.M. Howard, Guildford, UK, 5-10 September 1993, Vol. 2, London: Thomas Telford, pp. 1581-1590.

Rębielak, J. (1995) Bar space structures - Rules of shaping, in *Symmetry: Natural and Artificial*, Proceedings of the Third Interdisciplinary Symmetry Symposium and Exhibition, Washington, USA, August 14-20, also in: *Symmetry: Culture and Science*, 6, 3, 442-445.

Rębielak, J (1996) Special group of space structures shaped on the basis of the B{T – T}A space structure, Proceedings of Local Seminar of IASS Polish Chapter, *Lightweight Structures in Civil Engineering*, pp. 146-147.

Rębielak, J (1998) Structural proposals for long span roofs and high-rise buildings, *Raport serii* PRE Nr. I-1/P-822/98, Department of Architecture, Wrocław University of Technology.

Symmetry: Culture and Science
Vol. 9, Nos.2-4, 408-414, 1998

QUANTITATIVE ANALYSIS OF THE SYMMETRY OF ARTIFACTS

Idit Saragusti[(a)*], Ilan Sharon[(a)],
Omer Katzenelson[(b)] and David Avnir[(b)]

Address. Institute of Archaeology[(a)] and Institute of Chemistry[(b)] The Hebrew University of Jerusalem, Israel
* (to whom correspondence should be addressed) Institute of Archaeology, Hebrew University, Mt. Scopus, Jerusalem, 91905, Israel. E-mail· mssidit@pluto.mscc huji ac.il

INTRODUCTION

The concept of symmetry has attracted virtually all domains of intellectual activity, archaeology being no exception. For example, symmetry has been at the focus of many Lower Paleolithic studies because it is one of the main shape characteristics of handaxes from Acheulian assemblages. Scholars investigating this period have often argued that the ability to perceive symmetric shapes, is meaningful in attempts to reconstruct the evolution of human cognitive, behavioral and technological capacities. For instance, Isaac (1986) saw the ability to conceive and produce symmetric shapes, as manifested in many Acheulian handaxes, as one of the indications of the higher complexity of this industry complex, compared to earlier traditions. Wynn (1985) even proposed a qualitative correlation between the *degree of symmetry* of the artifacts and the degree of intelligence of their makers. He used the ability to perceive bilateral symmetry, as manifested in typical Late Acheulian handaxes, as a major argument for relating a higher degree of intelligence to the makers of these tools compared to the degree of intelligence demonstrated by the earlier Oldowan industrial tradition.

Despite the central importance of the concept of symmetry in these and many other studies, it has usually been treated in qualitative rather than quantitative way, and sentences like '...these early handaxes are *roughly* symmetrical – none approaches the *fine* symmetry of later example...' (Wynn, 1985, p. 40 our emphasis) are often found in the literature. The need for objective means for symmetry measurement is the main motivation for the current study.

The concept and the methodology of symmetry content measurement, were developed in a recent series of publications (Zabrodsky & Avnir, 1995 and references cited therein; Avnir, et al., 1996 – a review) and applied successfully to a variety of symmetry related problems on the molecular and supra molecular levels (Buch, et al., 1995; Kanis, et al., 1995; Katzenelson, et al., 1996; Keinan, et al., 1996; Pinto, et al., 1996). This approach of Continuous Symmetry Measures (CSM), is general and not limited to the molecular scale.

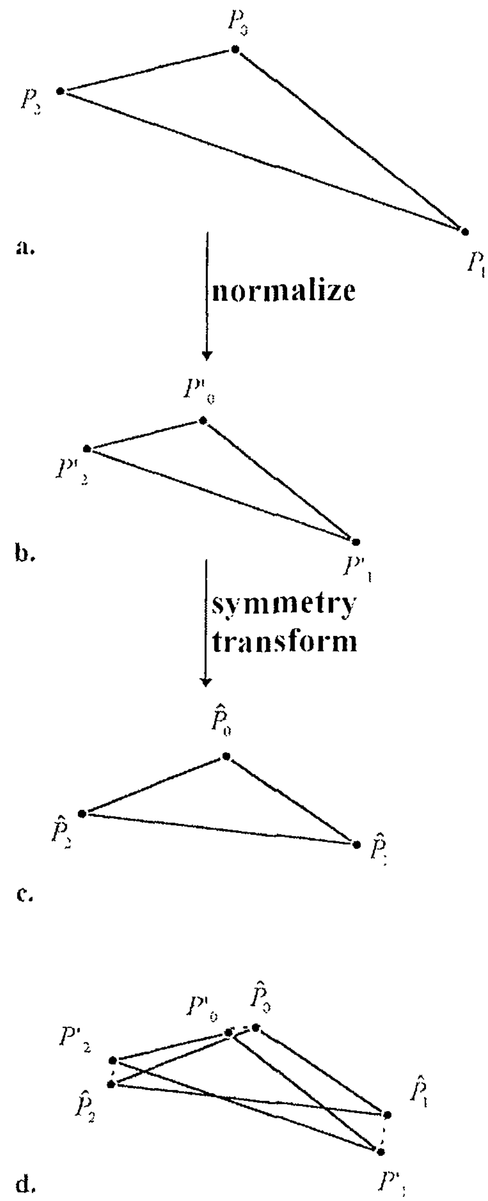
The main aim of this study was to test whether quantitative analysis of symmetry is applicable to archaeological research as an objective tool for assessment of shape through symmetry.

The general principles of the method are briefly described below, with special attention given to the details relevant for the archaeological research. The application of this method to the quantitative analysis of the symmetry of handaxes is described thereafter.

THE CONTINUOUS SYMMETRY MEASURE APPROACH

The Continuous Symmetry Measure (CSM) approach is based on the notion that quite often, it is natural to evaluate "how much" of a given symmetry there is in a structure rather than treating it as a "more" or "less" property. Consequently, this implies treating symmetry as a structural property of a *continuous* behavior, the quantity of which should be evaluated by a measurement tool. The design of a measurement tool involves some degree of arbitrariness, in the sense that one has to decide on issues such as how should the zero-reference level be set, what should be the maximal value, what should be the actual measurement yard-stick, what normalization procedures should one employ, and so on. Any such decision is open to debate. Keeping this in mind, the symmetry measure is based on a definition which is minimalistic as one could practically get. Our proposed answer (Zabrodsky & Avnir, 1995) to the question "How much of a given symmetry is there in a given structure?" has been:

Find the minimal distances that the vertices of a shape have to undergo, in order for the shape to attain the desired symmetry.



In a formal way:

Let G be a given symmetry group.

Let $P_{i=1\ldots n}$ be the vertices of the original configuration and $\hat{P}_{i=1\ldots n}$ the corresponding points in the nearest G-symmetric configuration then

$$S(G) = \frac{100}{n}\sum_{i=1}^{n}\left\|P_i - \hat{P}_i\right\|^2 \qquad (1)$$

expresses the amount of symmetry in the original configuration.

If a shape has the desired symmetry, $S(G) = 0$. A shape's symmetry measure increases as the shape departs from G-symmetry and it reaches a maximal value. As seen above, eq. (1) is normalized to the number of vertices (n) defining the object's outline.

Eq. (1) is general and allows one to evaluate the symmetry measure of *any* shape relative to *any* symmetry group or element. In order to avoid size effects, the size of the original structure is normalized, enabling the comparison of the degree of symmetry of shapes of different size.

Figure 1: The basic features of the Continuous Symmetry Measure (CSM): In order to evaluate how much bilaterality (mirror symmetry) is there in the triangle (a), its size is normalized (b), and the mirror symmetric structure (c) which is nearest to (b) is found using the symmetry transform described in the text and in the Appendix. (d): The $S(\sigma)$ value is calculated from the minimal distance between (b) and (c), using equation (1). In this case, $S(\sigma) = 4.96$

An important feature of the CSM approach, is that no pre-selected specific reference shape is assumed at the beginning of the analysis, though it is obtained as an end outcome. The nearest set of $\hat{P}_i$'s, i.e., the set of coordinates describing the nearest symmetric shape is computed with an algorithm, described in detail in Zabrodsky and Avnir (1995). The algorithm thus also provides one with the shape of the nearest symmetric object (Fig. 1). An additional important feature is that the $S(G)$ values are on the same scale, so that one can ask, for example which member of a set of polygons is the closest to being symmetric (having the lowest $S(G)$ value), or the farthest (having the highest value).

BILATERAL SYMMETRY ANALYSIS OF HANDAXES

As a part of a research which aims to analyze changes in the symmetry of handaxes over time, samples of handaxes from three Lower Paleolithic sites in Israel (Ubeidiya, Gesher Benot Ya'aqov and Mayan Barukh) were used as a test-case, through which we were trying to test the applicability of the method presented above to the archaeological research.

In the archaeological literature the term "symmetry" has been usually limited to bilateral symmetry, namely to symmetry around a mirror that bisects the body, also known as *achirality.* Many other symmetry operations (e.g., symmetry of rotation, tetrahedral symmetry and so on, which may be of relevance for other archaeological problems as well) are treatable by eq. (1), but we limit our discussion here to bilateral symmetry, to which we continue to refer in this report as simply "symmetry". The meaning of eq. (1) is then, the minimal distance of a given shape from exact bilateral symmetry.

Also in the current study we limited ourselves to the analysis of the two dimensional (2D) symmetry of the plan view of the artifacts. In the archaeological research, the 2D presentations of plan view and various cross sections of the 3D handaxes are standardly considered to represent reliably the morphological features of the artifact, including its symmetry features. The plan view is the one most commonly used for shape evaluation of handaxes, although the degree of symmetry of longitude and latitude cross sections maybe of interest as well.

CSM analysis requires that the shape of the object to be analyzed be represented by a set of n vertices or boundary points. For the purpose of this pilot study, we scanned hand-drawings of the handaxes outlines into a PC and then automatically traced these outlines, resulting in sets of x, y coordinates, each set of which served as the desired boundary points.

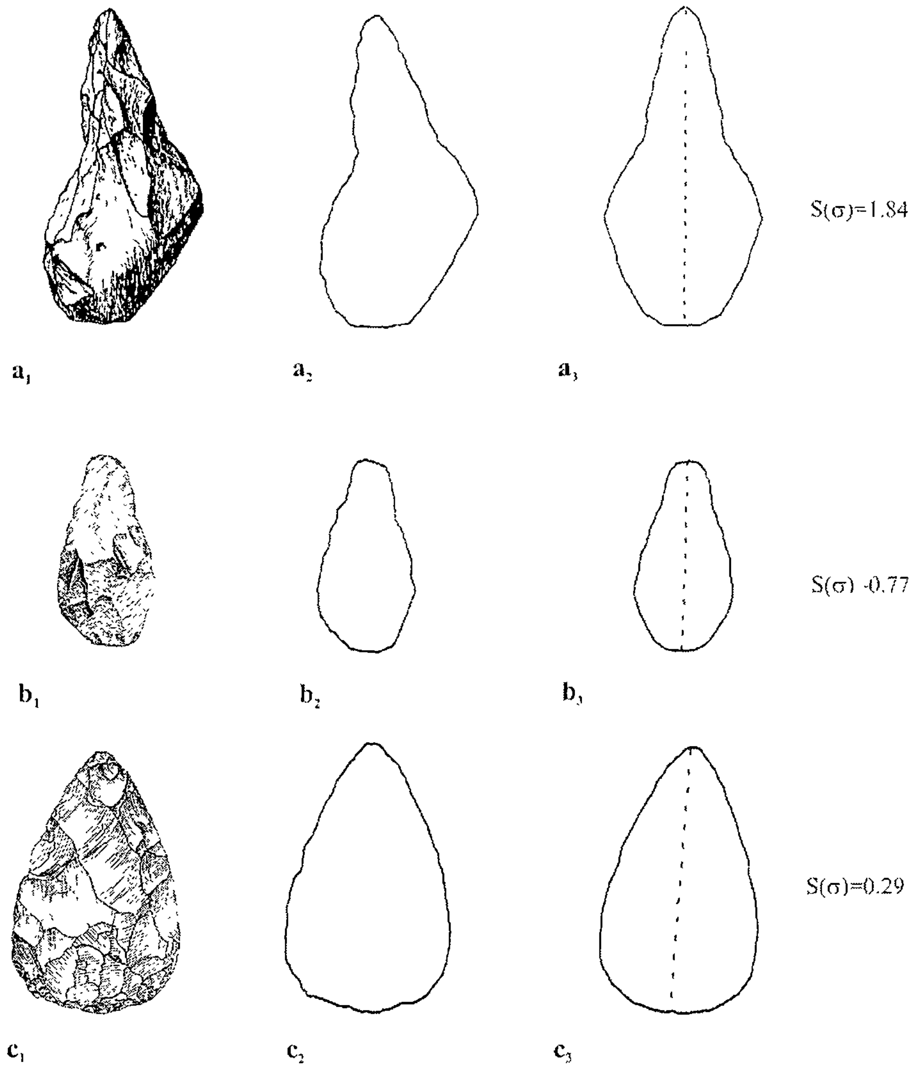


Figure 2: Handaxes samples from the three sites: a - Ubeidiya (Bar-Yosef & Goren-Inbar, 1993); b - Gesher Benot Ya'aqov and c - Ma'ayan Barukh (Stekelis & Gilead, 1966); Original drawings: a_1, b_1, & c_1; Outlines redrawn, using x, y coordinates: a_2, b_2 & c_2. The nearest symmetric shapes, their $S(\sigma)$ values and the mirror lines: a_3, b_3 & c_3.

Fig. 2 shows that redrawing of the outlines with these coordinates preserves the information and the redrawn outlines are very close to the original ones. The symmetry values of each of the outlines was computed according to eq. (1), using the sets of coordinates. The output is the minimal $S(\sigma)$ value, namely, the minimal distance between the original outline and the *nearest* mirror-symmetric shape. An additional output is, as explained above, the coordinates of the symmetrized shape and the position and inclination of the mirror line, σ (Fig. 2).

Preliminary results of this research have shown that the overall symmetry of handaxes generally increases over time, and that the variability decreases. Although this tendency is commonly mentioned in the archaeological literature, we provide here – to the best of our knowledge for the first time – quantitative demonstration of this phenomena. These preliminary results and a more extensive discussion of this issue can be found elsewhere (Saragusti, et al. 1998).

CONCLUSIONS

Our starting point was the prevailing situation where symmetry often serves as an approximate descriptive 'language' in archaeological studies. We proposed that while it is true that an imprecise language helps grasp complex situations and identify first order trends, there is always a danger of missing the full picture because of vagueness. Thus, it has been our aim here to provide a method for a quantitative evaluation of symmetry in archaeology. This method, when used to evaluate the degree of symmetry of handaxes, was found to be sensitive enough for measuring observed differences in symmetry. The CSM method may, therefore, be applied as a means to achieve a quantitative, objective and more accurate description of an important global shape feature of artifacts in archaeological research.

REFERENCES

Avnir, D., Katzenelson, O., Keinan, S., Pinsky, M., Pinto, Y , Salomon, Y. & Zabrodsky Hel-Or, H. (1996). The Measurement of Symmetry and Chirality. In: D. H. Rouvray, Ed., *Concepts in Chemistry* Somerset: Research Studies Press, pp. 283-324.

Bar-Yosef, O. & Goren-Inbar, N. (1993). *The Lithic Assemblages of 'Ubeidiya, A Lower Palaeolithic Site in the Jordan Valley.* Jerusalem: The Hebrew University of Jerusalem.

Buch, V , Greshgoren, E., Zabrodsky Hel-Or, H. & Avnir, D. (1995). Symmetry Loss as a Criterion for Cluster Melting, with Application to (D2) 13. *Chemical Physics Letters* 247, 149-153.

Isaac, G. L. (1986) Foundation Stone: Early Artefacts as Indicators of Activities and Abilities In Bailey, G. N. and Callow, P. Eds. *Stone Age Prehistory.* Cambridge: Cambridge University Press, pp. 221-241.

Kanis, D. R., Wong, J. S , Marks, T. J., Ratner, M. A., Zabrodsky, H , Keinan, S. & Avnir, D. (1995). Continuous Symmetry Analysis of Hyperpolarizabilities. Characterization of Second Order Non-Linear Optical Response of Distirted Benzene. *Journal of Physical Chemistry* 99, 11061-11066.

Katzenelson, O., Zabrodsky Hel-Or, H. & Avnir, D. (1996). Chirality of Random Supramolecular Structures. *Chemistry European Journal* 2. No 2, 174-181.

Keinan, S., Zabrodsky Hel-Or, H. & Avnir, D. (1996). How Much Chiralıty was there in Pasteur's Tartrate? The Quantitative Evaluatıon of the Chirality of Crystals. *Enantiomer* 1, 351-357.

Pinto, Y., Zabrodsky Hel-Or, H. & Avnır, D J. (1996). Continuous Chirality Analysis of Interconversion Pathways of the Water-Trimer Enantiomers. *Journal of the Chemical Society, Faraday Transaction* 92, 2523 - 2527.

Saragusti, I., Sharon, I., Katzenelson, O. & Avnir, D. (1998). Quantitative Analysis of the Symmetry of Artifacts: Lower Paleolithic Handaxes. *Journal of Archaeological Science* 25, No. 8, 817-825.

Stekelis, S. & Gilead, D. (1996). *Ma'ayan Barukh - a Lower Paleolithıc Site ın Upper Galilee.* Jerusalem: Israel Prehistoric Society.

Wynn, T. (1985). Piaget, Stone Tools and the Evolution of Human Intelligence. *World Archaeology.* 17, No. 1, 32-43.

Zabrodsky, H. & Avnir, D. (1995). Continuous Symmetry Measures. *Journal of American Chemical Society* 117, 462-473.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 415-420, 1998

COMMON MECHANISMS OF DIFFERENT FUNCTIONAL LEVELS OF SELF-ORGANIZATION

A. M. Segal[1] and E. G. Rapis[2]

[1] psychophysiologist, (b Chemovitz, Ukraınıa,1944).

Address: Biofeedback Unit, Psychıatric Servıces of Sourasky Medical Center, 6 Weizman St., Tel Aviv 64928, Israel and Sakler Medical School of Tel Aviv University.

Fields of interest: neurofeedback, self-regulation of physiological functions, applicatıons of theory of chaos in biology.

Publications: Effects of naloxone on the neuropsychiatric symptoms of a woman wıth partial adrenal 21-hydroxylase deficıency (with I. Blum et al.), *Am. J. Psychıatry* (1983) 140, (8); Sudden death, neuroleptıcs and psychotic agitation (with J. Leria and A Elitzur), *Prog. Neuro-Psychopharmacol. & Biol. Research* (1995), 19, 229 -241; Stress management for psoriatic patients. the effectiveness of biofeedback and relaxation techniques (with G. Keınan, U. Gal and S. Brener), *Stress Medicine* (1995), 11, 235-241.

[2] ophthalmologist, biophysicist (b Tomsk, Russia, 1927).

Address: 210, Hayarkon st , Tel Aviv, Israel.

Fields of interest: ophthalmolymphology , haemophthalm, self-organization of protein films.

Publication: The self-organization of protein, *Tech. Phys. Lett.* (1995) 21, 321-324; Symmetry and self-organization of protein, *Symmetry: Culture and Scıence* (1995), 439-441; Magnetic sensitivity of protein, *Tech. Phys. Lett.* (1997) 23, 263-267.

Abstract: *Possible common mechanisms of self-organization of bioelectrical activity in visual cortex and in simple in vitro protein-water systems (under non-equilibrium conditions) are discussed in the light of experimental data obtained.*

In our collaborative study a question of common mechanisms of self-organization of extremely different physiological systems (in terms of hierarchical levels of function) was investigated experimentally. Bioelectrical signals of visual cortex of learning brain (occipital lead of EEG in the process of thermal biofeedback) were compared with protein film formation in a protein-water system under certain in vitro non-equilibrium conditions.

Thermal biofeedback for many years is known to be a well-established therapeutic procedure (Schwarts 1995) but relatively little is known yet in terms of central mechanisms *of its action. In this study we tried to learn some aspects of this question in terms of central nervous system activity along the process of therapeutic skills acquiring.*

This subject might be viewed as well as a deviation in chaotic quasisymmetric order in electrical brain wave activity and its dynamic characteristics may be compared with those achieved by altered conditions in self-organizing systems of different hierarchical levels.

METHODS

Five subjects (four women and one man, ranging between 20 and 48 years of age) were involved in this study. All of them were suffering from a kind of a disorder in which thermal biofeedback was indicated as a therapeutic procedure. Three had vascular headaches, one was suffering from anxiety and one young woman had unstable essential hypertension. We decided to conduct this pilot investigation on clinical population in order to have a good back up in terms of patients' motivation. All the patients after a clinical intake procedure were given a session *of* combined Jakobson - Schultz relaxation and were instructed to proceed with this 15-minute relaxation three times a day. Then thermal biofeedback half an hour sessions were conducted on a twice a week basis. Two of the participants underwent 20 sessions each. Three others had received altogether 23 sessions. The number of all the sessions which had been conducted and analyzed was 63.

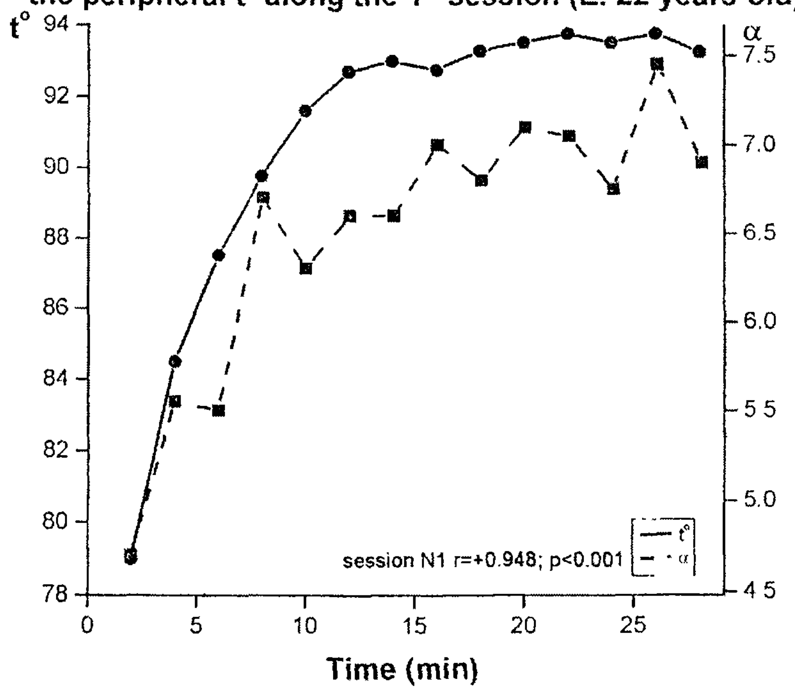
For physiological data acquisition The API Neurodata Physiograph was used. It has the possibility to monitor, store and process the thermal and the EEG information concomitantly. The EEG recordings were obtained from dominant occipital area. They were performed unipolarly and the patients were sitting with open eyes. The thermal biofeedback training was conducted classically by attaching the dominant index to a thermistor. The data was processed and presented in statistical charts. Correlational analysis of various variables was applied in order to determine possible statistical bonds between different phenomena.

RESULTS

As we knew already from our previous clinical work in biofeedback the majority of our patients are able to elevate the peripheral temperature of their hands or in other words, they are able to diminish the sympathetic flow. Therefore they relax the smooth muscles of the arterial vascular bed in the extremity attached to the thermistor. The overall change across all the 63 sessions was about 3,7 degree F. The main gain was obtained at the very beginning of the learning process (the very first couple of sessions). The time span from the start of the session up to the maximal elevation was shortened gradually from session to session from 15 to 7-8 minutes and remained so after session 11-12. The base line temperature of the dominant hand was climbing slowly up. The alpha-wave energy changes during the process of thermal training were even more demonstrative. At the first session across all 5 cases studied the gain during the session exceeded 100% (from 4,3 to 9,0 units of Neurodata scale). At the point of the 5th session this difference did not exceed even 1,5 units and this number was kept more or less on the same level up to the end of the procedure. No significant changes were observed in other aspects of brain-wave activity.

The direction of the change in temperature and in alpha-wave energy along the starting sessions was similar and the correlation between those variables appeared to be very high.

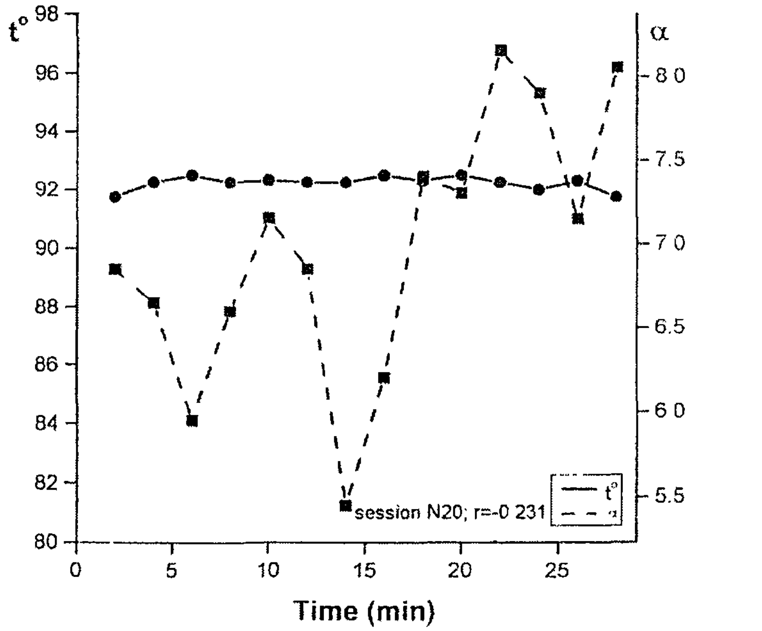
The dynamics of change of α-band and the peripheral t° along the 1st session (E. 22 years old)



x-time in minutes
y l-temperature
*y*2-*α* band energy

Those strong correlations had undergone very significant changes and as a paradigm the last 20th session of the same patient witnesses this fact: there is no correlation at all between those 2 parameters.

The dynamic of change of α-band and the peripheral t° along the last 20st session (E. 22 years old)



x-time in minutes
y l-temperature
y2-a band energy

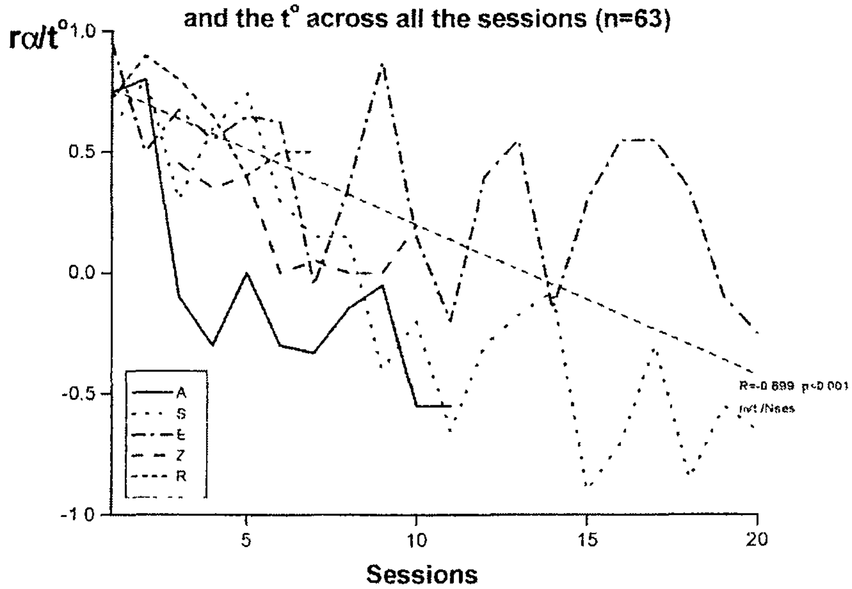
In general, the alpha-wave energy changes preceded the changes in temperature.

The second part of the study was conducted in an open protein-water system were non-equilibrium conditions were created in vivo by rapid water evaporation. In a series of experiments (more then 150) different systems of protein-water were placed on a solid substrate (glass, plastic etc.) and the process of protein-condensing was visualized in dynamics by means of several kinds of microscopy (for details see Rapis 1995a, 1995b, 1997).

The qualitative results of the experiments were repeated with high accuracy. The experiments showed the following dynamics of protein condensing: front movement of self-supporting oscillations (autowaves) has been constantly observed with alterations of attraction and relaxation zones. In liquid crystal phase at critical increase in density the appearing structural defects created spontaneous intermittent spiral and chiral symmetry and self-assembly nucleation of three-dimensional protein cluster film. Morphological structures of "cel1s" with nuclei or domains have been formed in them. In denser material avalanche-like spatial sharp-ended dissipative structures or folding films-solutions with fractal properties appeared suddenly. Uniting around the nucleus, they formed pairs of whirl tunnels. In liquid crystal phase jump-like correlated color changes (red, yellow and green) have been observed. Those changes were reflecting pulsating in film elasticity or reverse elasticity inherent to biopolimer gels (Aggeli 1997; Pouline 1997). On the whole we succeeded to find out a non-linear, chaotic condensation process dynamics and 3D protein film self-organization. This spontaneous dynamic process had a lot of properties inherent to biological living objects: morphological self-similarity, spontaneity, fractality, coherency, synchrony, intermittent spiral and chiral symmetry etc. The picture of such self-organizing protein-water films generally might be characterized as "chaotic symmetry".

DISCUSSION

Correlational relationships between the α-rythm energy and the t° across all the sessions (n=63)



x-number of sessions

y-correlation coefficient between α band and penpheral temperature

As can be concluded from our material the peripheral temperature changes along the thermal biofeedback training are accompanied from the very beginning by a concomitant and significant elevation in alpha-wave energy in the visual cortex of the brain. Those strongly bound statistically relationships relatively early underwent dramatic changes: the correlational index drops promptly reflecting, most probably, a different (comparing with the very beginning of the temperature learning) functional state of the occipital cortex.

We feel that at the initiation of the learning process the brain has to work deterministically but very quickly its activity returns to its previous patterns (attractors) in spite of the fact that the learning process is still going on. The statistical bonds between alpha energy and the peripheral physiological signal disappear. The whole-body homeostatic obligations are obviously far more important than serious devotion to a local peripheral task. Low dimensional goal-oriented activity is promptly replaced by a highly dimensional and functionally flexible state of high performance (Heffeman 1995). The chaotic and quasisymmetric equilibrium returns to dominate in the picture of EEG.

From the other hand, the self-organizing process in non-equilibrium systems "protein-water" suggests strongly that it follows the same rules of the theories of fractals and chaos (Rapis 1995a, 1995b, 1997). Because the fact that protein films and membranes are one of the milestones of almost any living organism we can suppose that its rules of self-organization determine the fundamental laws of self-organization in higher functional levels of hierarchy.

Such a postulate seems to be obvious because if a lower level and a higher one would not obey the same rules – the system would not be able to function properly. This study gives us some experimental evidences supporting this idea.

REFERENCES

Aggeli et al. (1997) Responsive gels fonned by the spontaneous self-assembly of peptides into polymeric betha-sheet tapes, *Nature*, 386, 20, 259-262.

Heffernan, M. (1995) Using chaos to control brainwaves, *Megabrain Report*, 3 (1), 14-21.

Poulin, P et al (1997) Novel colloidal ınteractions in anisotropic fluids, *Scıence*, 275, 1770-1773.

Rapis, E. (1995a) The self-organization of proteın, *Tech. Phys. Lett.*, 21, 321-324.

Rapis, E. (1995b) Symmetry and self-organization of proteın, *Symmetry: Culture and Scıence*, 439-441.

Rapis, E. (1997) Magnetic sensitivity of protein, *Tech. Phys. Lett.*, 23, 263-267.

Schwartz, M (1995) *Biofeedback Practıtıoner Guıde*, Guilford Press, Second Edition

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 421-430, 1998

A SYMMETRIC PATTERN IN FINANCIAL MARKETS

George G. Szpiro

Address: The Israeli Centre for Acadentic Studies, Kiriat Ono, Israel (affiliated with the University of Manchester), P O.Box 6278, Jerusalem, Israel. Fax· ++972-2-6230201. E-mail: george@netvision.net.il.

Abstract: *A subtle pattern is hidden in the markets for financial assets which is due to the finite size of the "ticks" by which prices can change. The nearly symmetric pattern arises when the values of a time-series are plotted against their delayed values. This paper clarifies the reason for this phenomenon. High frequency data from a foreign exchange market are used to illustrate this phenomenon.*

1. INTRODUCTION

For generations so-called chartists and technical analysts tried to find structure in financial and economic data but, as the generally accepted "perfect market hypothesis" suggests, most of these endeavors proved futile. However, the study of chaotic phenomena in the physical sciences suggested new avenues in this research area. For example, Brock, Dechert, LeBaron and Scheinkman (1996) introduce a test for the whiteness of time-series observations (i.e., the absence of structure) based on the correlation integral, a tool originally introduced by Grassberger and Procaccia (1983) to determine the dimension of a chaotic system. Attempts to forecast time-series have been made using a technique that is related to the Grassberger-Procaccia method and have been successful in the physical sciences (Farmer and Sidorowich 1987). They failed in economics and finance, however, and Szpiro (1997) suggested using the method to measure the amount of noise that is present in economic data.

The so-called phase portrait, a well-known instrument in the theory of dynamical systems, is a tool that has been utilized to detect non-linearities in time-series.

This method requires embedding the data in two- or higher-dimensional space – in the most elementary version this simply means plotting x_{t+1} against x_t – and inspecting the resulting graph. It was generally thought that a non-uniform distribution of the plotted points indicates the presence of some underlying structure. The inference was that in such a case the data are, in principle, forecastable. However, rounding errors in the measurement process or the non-continuity of price quotations can significantly affect the above conclusion. Crack and Ledoit (1996, henceforth CL) described the effect that emerges when the returns R_t of a financial time-series are embedded in two-dimensional space. By plotting the return of a share in one period, against its return during the previous period, CL discovered that rays, originating at the origin, radiate in all directions. Major directions are more pronounced than minor directions, and CL named the figure a "compass rose". No predictive power can be drawn from the existence of this phenomenon, however, and, even more seriously, Kramer and Runge (1997) showed that the existence of the compass rose seriously distorts tests that are meant to detect the presence of chaos in time-series.[1]

In view of recent developments, for example NYSE's announcement to move the minimum increment in the price of traded stock (the tick size) from eighths of a dollar to sixteenths and then to tenths (NYSE 1997), or the emerging tendency to charge minute fractions of cents for services on the Internet, it is important to analyze the effects of rounding errors in a rigorous fashion. This paper – together with its predecessor (Szpiro 1998) – puts the analysis of non-continuities in market prices on a sound basis. In the following section the mathematics of a formal model is set out, and Section 3 uses data from a foreign exchange market to show that the symmetric patterns arise in actual economic time-series.

2. THE MATHEMATICAL MODEL

The return and the closing price of financial asset, say a stock, at date t will be denoted by R_t and P_t respectively, and h is the size ofthe tick by which the stock can rise. We have,

$$\frac{R_{t+1}}{R_t} = \frac{(P_{t+1} - P_t)/P_t}{(P_t - P_{t-1})/P_{t-1}} = \frac{n_{t+1}h/P_t}{n_t h/(P_t - n_t h)}. \tag{1}$$

[1] In a different context, it has been noted that finite measurements, or rounding errors, may cause interesting patterns in the physical sciences (Szpiro 1993)

The right hand side of the equation arises because P_{t+1} equals $P_t + n_{t+1}h$ and P_t equals $P_{t-1} + n_t h$ where n_t is the number of ticks by which the price of the stock increased in the time period t-1 to t. The behavior of this equation is analyzed in the following section.

To investigate equation (1), the locus of

$$X\colon\ R_t = \frac{n_t h}{P_t - n_t h} \quad \text{and} \quad Y\colon\ R_{t+1} = \frac{n_{t+1} h}{P_t} \tag{2}$$

will be plotted in two-dimensional space. It should be noted that this plot, which is often called a phase-portrait in the physical sciences, is not single-valued, since different combinations of n_t , n_{t+1}, and P_t, may result in identical loci in the figure. Hence, to begin the analysis, I hold the price of the share constant (P_t = 100). (Throughout the paper the tick size h is set equal to 1.0, and the integers n_t, and n_{t+1} vary between -5 and +5.) As CL have pointed out, the system defines a grid (Figure 1). However, on closer inspection ofthe figure, one may note that this grid is not regular: as follows from equation (1) the distance between the vertical gridlines expands with n_t.[2]

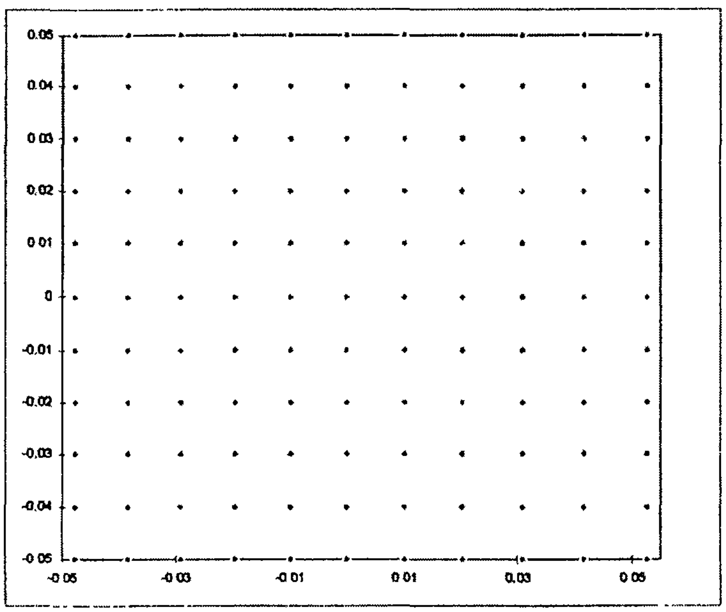


Figure 1: The grid (P_t = 100)

---

[2] This effect becomes more pronounced when the stock price is low in relation to the tick size (i.e., when the ticks are coarse).

Now in addition to n_t and n_{t+1}, let P_t also vary. CL showed that this results in a "smeared grid". If P_t varies, but not too much, the pattern begins to emerge. Figure 2 depicts the loci of system (2) for four prices (P_t = 100, 105, 110, 115). The most prominent feature of the figure is that the compass rose is actually made up of separate segments. For each combination of n_t, and n_{t+1}, the loci corresponding to a collection of P_t-values from a cluster which, in effect, is a smeared gridpoint. Note that smearing does not quite occur in the directions of the compass rose: since the distance between the horizontal gridlines expand horizontally the rays exhibits curvature. Again, this phenomenon becomes more pronounced if the tick size is coarse.

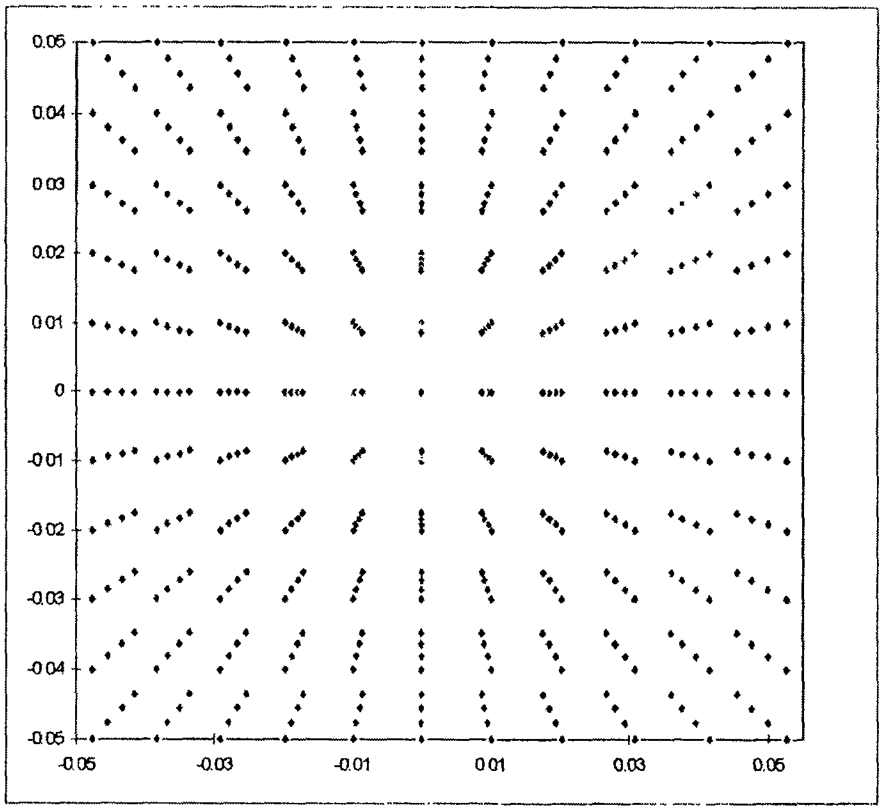


Figure 2: Smeared grid (P_t = 100, 105, 110, 115)

If the P_t-values vary over a sufficiently wide range the clusters overlap. To show this, it must be proved that the "endpoints" of the clusters connect, and that their slopes are equal whenever the loci coincide. Let us look at two clusters that lie on the same ray δ/ε, one belonging to $n_{t+1} = \lambda\delta$, $n_t = \lambda\varepsilon$, and the other belonging to the following cluster $n_{t+1} = (\lambda+1)\delta$ and $n_t = (\lambda+1)\varepsilon$. The corresponding loci are

$$X_1 = \frac{\lambda\varepsilon}{P} \qquad Y_1 = \frac{\lambda\delta}{P - \lambda\delta}$$
$$X_2 = \frac{(\lambda+1)\varepsilon}{Q} \qquad Y_2 = \frac{(\lambda+1)\delta}{Q - (\lambda+1)\delta} \tag{3}$$

Obviously , X_1 coincides with X_2, and Y_1 with Y_2 if

$$Q = \left(1 + \frac{1}{\lambda}\right)P. \tag{4}$$

Hence when the share prices span values between P and $P(1+1/\lambda)$ the clusters connect; when they span a wider range the clusters overlap and the points belonging to them intersperse. Note that when P grows the higher-order clusters (those with large λ) are the first ones to connect. For all clusters to connect, the price must vary between P and $2P$. The slopes at a certain locus are identical even if the clusters belong to different rays. The slopes at the loci defined by price P in the first cluster and by price Q in the second cluster are

$$Slope_1 = \frac{\varepsilon}{\delta}\left(\frac{P}{P - \lambda\delta}\right)^2 \quad \text{and} \quad Slope_2 = \frac{\varepsilon}{\delta}\left(\frac{Q}{Q - (\lambda+1)\delta}\right)^2. \tag{5}$$

If $Q = P(1+1/\lambda)$, the slopes are equal.

If the P_t-values vary sufficiently during the observation period, curvature and interspersion cause the phase-portrait to become "smudged" (in addition to being smeared), and the compass rose seems to disappcar. Only the major directions remain delineatcd. In Figure 3 the range of share prices is larger than in Figure 2 *(P_t = 100, 118, 136, 154)*, and only the major rays can be made out in the disarray. The question is why major directions stand out even after the compass rose has become smeared and smudged. As I will now show, the answer is based on an optical illusion.

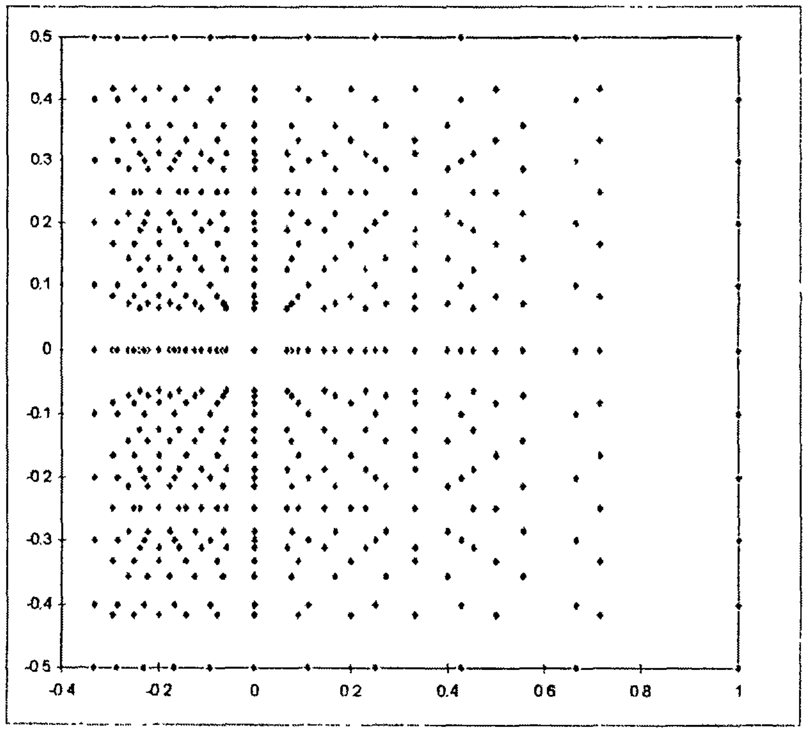


Figure 3: Smeared, smudged grid (P_t = 100, 118, 136, 154)

When stock prices span values between P and Q the points in the phase-portrait cover a distance of $\varepsilon(1/P - 1/Q)$ in the y-direction, and of $\delta(1/(P-\delta) - 1/(Q-\delta))$ in the x-direction. Hence, the larger ε and/or δ are, the more the points in the phase-portrait are pulled apart. As an example let us analyze the first two clusters on a major ray and on a minor ray, respectively. In the North-West direction the first two clusters belong to $\varepsilon/\delta = 1/1$ and to $\varepsilon/\delta = 2/2$. In a minor direction, say WWNW[3], the first two clusters belong to $\varepsilon/\delta = 1/4$ and to $\varepsilon/\delta = 2/8$. It follows from equation (7) that the clusters on each of the two rays connect when the stock prices span, say, the values from 100 to 200. However, in the phase-portrait the clusters on the WWNW ray are separated by a distance that is four times as large as the distance for the NW ray. Hence by increasing the stock prices from 100 to 200, the clusters of WWNW are pulled asunder to such an extent, that the human

[3] As CL pointed out, the rays do not correspond exactly to the compass rose. WWNW in the compass rose corresponds to an angle of $^1/_4(45°) = 11.25°$, while WWNW , as used in this paper , corresponds to $\tan^{-1}(^1/_4) = 14.03°$

eye no longer identifies them as lying on a ray.[4] The mentioned curvature would further obscure the rays. To make this point in another way, note that it is much more difficult for minor rays to get close to the origin in the phase-portrait than it is for major rays: in order for the NW ray to reach the point $x = 0.01$, a share price of 100 suffices, while a share price of 400 would be required for the WWNW ray to reach that point.

There is some asymmetry about the y-axis which follows from equation (2) when n_t is taken once as positive and once as negative. The deviation from symmetry for two members of a pair of rays can be computed as,

$$\Delta = \frac{(n_t h)}{(P - n_t h)} - \frac{(n_t h)}{(P + n_t h)} = \frac{2}{\left(\frac{P}{n_t h}\right)^2 - 1}. \tag{6}$$

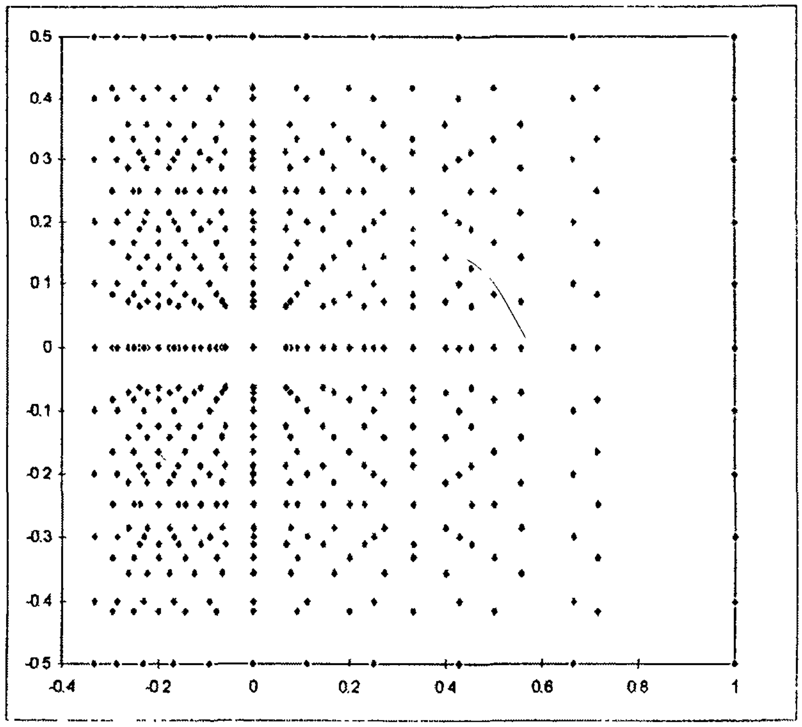


Figure 4: Coarse tick size (P_t = 10, 12, 14, 16)

[4] Furthermore, for large ε and/or δ and/or small variations in the price of the stock the connections often lie outside the range that is depicted in the figure.

To portray the discussed features in a salient manner, I now plot the situation with a coarse tick size (P_t = 10, 12, 14, 16). In Figure 4 the compass rose, curvature, asymmetry and the disappearance of the minor rays can be made out.[5]

3. EMPIRICAL EVIDENCE

In order to illustrate the phenomena that were predicted in the previous section, we use data from the Dollar-Deutschmark foreign exchange market. The data consist of a high-frequency time- series of bid and ask quotes, whose arithmetic means are taken as proxies for actual trades. Nearly 1.5 million observations were collected for the period between October 1st 1992 to September 31st 1993, and the series was made available to the academic community.[6] During the observation period the exchange rate varied from a low of 1.3950 DM/$ to a high of 1.7455, that is, the rate fluctuated by about 25 percent. This implies, by equation (4), that the fourth, fifth and higher clusters connect.

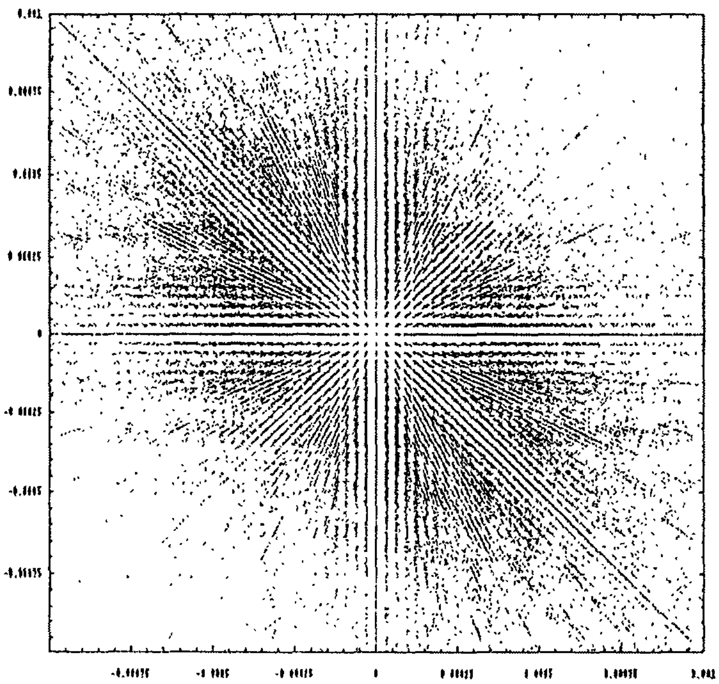


Figure 5: Compass Rose of the $/DM market

[5] Some new pattems also emerge, which are not germane to the discussion, however.
[6] Olsen & Associates, Institute for Applied Economic Research, Zürich.

In Figure 5 the values of R_{t+1} are plotted against R_t for a sample of 250,000 observations. A smeared and smudged grid-like system, resembling the compass rose is visible. When inspecting the blown up part of the figure (Figure 6), some of the features that were discussed in the previous section become evident. Especially the clusters are clearly visible. On the NW ray it can be seen that the fourth and the fifth cluster do, in fact, connect, while clusters of lower order remain separated.

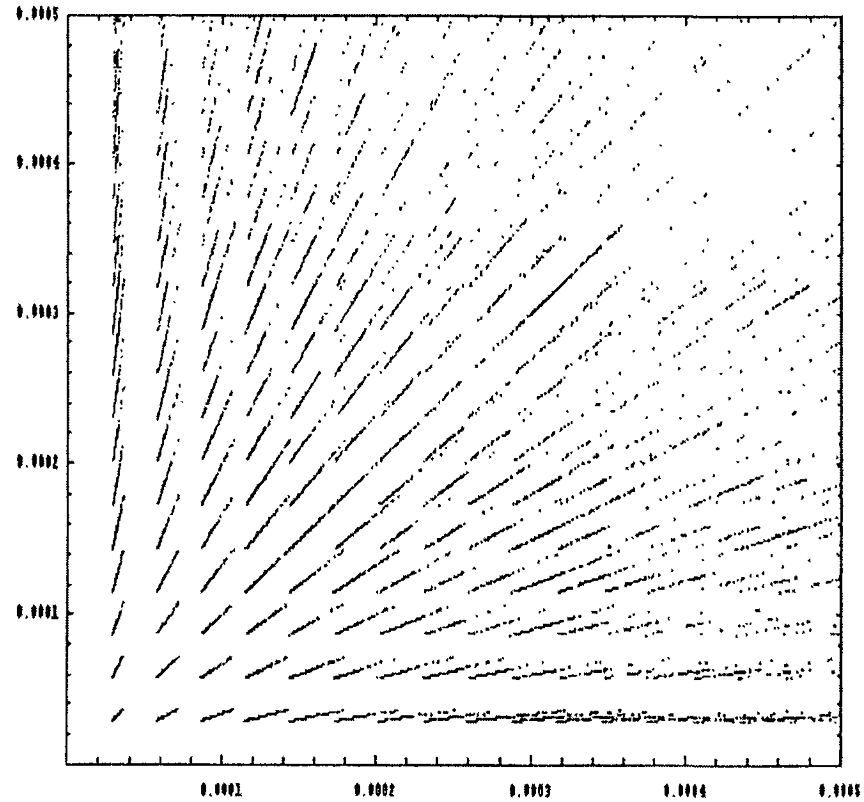


Figure 6: Detail of Figure 5

REFERENCES

Brock, W. A., Dechert, D., LeBaron, B. and Scheinkman, J. A. (1996) A test for independence based on the correlation dimension, *Econometric Reviews* 15, 197-235.

Crack, T. F. and Ledoit, O (1996) Robust structure without predictability: the 'Compass Rose' pattern in the stock market, *J. of Finance* 51, 751-762

Farmer, J. D. and Sidorowich, J. J: (1987) Forecasting chaotic time-series, *Physical Review Letters* 59, 845-848.

Grassberger, P. and Procaccia, I. (1983) Characterization of strange attractors, *Physical Review Letters* 48, 346-349.

Kramer, W., and Runde, R. (1997) Chaos and the compass rose, *Economics Letters* 54, 113-118.

NYSE (1997) Press Release, June 5th, 1997.

Szpiro, G. G. (1993) Cycles and circles in roundofferrors, *Physical Review* E 47, 4560-4563.

Szpiro, G. G. (1997) Noise in unspecified, non-linear time-series, *Journal of Econometrics* 78, 1997 , 229-255.

Szpiro, George G. (1998) Tick Size, the Compass Rose and Market Nano-Structure, *Journal of Banking and Finance*, forthcoming.

Symmetry: Culture and Science
Vol 9, Nos. 2-4, 431-440, 1998

OPTIMIZING LOTUS FLOWERS

Tibor Tarnai

Address: Technical University of Budapest, Department of Structural Mechanics, Budapest, Műegyetem rkp 3., H-1521, Hungary. E-mail: tarnai@ep-mech.me bme.hu

Abstract: *It has been discovered that in many cases the fruits in the receptacle of the sacred lotus* (Nelumbo nucifera) *are arranged in accordance with the solution to the following geometrical problem: How must* n *non-overlapping equal circles be packed in the unit circle so that the diameter of the circles will be as large as possible? In the paper an account of this problem (its putative solutions, and related configurations in lotus receptacles, traditional Japanese mathematics and family crests) is presented.*

1. INTRODUCTION

Design of ropes and cables has raised the following mathematical problem [1] which became one of the classical problems of discrete geometry: How must n non-overlapping equal circles be packed in the unit circle so that the diameter d of the circles will be as large as possible? [2, 3] This problem originated in engineering can be investigated in an engineering way. To an arrangement of circles there corresponds a *graph*. The vertices of the graph are the centres of the circles, and the edges of the graph are straight-line segments joining the centres of the touching circles. This graph is considered as a bar-and-joint structure supported along the boundary of a circle of radius $1 - d/2$, and a configuration is looked for where the bar length is a maximum, the distances between joints are not smaller than the bar lengths, the structure is rigid and it can have a stable state of self-stress without tensional force in bars. The solution of the structural problem corresponds to a locally optimal solution of the mathematical problem.

Over the years other methods have also been applied to solve this problem. Very recently Graham et al. [4] applied a billiard algorithm with which they produced putative solutions of the above packing problem for $n \leq 65$. Interestingly enough the solutions for small values of n can be seen in old Wasan (Japanese classical mathematics) books [5, 6] and Mon (Japanese family crest) collections [7] made long ago. In ancient Egypt, packing of equal circles in a circle was also used to measure the area of the circle [8]. Recently we discovered that the carpels in the receptacle of lotus flowers are arranged in accordance with dense packing [9]. This property is revealed even better for the arrangement of the lotus fruits in the receptacle.

The aim of this paper is to make a short comparison between the lotus fruit arrangements and the solutions of the circle packing problems, and to have a look at the related configurations in Wasan, and Mon design.

2. LOTUS

Lotus (*Nelumbo nucifera*) flowers (Figure 1(a)) have a conical receptacle with a circular even upper surface. In that circular face there are several small depressions, each of which contains one carpel. At first sight it seems that the carpels are arranged spirally [10]. On close inspection, however, we have found that the arrangement of carpels, and later the arrangement of fruits, in the receptacle is in accordance with the solution to the above circle packing problem. The solution provides the most economical configuration of circles because the area of the given circular domain covered by the circles is a maximum. So, *Nelumbo* optimizes the fruit arrangement in the receptacle. Since in living objects there are always imperfections and mutations, this is not a rigid law but only a prevalent tendency. Comparing actual fruit arrangements (Figure 2) with the mathematical solutions (Figures 6 and 7), the agreement is striking, disagreement is found mainly in those cases where the upper surface of the receptacle is distorted and/or the fruits are not equal. This is in accordance with the fact that the solutions of mathematical packing problems are sensitive for changes in data. Close packings of n circles, for instance, result in different configurations in a circle and in a domain different from a circle.

It is interesting to mention that in artistic representations of lotus sometimes not the optimal fruit arrangements are given. For example, the decoration in the Golden Pavilion, Horyuji temple in Nara, Japan, shows lotus flowers with 19 carpels (Figure 1(b)) arranged concentrically with carpel numbers: 1, 7, 11. In the optimal case, however, the arrangement is: 1, 6, 12 (Figure 2, n = 19).

Figure 1: Lotus flower and its representation as decoration in Horyuji temple (courtesy of Professor K. Miyazaki)

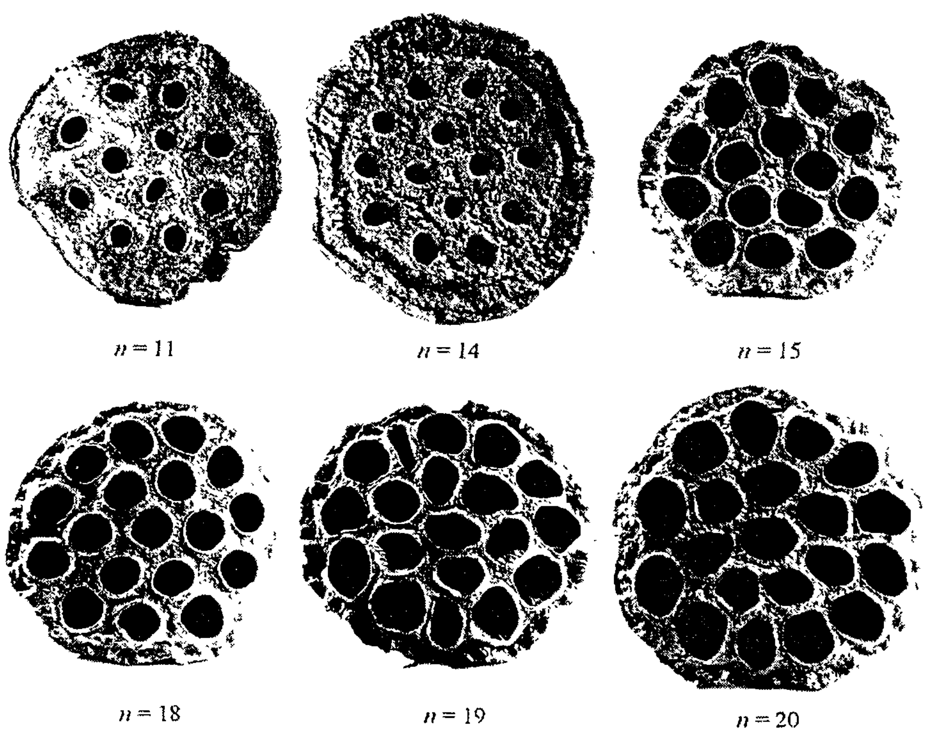


Figure 2: Arrangement of different numbers of fruits in the lotus receptacle (depressions in dried specimens)

3. MON (JAPANESE FAMILY CREST)

Circles have been very frequently used as motifs in Japanese family crests [7], often in close packings. Packing of circles in a circle is quite common (Figure 3). Many of them are closely related to the solution to the problem of the densest circle packing in a circle for $n = 2$ to 9. Although lotus is an important flower in Japan, an important symbol in Buddhism, it seems that its carpel does not appear among the family crests, circles do not represent fruits of lotus, at least it is not declared. Circles in Figure 3 have different meanings: stars (1, 2, 4-6, 14-17, 22-30), dragon's eye or snake's eye (7-12), weight used on balance-type scales (13), plum blossom (3, 15), tomoe (18-21).

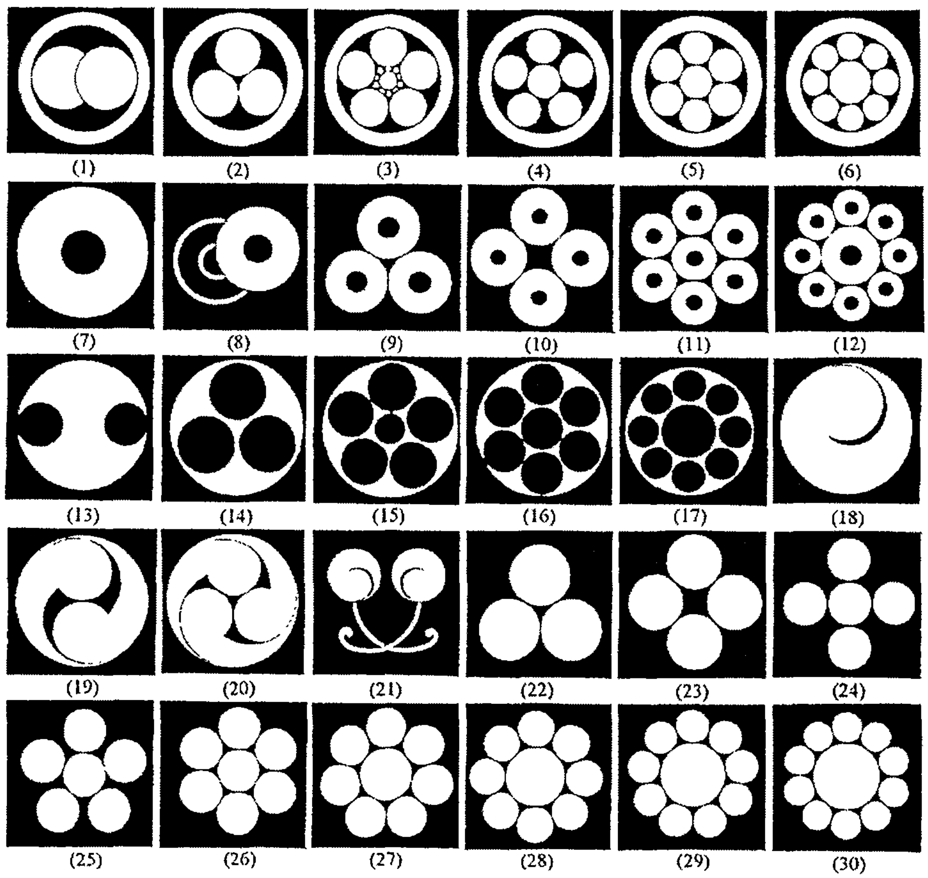


Figure 3: Circle packings in a circle in Japanese family crests

4. WASAN (JAPANESE CLASSICAL MATHEMATICS)

The Japanese mathematics wasan in the 17^{th}-19^{th} centuries discussed different problems where one had to find an unknown quantity from given quantities under certain conditions. In geometrical problems such quantities were, for instance, length of line segments or diameter of circles. At that time if someone had a nice mathematical discovery, then he put it down on a wooden board, called *sangaku*, dedicated it to gods and hung it up under the roof of a shrine or a temple [11]. Most sangaku problems are geometrical. The sangaku problems were often collected and published in a book. Many of these geometrical problems are related to circles, quite frequently to close packings of circles, particularly to densest packings of equal circles in a circle. In general the problem is not what the densest packing is, but to determine a distance, a circle radius or area of a domain in a close packing under constraints. Figure 4 shows figures of problems in different old wasan books. (a) [6] and (b) [5] are the densest packings of 5 and 7 equal circles. (c) [5] gives the densest packing of 9 equal circles if the diameter of the central circle is reduced. (d) [12] provides the densest packing of 10 equal circles with two axes of symmetry. It is interesting to mention that Kravitz [1] has considered this configuration as conjectured solution of the densest packing problem without symmetry constraints. The correct solution was given later by Pirl [13] (Figure 6, n = 10). (e) [14] shows a 3-dimensional configuration, but in the picture there is a packing of 12 circles of two different sizes in a circle. If the diameter of the central three circles is increased and that of the nine circles along the boundary is decreased, then with a slight modification, the densest packing of 12 equal circles (Figure 6, n = 12) is obtained. (f) [14] is one of the densest packings of 18 equal circles (Figure 7(a)).

Why did the Japanese have an interest in the densest packing problems? D. Nagy [15] asks this question in one of his papers analysing the connection between old Japanese mathematics and modern discrete geometry. He thinks that one of the probable reasons can be that the Japanese posed practical problems concerning economical cuttings and arrangements of things. One of the earliest problems published in wasan books [16] is counting barrels piled up in a triangular form. If the pile contains k rows of barrels, then the solution is the triangular number $t(k) = k(k+1)/2$. Knowing this result one can answer the question: How many bamboo sticks are there in a bundle containing one stick in the middle and k concentric layers of sticks hexagonally packed about it? [16] The answer is the hexagonal number $h(k) = 3k(k+1)+1$. In the case k = 1, we have $h(1) = 7$ (Figure 4(b)); in the case k = 2, we have $h(2)$ = 19 (Figure 4(f), one circle in the middle is added). The interesting point is that the Japanese have known in practice that equal bamboo sticks of hexagonal number can be arranged in a circular cylinder form (Figure 5) [16, 17].

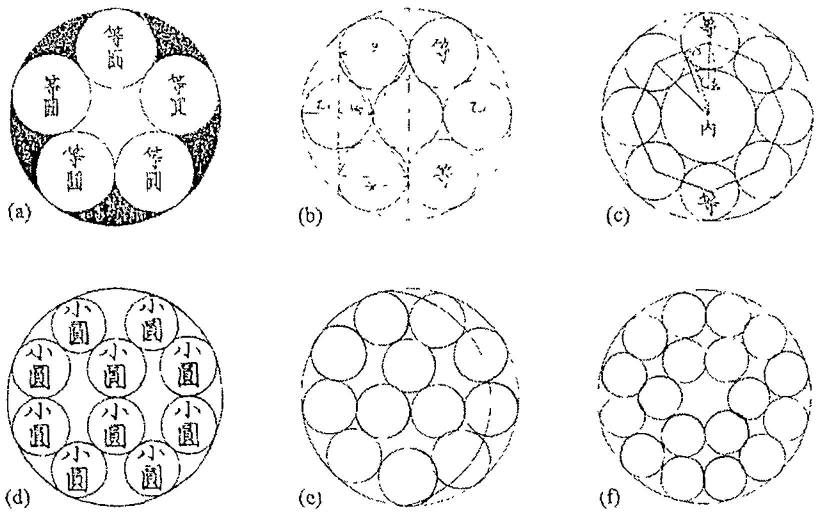


Figure 4: Packings of circles in a circle illustrations in old wasan books

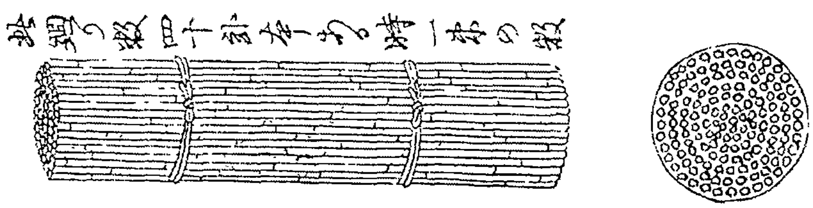


Figure 5: Curved hexagonal packing of equal bamboo sticks: illustrations in old wasan books

5. DISCRETE GEOMETRY

Discrete geometry [2, 3, 22] discusses among others tilings, packings and coverings in 2, 3 or higher dimensions. Extremal problems like densest packings and thinnest coverings are of the utmost importance. Problems of densest packing of equal circles in domains of different shapes in the plane are intensively studied nowadays thanks to the rapidly developing computer-aided methods.

Proven solutions of the problem of densest packing of n non-overlapping equal circles in the unit circle are known up to n = 11 circles [13, 18, 19] and conjectured solutions are known for $12 \leq n \leq 24$ [1, 13, 20, 21]. Graham et al. [4] extended this range recently and provided putative solutions up to n = 65. The packing algorithm used by them is based on the *billiards* model of computational physics, worked out by B.D. Lubachevsky. They compiled the numerical data in a table and presented the figures of the configurations. Subfigures of our Figures 6-8 are also taken from there [4].

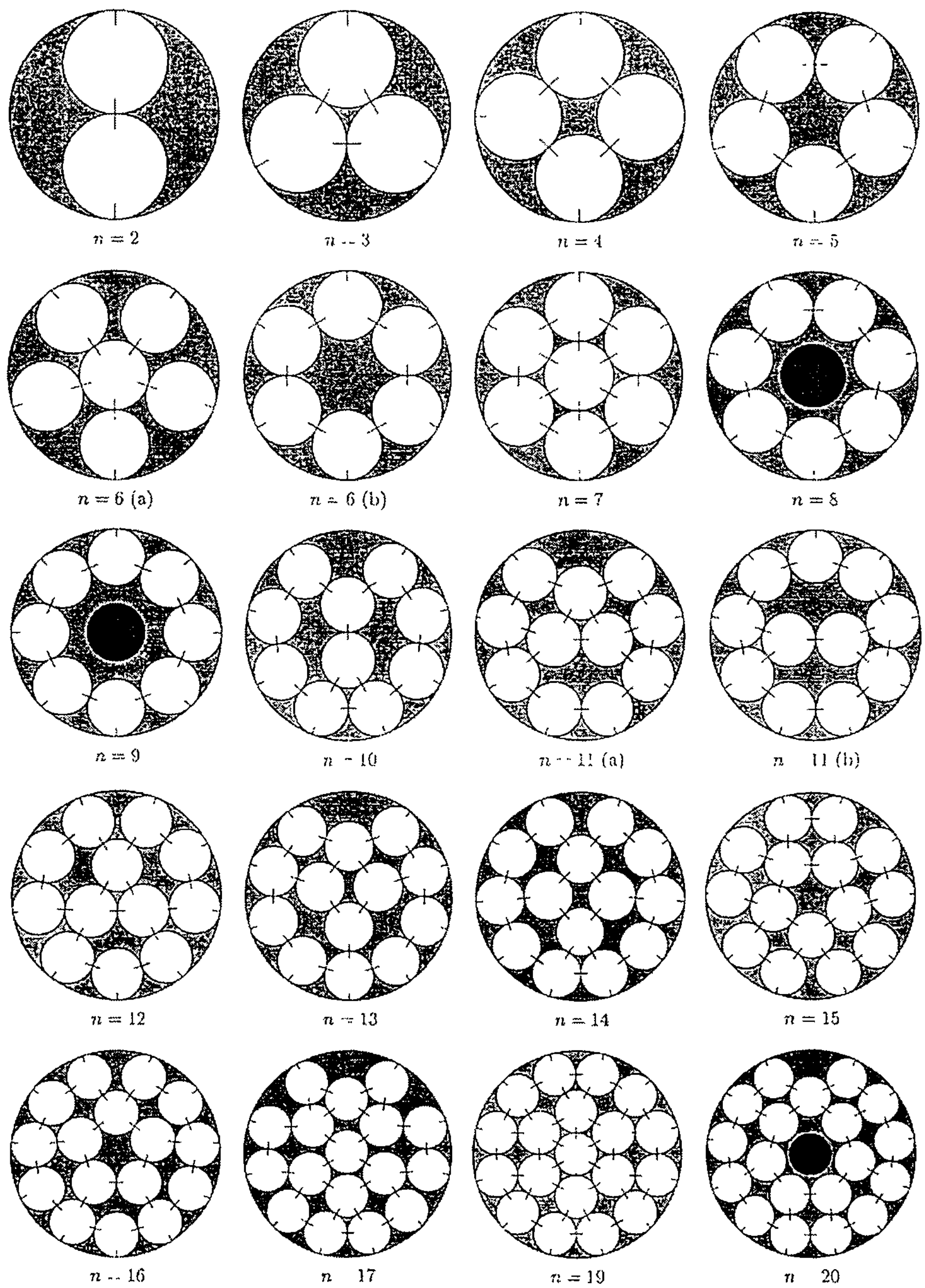


Figure 6: Proven and conjectured best packings of 2 to 17 and 19 to 20 equal circles in a circle (courtesy of Dr. B D. Lubachevsky)

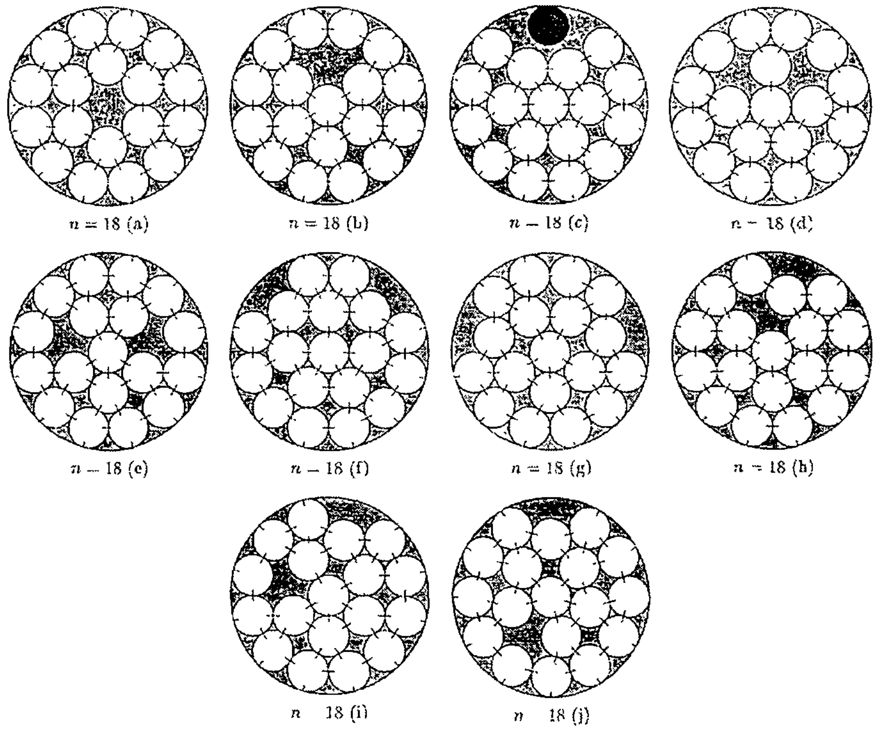


Figure 7: Conjectured best packings of 18 equal circles in a circle (courtesy of Dr. B.D. Lubachevsky)

Lubachevsky and Graham [23] have studied hexagonal dense packings of $h(k) = 3k(k+1)+1$ equal circles in a circle for up to k = 5. They have found that a dense circle packing in a circle can be obtained from that in a regular hexagon by deformation such that the number of circles in each circumferential layer is preserved, and the circumferential paths in the graph are closed (Figure 8). This explains why the Japanese could arrange equal bamboo sticks in a circular cylinder form.

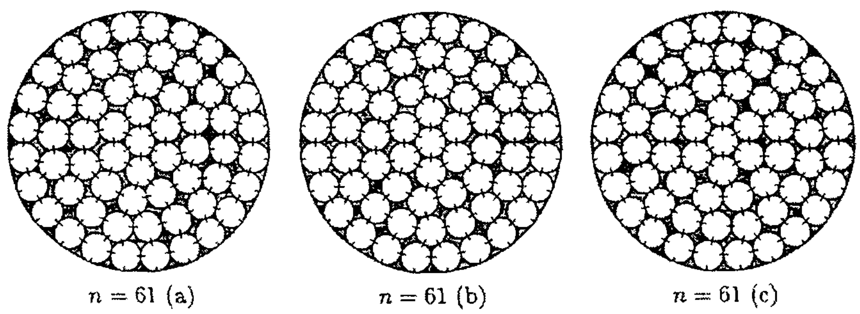


Figure 8: Curved hexagonal packings of 61 circles in a circle (courtesy of Dr B.D. Lubachevsky)

6. CONCLUSIONS

If the circles are represented by their centres, then the problem of the densest packing of equal circles in a circle is equivalent to the following problem: How must n points be distributed in a circle so as to maximize the least distance between any two of them? Therefore, lotus maximizes the minimum distance between fruits. This remarkable extremal property of *Nelumbo* receptacle is similar to that of certain spherical pollen grains, e.g. those of *Fumaria capreolata*, discovered earlier by Tammes, where the orifices on the surface of the pollen grain are arranged so [22]. This is a novel example that living nature can provide solutions to abstract mathematical problems, and so can give inspiration to solve practical (e.g., structural) problems.

Acknowledgements. I thank Professor Koji Miyazaki for help with collecting material and for advice, and Dr. B.D. Lubachevsky for sending me copies of his papers before their publication. The research work was supported by OTKA Grant No. T024037 awarded by the Hungarian Scientific Research Fund.

REFERENCES

[1] Kravitz, S. (1967) Packing cylinders into cylindrical containers, *Math. Mag.* 40, pp. 65-71.
[2] Croft, H.T., Falconer, K.J. and Guy, R.K. (1991) *Unsolved Problems in Geometry*, New York: Springer-Verlag.
[3] Moser, W and Pach, J. (1993) Research Problems in Discrete Geometry: Packing and Covering, *DIMACS Technical Report* 93-32, May.
[4] Graham, R.L., Lubachevsky, B.D., Nurmela, K.J. and Ostergard, P.R.J. (1997) Dense packings of congruent circles in a circle, *Discrete Math.* (to appear).
[5] Chiba, T. (1830) *Sanpo-shinsho* (in Japanese).
[6] Yamamoto, Y. (1841) *Sanpo-jyojyutsu* (in Japanese).
[7] Dower, J.W. (1991) *The Elements of Japanese Design*, New York: Weatherhill.
[8] Joseph, G.G. (1994) *The Crest of the Peacock*, p. 84, London: Penguin.
[9] Tarnai, T. (1998) Packing of equal circles in a circle, *Hasu no Hanashi* (in Japanese) 3, pp. 24-31.
[10] Watanabe, S. (1990) *The Fascinating World of Lotus*, Tokyo: Parks and Open Space Association of Japan.
[11] Fukagawa, H. and Pedoe, D (1989) *Japanese Temple Geometry Problems*, Winnipeg Canada: Charles Babbage Research Centre.
[12] Fujita, K. (1807) *Zoku-shinpeki-sanpo* (in Japanese).
[13] Pirl, U. (1969) Der Mindestabstand von n in der Einheistkreisscheibe gelegenen Punkten, *Math. Nachr.* 40, pp. 111-124
[14] Aida, Y (1800?) *Sanpo kiriko-shu* (Mathematical book on polyhedra, in Japanese).
[15] Nagy, D. (1994) Wasan (old Japanese mathematics) and discrete geometry: from ethnomathematics and aesthetics to education and research, *Extended Abstracts*, Int. *Katachi* $\cup$ *Symmetry Symp.*, Tsukuba, 1994, pp 328-332.
[16] Yamada, M. (1659) *Kaizanki* (in Japanese).
[17] *Sanpo-taizen-shinansho* (in Japanese), (the author is unknown, Edo era).
[18] Graham, R.L. (1968) Sets of points with given minimum separation (Solution to Problem E1921), *Amer. Math. Monthly* 75, pp 192-193.
[19] Melissen, H. (1994) Densest packing of eleven congruent circles in a circle, *Geom Dedicata* 50, 15-25

[20] Goldberg, M (1971) Packing of 14, 16, 17 and 20 circles in a circle, *Math Mag*. 44, pp. 134-139.
[21] Reis, G.E. (1975) Dense packing of equal circles within a circle, *Math Mag*. 48, pp 33-37.
[22] Fejes Tóth, L. (1964) *Regular Figures*, New York: Pergamon, Macmillan.
[23] Lubachevsky, B D and Graham, R L (1997) Curved hexagonal packings of equal discs in a circle, *Discrete & Computational Geometry* 18, pp. 179-194.

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 441-456, 1998

GEOMETRICAL INTERPRETATION: RECURRENCE FORMULA OF NATURAL NUMBER'S POWER SERIES SUMMATION

Yasunari Watanabe

Address: Department of Information Systems Teikyo Heisei University, Uruido Ichihara, Chiba 290-0153 Japan

Abstract: *The geometrical interpretation of natural number's* n-*th power series summation is presented by the superposition of the* j-*dimensional lattice of hyperspace regular simplexes,* α_j *in which* α_{j-1} *stacks with a linear increment weight of the lattice points. The superposition of* α_j *is carried out every* (j+*1*) *times rotation around the* (j+*1*)*-fold axis of* α_j. *The superpositions result in homogeneous density in* α_j. *The recurrence formula of sum of the power series of natural numbers is presented using the mean value of the superposed weight of the lattice points in the* α_j.

1. INTRODUCTION

Some geometrical interpretations of natural numbers' n-th power series summation were presented for n = 1, 2, 3 (Gardner 1986). On the other hand, a geometrical properties are discussed for second moment of inertia of α_j Voronoi cell around the origin (Conway & Sloane 1993). However, a geometrical interpretation of generalized number theory using α_j has not been proposed so far as the author is aware. In this paper a new geometrical scheme for the sum of the n-th power of natural numbers n = 1, 2, 3 is presented by using the geometry of the closest packing of circles (1- and 2-D cases) and that of spheres (3-D case) which is called face-centered close packed structure (Conway & Sloane 1993). As Coxeter pointed out, any j+1 points with equal distance between any two points which do not lie in a (j-1)-space are the vertices of a j-D simplex, as point α_0, line-segment α_1, regular triangle α_2, regular tetrahedron α_3, regular pentatope

α_4, ... , regular simplex α_j (Coxeter 1973). It is shown that higher-order power series of natural numbers can be obtained by considering further extension of the dimension, namely the lattice points in the hypersimplex α_j.

2. CASE STUDY OF SUPERPOSITION OF SIMPLEXES IN 1-D TO 4-D

First, the geometrical interpretation of the sum of low-dimensional (n = 1, 2, 3) power of natural numbers is presented concretely. Consider 1-D lattice L_1 with linear increment weight 1, 2, 3, ... , n on the lattice points. Let $R(\emptyset)$ be the rotation operator of α_1 with rotation angle $\emptyset_1$ around the [1, 1] zone axis in the 2-D square lattice. α_1 operated on by $R(\emptyset_1)$ becomes $\alpha_1(\emptyset_1)$, if α_1 is rewritten as $\alpha_1(0)$, then

$$\alpha_1(\emptyset_1) = R(\emptyset_1) \cdot \alpha_1(0). \tag{1}$$

The superposition of $\alpha_1(\emptyset_1) + \alpha_1(0)$ is represented as

$$S_1 \cdot \alpha_1(0) = \{R(\emptyset_1) + I\} \cdot \alpha_1(0) \tag{2}$$

where S_1 is the superposition operator and I is the identity operator of $\alpha_1(0)$. In the 1-D case $\emptyset_1$ is restricted to π, then

$$\cos \emptyset_1 = -1 \tag{3}$$

$$S_1 = R(\pi) + R(0) = R(\pi) + I \tag{4}$$

The weight W_1 of all the lattice points. $S \cdot \alpha_1(0)$ is the same weight $W_1 = n$ as shown in Figure 1.

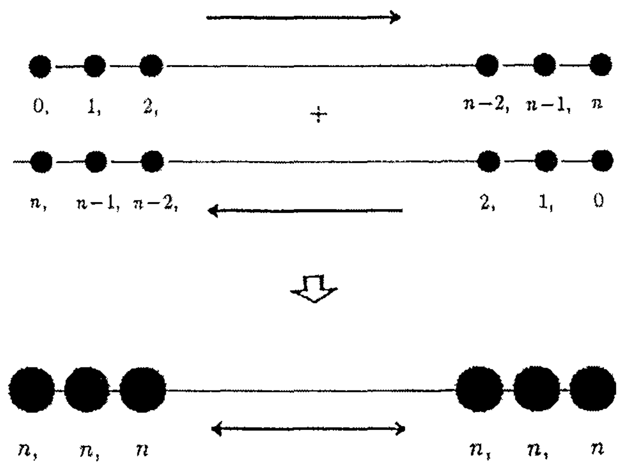


Figure 1: Line segment α_1 with linear increment weight and superposed α_1 with homogeneous density n.

If $\alpha_1(0)$ has nonlinear increment weight, $S_1\alpha_1(0)$ should have an inversion center or 2-fold rotational symmetry, but it has homogeneous density $W_1 = n$ for $\alpha_1(0)$ with linear increment weight. Let the number of lattice points of α_1 be l_1 and the multiplicity of superposition be M_1, so $l_1 = n + 1$ and $M_1 = 2$ for two superpositions.

The mean mass of $S \cdot \alpha_1(0)$,

$$l_1 = n + 1 \qquad W_1 = n \qquad M_1 = 2 \qquad \overline{m_1} = \frac{l_1 W_1}{M_1} = \frac{1}{2}n(n+1). \tag{5}$$

$\overline{m_1}$ m is equal to the mass of $L_1(0)m_1$, since the weight of the k-th point is k and the number of k-th points l_0 is 1,

$$l_0 = 1 \qquad m_1 = \sum_{k=0}^{n} k \bullet l_o. \tag{6}$$

Therefore,

$$\sum_{k=1}^{n} k = \frac{1}{2}n(n+1). \tag{7}$$

Hereafter the notations of α_j, S_j, W_j, M_j, l_j, m and $\overline{m_j}$ are introduced for j-D parameters.

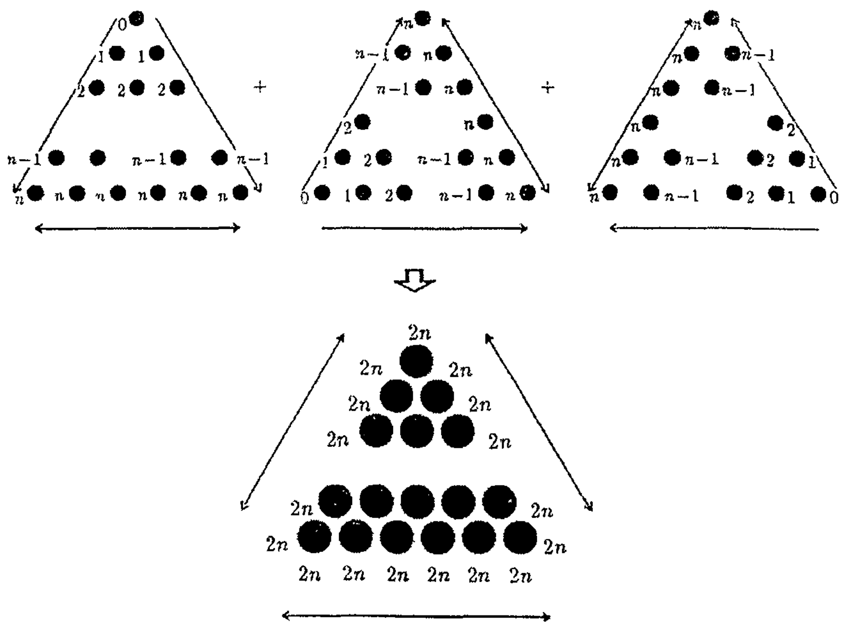


Figure 2: Regular triangle α_2 operated on by 0, $2\pi/3$ and $4\pi/3$ rotations and superposed α_2 with homogeneous density $2n$. Stacking layers of line segments α_1 with linear increment weight are shown by black circles.

Second, consider 2-D lattice $\alpha_2(0)$ which is the stacking of $\alpha_1(0)$, the length of which increases linearly as shown in Figure 2 to form a regular triangle lattice $\alpha_2(0)$. The superposition is applied to $\alpha_2(0)$ in the same way as in the 1-D case. $R\emptyset_2$ is considered using the [1 1 1] plane of the cubic lattice in the first quadrant which takes the form of a regular triangle. The superposition could be performed every $2\pi/3$ rotation successively around the [1 1 1] axis through the center of gravity of a regular triangle, that is,

$$\cos \emptyset_2 = -1/2 \tag{8}$$

$$\mathbf{S}_2 \cdot \alpha_2(0) = \{\mathbf{R}^2(\phi_2) + \mathbf{R}(\phi_2) + \mathbf{I}\} \cdot \alpha_2(0). \tag{9}$$

Here $R^2(\emptyset_2)$ is equal to $R(2\emptyset_2)$, thus

$$\mathbf{S}_2 = \mathbf{R}(\frac{4\pi}{3}) + \mathbf{R}(\frac{2\pi}{3}) + \mathbf{I}. \tag{10}$$

For S_2, three superpositions of $\alpha_2(0)$, a regular triangle with homogeneous density is obtained as shown in Figure 2; here,

$$M_2 = 3, \quad l_2 = \sum_{k=0}^{n}(k+1) = \frac{1}{2}(n+1)(n+2), \quad W_2 = 2n \tag{11}$$

are given as shown in Fig.2. The mean mass of $S_2 \cdot \alpha_2(0)$ is

$$\overline{m_2} = \frac{l_2 W_2}{M_2} = \frac{1}{3}n(n+1)(n+2). \tag{12}$$

The number of lattice points on the k-th layer line-segment in the triangle $\alpha_2(0)$ is l_2 and each lattice point of l_2 has unit weight, thus

$$l_0 = 1 \quad l_1 = \sum_{l_0=0}^{k} l_0 = k+1, \quad m_2 = \sum_{k=0}^{n} k l_1 = \sum_{k=0}^{n} k(k+1) \tag{13}$$

$$\sum_{k=1}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1). \tag{14}$$

Third, consider 3-D lattice $\alpha_3(0)$ by analogy with the 2-D lattice. A regular triangle of $\alpha_2(0)$ in 2-D space corresponds to a tetrahedron in which a face-centered close-packed lattice is formed. This geometry of $\alpha_3(0)$ in a regular tetrahedron is interpreted as n stacking layers of a regular triangle of $\alpha_2(0)$. That is,

$$M_3 = 4, \quad l_3 = \sum_{k=0}^{n}\frac{1}{2}(k+1)(k+2) = \frac{1}{6}(n+1)(n+2)(n+3), \quad W_3 = 3n. \tag{15}$$

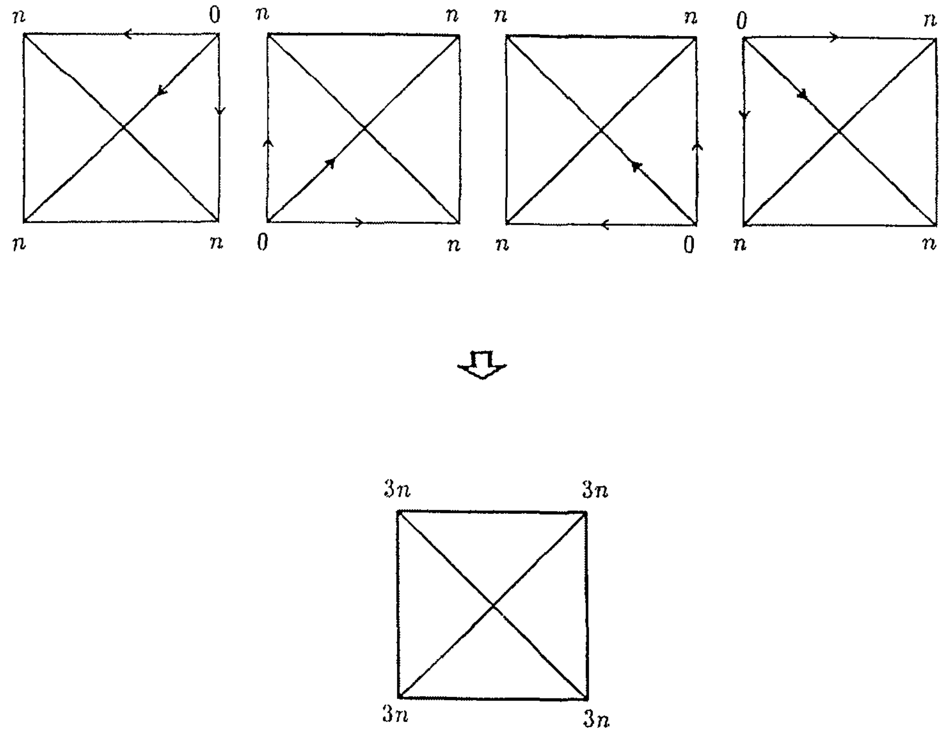


Figure 3: Regular tetrahedra (α_3) operated on by 0, $\emptyset_3$, $2\emptyset_3$, $3\emptyset_3$ rotations ($\cos \emptyset_3 = -1/3$) and superposed α_3 with homogeneous density $3n$. A sense of stacking layers of regular triangle (α_2), with increment weight is denoted by arrows

Thus, the mean mass of $S_3 \cdot \alpha_3$ is

$$\overline{m_3} = \frac{l_3 W_3}{M_3} = \frac{1}{8} n(n+1)(n+2)(n+3). \tag{16}$$

Let the number of lattice points in the k-th stacking layer be l_2, and the weight of all the lattice points in the k-th layer be k, then

$$l_2 = \sum_{l_1=0}^{k} l_1 = \frac{1}{2}(k+1)(k+2), \quad m_3 = \sum_{k=0}^{n} k l_2 = \sum_{k=0}^{n} k \sum_{l_1=0}^{k} \sum_{l_0=0}^{l_1} l_0 = \sum_{k=0}^{n} \frac{1}{2} k(k+1)(k+2). \tag{17}$$

From equations (14), (16) and (17),

$$\sum_{k=1}^{n} k^3 = \frac{1}{4} n^2 (n+1)^2. \tag{18}$$

As for $R(\emptyset_3)$, consider the rotation of a tetrahedron in 4-D space, because it is more comprehensive to select the rotation about any plane including the [1 1 1 1] axis in 4-D space than in 3-D space. Four kinds of tetrahedra obtained for every rotation by the tetrahedral angle are superposed as shown in Figure 3.

$$\cos \emptyset_3 = -1/3 \tag{19}$$

$$\mathbf{S}_3 \cdot \alpha_3(0) = \{\mathbf{R}^3(\phi_3) + \mathbf{R}^2(\phi_3) + \mathbf{R}(\phi_3) + \mathbf{I}\} \cdot \alpha_3(0) \tag{20}$$

where $\emptyset_3$ is equal to the tetrahedral angle.

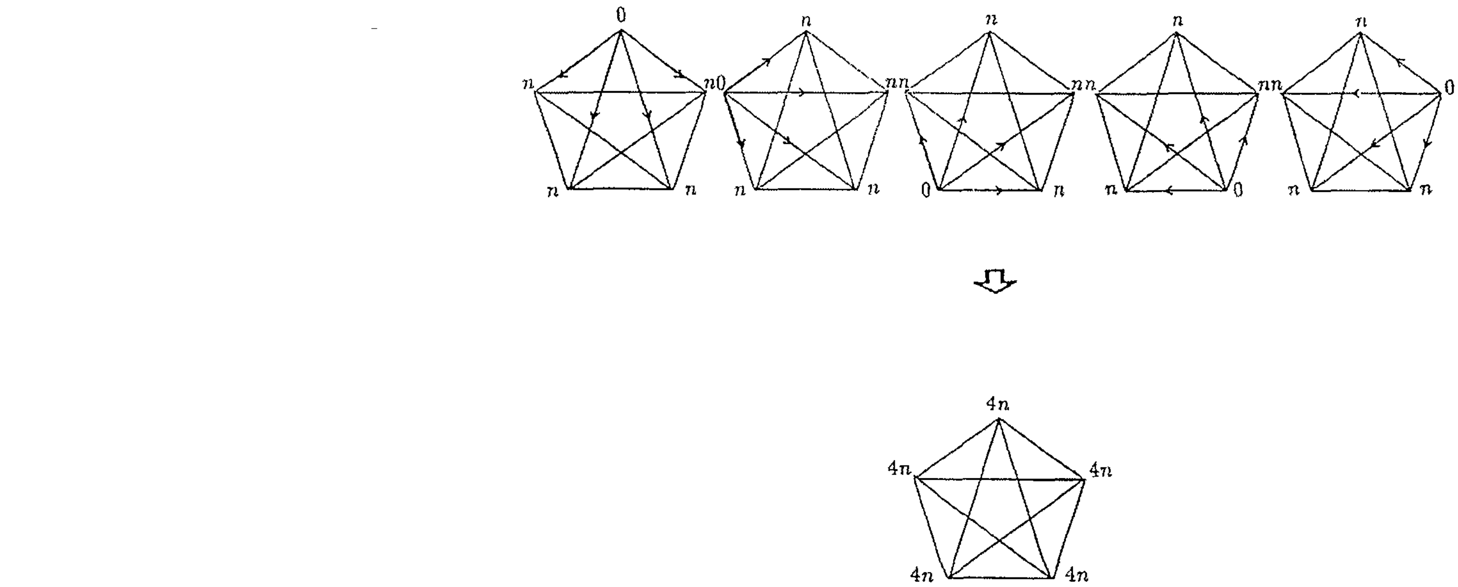


Figure 4: Regular pentatopes α_4 operated on by 0, $\emptyset_4$, $2\emptyset_4$, $3\emptyset_4$ and $4\emptyset_4$ rotations ($\cos \emptyset_4 = -1/4$) and superposed α_4 with homogeneous density $4n$. Sense of increment weight is denoted by arrows.

The practical examples of the 1-D, 2-D and 3-D superpositions of the lattice facilitate deduction of the 4-D and higher-dimensional geometry by analogy. It is necessary that the figure used for the superposition have a high rotational symmetry for the purpose of getting a homogeneous density. A regular polytope (regular hyper-simplex), α_j, is available to satisfy such a condition.

In the 4-D case five regular pentagonal simplexes α_4 are considered as superposition simplexes as shown in Figure 4. Stacking illustrations of j-D simplexes in j+1 dimensional space j = 2, 3, 4 are shown in Figure 5.

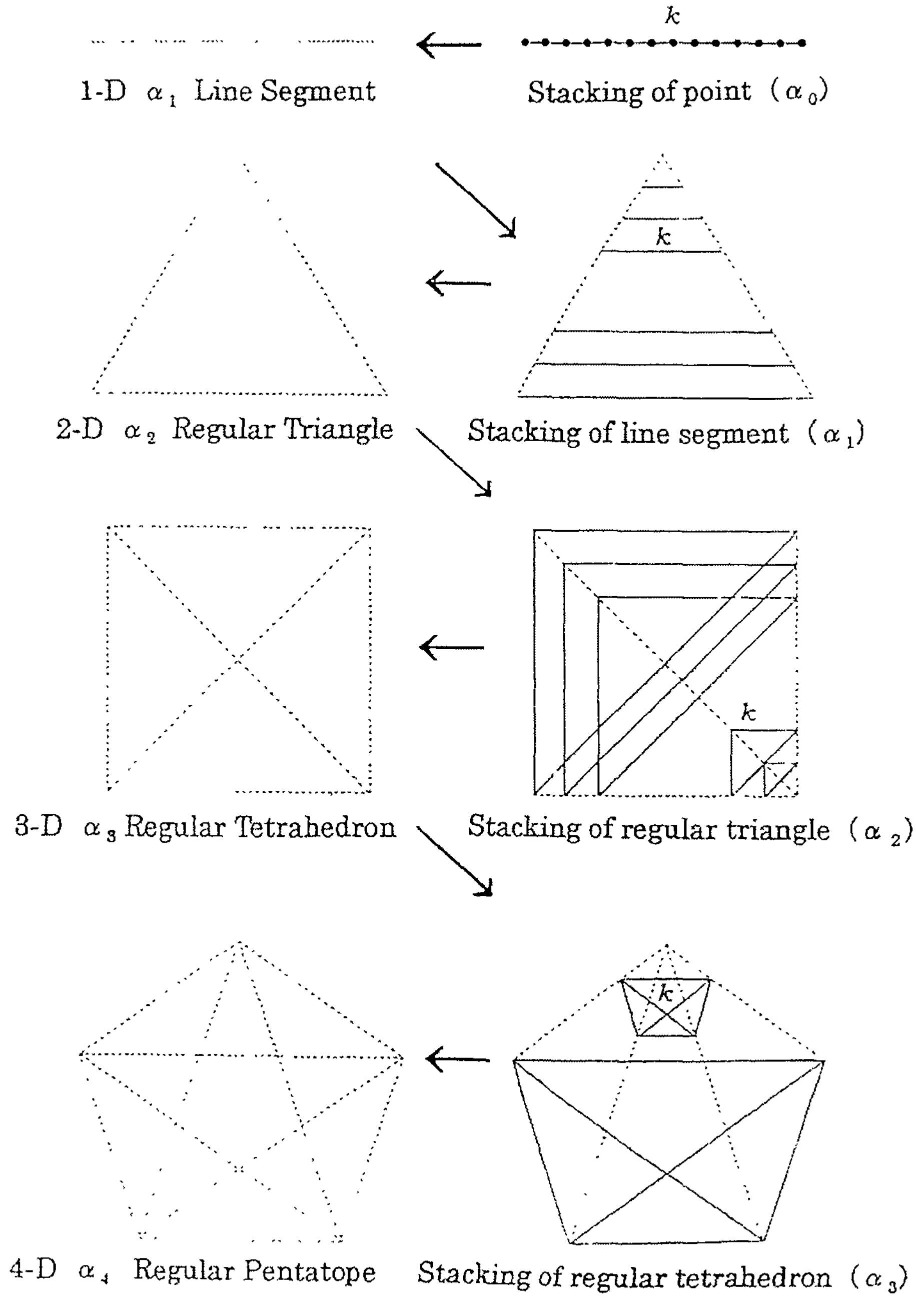


Figure 5: Stacking illustration of α_{j-1} in α_j. Stacking of line segments (α_1) in regular triangle (α_2), those of α_2 in regular tetrahedron (α_3) and those of α_3 in regular pentatope α_4.

3. DERIVATION OF $\overline{m_j}, m_j$

The sum of 1st, 2nd and 3rd power of natural numbers is derived by the superposition of line segment, triangle and tetrahedron, respectively. Here, consider j-D simplex α_j; since α_j is composed of n stacking layers of α_{j-1} and all lattice points in its k-th layer l_j have the same weight k, $l_{j,n}$ is represented as follows.

$$l_{j,n} = \sum_{k=0}^{n} l_{j-1,k} \tag{21}$$

where $l_{j-1,k}$ is the lattice point of the k-th layer with linearly increasing weight in the α_{j-1}. The j-D mean mass, $\overline{m_j}$, is deduced as follows from the superpositions of the 1-D to 3-D regular simplex; here l_j is a lattice point in α_j.

$$M_j = j + 1, \quad l_j = \frac{1}{j!}(n+1)(n+2)\cdots(n+j-1)(n+j), \quad W_j = jn \tag{22}$$

$$\overline{m_j} = \frac{l_j W_j}{M_j} = \frac{j}{(j+1)!} n(n+1)(n+2)\cdots(n+j) \tag{23}$$

The parameters of superposed hyper regular simplexes are summarized in Table 1.

j	W_j	$l_{j,n}$	M_j	$\overline{m_j}$
1	n	$n+1$	2	$\frac{1}{2}n(n+1)$
2	$2n$	$\frac{1}{2}(n+1)(n+2)$	3	$\frac{1}{3}n(n+1)(n+2)$
3	$3n$	$\frac{1}{6}(n+1)(n+2)(n+3)$	4	$\frac{1}{8}n(n+1)(n+2)(n+3)$
4	$4n$	$\frac{1}{24}(n+1)(n+2)(n+3)(n+4)$	5	$\frac{1}{30}n(n+1)(n+2)(n+3)(n+4)$
j	jn	$\frac{1}{j!}(n+1)(n+2)\cdots(n+j)$	$j+1$	$\frac{j}{(j+1)!}n(n+1)(n+2)\cdots(n+j)$

Table 1: Parameters of superposed hyper regular simplex

On the other hand, m_j is formulated by stacking of the layers of (j-1)-th regular simplexes with weight k in each layer of the j-D space. For $j \geq 2$,

$$m_j = \sum_k^n k l_{j-1,k} = \sum_k^n k \sum_{l_{j-1}}^{k} \sum_{l_{j-2}}^{l_{j-1}} \cdots \sum_{l_3}^{l_4} \sum_{l_2}^{l_3} l_2 = \sum_k^n k \frac{1}{(j-1)!} \sum_l^k (l+1)(l+2) \cdots (l+j-2). \tag{24}$$

In order to obtain the summation term of the right side of equation (24), consider the product function $f_j(x)$,

$$f_j(x) = (x+1)(x+2)\cdots(x+j-1). \tag{25}$$

From equation (25),

$$f_{j+1}(n) - f_{j+1}(n-1) = (n+1)(n+2)\cdots(n+j) - n(n+1)\cdots(n+j-1) = j f_j(n). \tag{26}$$

Therefore,

$$\begin{aligned}\sum_{k=0}^{n} f_j(k) &= \frac{1}{j}\sum_{k=0}^{n}\{f_{j+1}(k) - f_{j+1}(k-1)\} \\ &= \frac{1}{j}\{f_{j+1}(n) - f_{j+1}(-1)\} = \frac{1}{j}\cdot\frac{(n+j)!}{n!}.\end{aligned} \tag{27}$$

Consequently,

$$\sum_k^n (k+1)(k+2)\cdots(k+j-1) = \frac{1}{j}\cdot\frac{(n+j)!}{n!} \tag{28}$$

can be obtained. Here, another product function $4xf_j(x)$ is introduced in order to represent the right side of equation (24) by the parameter n. Equation (29) is derived easily in the same procedure as obtained equation (28).

$$\sum_k^n k(k+1)(k+2)\cdots(k+j-1) = \frac{1}{(j+1)}\cdot\frac{(n+j)!}{n!} \tag{29}$$

Substituting the right side of equation (29) into equation (24) , equation (30) is obtained.

$$m_j = \frac{1}{(j-1)!}\sum_k^n k(k+1)(k+2)\cdots(k+j-1) = \frac{j}{(j+1)!}(n+1)(n+2)\cdots(n+j). \tag{30}$$

Thus $\overline{m}_j = m_j$ is confirmed by the geometrical interpretation of the superposition of the regular simplexes. The relation is derived from $\overline{m}_j = m_j$ as shown in equation (30).

Equations (23) and (24) are verified by mathematical induction using the stacking geometry. Let the number of lattice points in the k-th layer of (j-1)-D be $l_{j-1,k}$ as given in equation (24), in which all lattice points have the same weight k. $l_{j,k}$ is derived as follows using equation (28).

$$l_{j,k} = \sum_{l_{j-1}}^{k}\sum_{l_{j-2}}^{l_{j-1}}\sum_{l_{j-3}}^{l_{j-2}}\cdots\sum_{l_2}^{l_3} l_2 = \frac{1}{(j-1)!}\sum_{l}^{k}(l+1)(l+2)\cdots(l+j-2)$$
$$= \frac{1}{j!}(k+1)(k+2)\cdots(k+j), \quad (j \geq 2) \tag{31}$$

Then,

$$m_{j+1} = \sum_k^n k l_{j,k} = \sum_k^n k \sum_{l'}^{k}\frac{1}{(j-1)!}(l'+1)(l'+2)\cdots(l'+j-2)$$
$$= \sum_k^n \frac{1}{j!}k(k+1)(k+2)\cdots(k+j-1)$$
$$= \frac{(j+1)}{(j+2)!}n(n+1)(n+2)\cdots(n+j+1), \tag{32}$$

therefore, equation (23) holds for (j+1)-D mass. Here, considering the expanded form of equation (30), the sum of j-th power of natural numbers is represented as a linear combination of sums of (j-1)-th~2nd power of natural numbers k as follows.

$$\sum_k^n k^j = \frac{1}{(j+1)} n(n+1)(n+2)\cdots(n+j) - \{1+2+\cdots+(j-1)\}\sum_k^n k^{j-1}$$
$$-\{1\cdot 2+1\cdot 3+\cdots+2\cdot 3+2\cdot 4+\cdots+(j-2)(j-1)\}\sum_k^n k^{j-2}$$
$$-\{1\cdot 2\cdot 3+1\cdot 2\cdot 4+\cdots+2\cdot 3\cdot 4+2\cdot 3\cdot 5+\cdots+(j-3)(j-2)(j-1)\}\sum_k^n k^{j-3}$$
$$-\cdots\cdots-(j-1)!\sum_k^n k, \qquad (j \geq 2). \tag{33}$$

The coefficients of sums of powers of natural numbers in the right side of equation (33) are given by the sum of combination products, $j\text{-}1C_1$, $j\text{-}1C_2$, ... , $j\text{-}1C_{j\text{-}1}$ of natural numbers 1, 2, 3, ... , j-1.

4. PROJECTION MATRIX FOR ROTATION ANGLE OF SUPERPOSITION

The rotation angle in the superposition operator is determined using body diagonal projection of a hypercube from j-D to (j-1)-D space. Consider a projection of an arbitrary vector of hyper cubic lattice $a(a_1, a_2, a_3, \ldots, a_j)$ along j-D projection vector $p(p_1, p_2, p_3, \ldots, p_j)$. Here projected vector $a_\perp$ is given by

$$\mathbf{a}_\perp = \mathbf{a} - \frac{\mathbf{p}}{|\mathbf{p}|}|\mathbf{a}|\cos\theta \tag{34},$$

where $\frac{p}{|p|}$ is a unit vectors along p and $\emptyset$ is an angle between p and a, then $a_\perp$ is written as

$$\mathbf{a}_\perp = \mathbf{a} - \frac{\mathbf{p}(\mathbf{p}\cdot\mathbf{a})}{|\mathbf{p}|^2}. \tag{35}$$

The geometry of a projection along $p_\perp$ is shown in Figure 6.

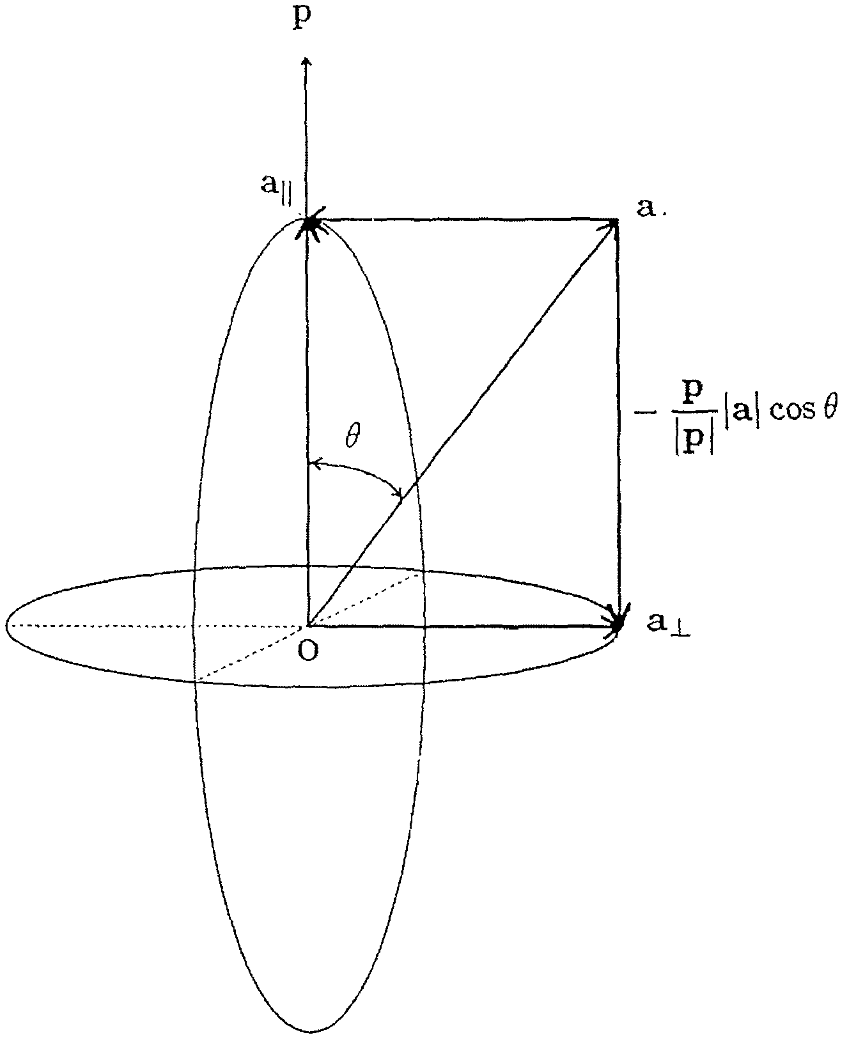


Figure 6: Geometry of j-dimensional orthogonal projection of a to planenormal to projection vector $p(p_1, p_2, p_3, \ldots, p_j)$, $a\perp$ and plane parallel p, $a_{\|}$

Geometry of j-dimensional orthogonal projection of a to plane normal to projection vector $p(p_1, p_2, p_3, \ldots, p_j)$, $a\perp$ and plane parallel to p, $a_{\|}$, j-D orthogonal base vectors a_i ($i = 1$~j) and projected vectors $a\perp_i$ are introduced where $p\perp$ is given by equation (36).

$$\mathbf{p} = \sum_{i}^{j} p_i \mathbf{a}_i, \qquad |\mathbf{p}|^2 = \sum_{i=1}^{j} p_i^2 \tag{36}$$

Substituting (36) for (35), a_i (i = 1~j) and projected vector $a\perp_i$ are introduced and the projection is formulated,

$$\mathbf{A}_\perp = \mathbf{P}_\perp \cdot \mathbf{A} \tag{37}$$

where column vectors of $A\perp$, $A_\parallel$ and projection matrix $P\perp$ are given as follows.

$$\mathbf{A}_\perp = \begin{pmatrix} \mathbf{a}_{1\perp} \\ \mathbf{a}_{2\perp} \\ \mathbf{a}_{3\perp} \\ \cdot \\ \cdot \\ \\ \mathbf{a}_{j-1\perp} \\ \mathbf{a}_{j\perp} \end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \cdot \\ \cdot \\ \\ \mathbf{a}_{j-1} \\ \mathbf{a}_j \end{pmatrix} \tag{38}$$

$$\mathbf{P}_\perp = \frac{1}{|\mathbf{p}|^2} \begin{pmatrix} |\mathbf{p}|^2 - p_1^2, & -p_1p_2 & -p_1p_3, & \cdot\ \cdot & -p_1p_{j-1}, & -p_1p_j \\ -p_2p_1, & |\mathbf{p}|^2 - p_2^2, & -p_2p_3, & \cdot\ \cdot & -p_2p_{j-1}, & -p_2p_j \\ -p_3p_1, & -p_3p_2, & |\mathbf{p}|^2 - p_3^2, & \cdot\ \cdot & -p_3p_{j-1}, & -p_3p_j \\ \cdot & & \cdot & \cdot\ \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot\ \cdot & \cdot & \cdot \\ \\ \cdot & \cdot & \cdot & \cdot\ \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot\ \cdot & \cdot & \\ -p_{j-1}p_1, & -p_{j-1}p_2, & \cdot & \cdot\ \cdot & |\mathbf{p}|^2 - p_{j-1}^2, & -p_{j-1}p_j \\ -p_jp_1, & -p_jp_2 & \cdot & \cdot\ \cdot & -p_jp_{j-1}, & |\mathbf{p}|^2 - p_j^2 \end{pmatrix} \tag{39}$$

The projection of $P_\parallel$ is obtained in the same way as that of $P\perp$. That is, a is decomposed into $a_\parallel$ and $a\perp$ orthogonal to each other. Then $P_\parallel$ is derived easily from the following relation.

$$A_\parallel + A\perp = A \tag{40}$$

$$A_\parallel + P_\parallel \cdot A \tag{41}$$

Substituting (37) and (41) into equation (40) then,

$$P_{\parallel} + P_{\perp} = I \qquad (42),$$

where I is the unit matrix. The conditions of the projection matrix are satisfied as shown in equation (43),

$$P^2{}_{\perp} = P_{\perp}, \qquad P^2{}_{\parallel} = P_{\parallel}, \qquad P_{\perp} \cdot P_{\parallel} = O \qquad (43)$$

where O is the zero matrix.

5. DETERMINATION OF ROTATION ANGLE FOR α_j SUPERPOSITION

First, consider a basis of projection axes of j-D hypercubic lattice to (j-1)-D space. If the body-diagonal axis [1, 1, 1, 1, ... , 1, 1] is chosen as a projection axis, the angle between any two basis vectors from center to vertex of j-D hyper regular simplex (α_j) is an equi-solid angle. This can be proven easily using the scalar product of two projected vectors presented by equation (37). Setting projection axis $(p_1, p_2, p_3, \ldots, p_{j-1}, p_j)$ in the orthonormal system as follows,

$$p_1 = p_2 = p_3 = \ldots = p_{j-1} = p_j = \frac{1}{\sqrt{j}} \qquad (44)$$

substitute (44) into projection matrix (39), then body-diagonal projection matrix (45) is obtained as follows.

$$\mathbf{P}_{\perp} = \begin{pmatrix} 1-\frac{1}{j}, & -\frac{1}{j}, & -\frac{1}{j}, & \cdot\ \cdot & -\frac{1}{j}, & -\frac{1}{j} \\ -\frac{1}{j}, & 1-\frac{1}{j}, & -\frac{1}{j}, & \cdot\ \cdot & -\frac{1}{j}, & -\frac{1}{j} \\ -\frac{1}{j}, & -\frac{1}{j}, & 1-\frac{1}{j}, & \cdot\ \cdot & -\frac{1}{j}, & -\frac{1}{j} \\ \cdot & \cdot & \cdot & \cdot\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\ \cdot & \cdot & \cdot \\ -\frac{1}{j}, & -\frac{1}{j}, & \cdot & \cdot\ \cdot & 1-\frac{1}{j}, & -\frac{1}{j} \\ -\frac{1}{j}, & -\frac{1}{j} & \cdot & \cdot\ \cdot & -\frac{1}{j}, & 1-\frac{1}{j} \end{pmatrix} \qquad (45)$$

The rotation angle $\varnothing_{j\text{-}1}$ for superposition of hypersimplexes corresponds to the equi-solid angle between projected basis vectors from j-D to (j-1)-D space. $\varnothing_{j\text{-}1}$ is determined by a scalar product of any two projected basis vectors, $A\perp_i$, $A\perp_{i'}$, obtained by equation (45).

$$\cos\phi_{j-1} = \frac{(\mathbf{a}_{\perp i} \cdot \mathbf{a}_{\perp i'})}{|\mathbf{a}_{\perp i}| \cdot |\mathbf{a}_{\perp i'}|} = -\frac{1}{j-1} \qquad (j \geq 2) \tag{46}$$

The verification of equation (46) is given in the case studies of j = 2, 3 and 4, as shown in section 2. That is, $\cos\phi_1 = -1, (\phi_1 = \pi)$, $\cos\phi_2 = -1/2, (\phi_2 = 2\pi/3)$ and $\cos\phi_3 = -1/3, (\phi_3 = 109.47$ deg: tetrahedral angle) are presented as a unit step of superpositional rotation angle for 1-D, 2-D and 3-D, respectively. In general, a body-diagonal projection of a hypercube from j-D to (j-1)-D space generates a (j-1)-D hyper regular simplex which has j-fold rotational symmetry as viewed from [1, 1, 1, 1, ... , 1, 1]. The j-fold rotational symmetry of the hyper regular simplex is confirmed by considering the superposition operator S_j. S_j can be represented as follows.

$$\begin{aligned} \mathbf{S}_j &= \mathbf{R}^j(\phi_j) + \mathbf{R}^{j-1}(\phi_j) + \cdots + \mathbf{R}(\phi_j) + \mathbf{I} \\ &= \mathbf{R}(j\phi_j) + \mathbf{R}\{(j-1)\phi_j\} + \cdots + \mathbf{R}(2\phi_j) + \mathbf{R}(\phi_j) + \mathbf{I} \end{aligned} \tag{47}$$

Equation (47) is equivalent to equation (48) by the Hamilton-Cayley theorem.

$$R^{(j+1)}{}_{(\varnothing j)} = I \tag{48}$$

Equation (48) means that the (j-1)-D hyper regular simplex has the j-D hypercubic body-diagonal axis in the form of a j-fold rotational axis. If j approaches infinity, cos $\varnothing_j$ converges to 0, then the projected vectors in j-D space approach an orthogonal system.

6. CONCLUDING REMARK

A geometrical interpretation of the sum of the j-th power of natural numbers is presented and visualized by considering the superposition of j-D hypersimplexes which are generated by every rotation by angle $\varnothing_j$ about the body-diagonal axis of a (j+1)-D hypercubic lattice. The expanded form of $\sum_{k=1}^{n} k^{j}$ have already been formulated using Bernoulli polynomials. The relation between the expanded form given in reference (Moriguchi 1957) and that in this work will be discussed elsewhere. This work is partly supported by the Special Coordination Funds for Promoting Science and Technology Agency of the Government of Japan.

REFERENCES

Gardner, M (1986) *Knotted Doughnuts and Other Mathematical Entertainments*, pp.192-204, Freeman.

Conway, J. H and Sloane, N. J A (1993) Sphere Packings, Lattices and Groups, *Grundlehren der mathematischen Wissenschaften* 290, pp.1-30, pp.452-456, New York. Springer-Verlag

Coxeter, H S M. (1973) *Regular Polytopes*, pp.120, 3rd edition, New York: Dover Publications, Inc.

Moriguchi, S., Udagawa, K. and Hitotsumatsu, S. (1957) *Mathematical Formula II - Series and Fourier Analysis*, pp.1-5, Iwanami Tokyo (in Japanese).

LIST OF SYMBOLS

Symbol		Meaning
$\alpha_j(\phi_j)$	:	j-dimensional hypersimplex at ϕ_j position
$l_{j,k}, l'_{j,k}$	:	Number of lattice points in k-th layer of $\alpha_j(\phi_j)$
$\mathbf{R}(\phi_j)$	:	Rotation operator of α_j
$\mathbf{I}$	:	Identity operator of α_j
$\mathbf{S}_j$	:	Superposition operator of $\alpha_j(\phi_j)$
ϕ_j	:	Rotation angle for superposition of α_j
W_j	:	Weight of lattice points of superposed α_j
M_j	:	Multiplicity of superposition of α_j
m_j	:	Mass of α_j
$\overline{m_j}$	:	Mean mass of superposed α_j
k	:	Ordinal layer number of stacking α_j
n	:	Natural number
$\mathbf{a}$	:	Arbitrary vector of hypercubic lattice
$\mathbf{p}$	:	j-dimensional projection vector
$\mathbf{P}_{\perp}$	:	j-dimensional projection matrix of $\mathbf{p}$
$\mathbf{P}_{\parallel}$	:	j-dimensional projection matrix of a perpendicular to $\mathbf{P}_{\perp}$
$\mathbf{a}_{\perp}$	:	Projected vector of a along $\mathbf{p}$
$\mathbf{A}_{\perp}$	:	Column vectors of $\mathbf{a}$
$\mathbf{A}_{\parallel}$	:	Column vectors of $\mathbf{a}_{\parallel}$ perpendicular to $\mathbf{a}_{\perp}$
θ	:	Angle between $\mathbf{p}$ and $\mathbf{a}$

Symmetry: Culture and Science
Vol. 9, Nos. 2-4, 457-463, 1998

TANGENTIAL SYMMETRIES OF PLANAR CURVES AND SPACE CURVES

Bernd Wegner

Address: Fachbereich Mathematik, TU Berlin, Strasse des 17. Juni 135, D - 10623 Berlin, E-mail. wegner@math.tu-berlin.de

1. INTRODUCTION

Mathematically symmetries occur for many objects and they have a quite general notion. They are automorphisms of a mathematical structure leaving invariant characteristic properties or quantities defined for this structure. For example, in Euclidean geometry symmetric figures mostly admit a non-trivial isometric self-map of the ambient Euclidean space, which also maps the figure onto itself. Hence this map, restricted to the figure, obviously preserves all the metric properties of that figure. But also combinatorial symmetries are considered where the self-map of the figure only preserves the combinatorial structure, while metric relations may change after having applied the map. Lots of different types of such symmetries are known, and all of them distinguish the shape of such a figure from the shape in the general case in a way which more or less immediately can be noticed by looking at that figure.

In the case of planar curves reflectional or rotational symmetries are considered as remarkable properties. Also self-similarities like in the case of spirals are of interest, and they are used to characterize certain spirals by admitting a big family of self-similarities. These properties have consequences for utilizing certain curves in the applications of geometry. Spatial generalizations of these notions are obvious, and in space the possibilities of possessing symmetries are even richer for curves. Compare a helix with a circle, for example. The aim of this short note is to explain the impact of so-called tangential symmetries on the shape of curves in the Euclidean plane and in Euclidean 3-space. As symmetry notions they seem to be quite weak, but nevertheless they have visible consequences for the shape of these curves.

2. TANGENTIAL SYMMETRIES IN THE PLANE

For closed smooth curves in the plane three types of tangential symmetries are of interest. These symmetries always consist of a group of smooth self-maps on the domain S^1 for the parametrization $c : S^1 \rightarrow E^2$ of the curve. The maps are assumed to assign points with parallel tangents to each other, and may be subject to additional conditions.

The general case without restrictive conditions has been studied by J. Shaer (unpublished) and F. J. Craveiro de Carvalho and S.A. Robertson [CR1]. Some classification results are obtained for open curves by the latter while J. Shaer provided a complete description of the shape of closed curves having a non-trivial group of tangent preserving self-maps. Mathematically spoken, the asignment of their unit tangents factorizes through a multiple covering map of the unit sphere for these curves, after having identified tangent vectors with opposite directions. For example in the locally convex case, the group Z_2 characterizes strictly convex ovals. Centrally symmetric curves may serve as examples for the nonconvex case. More general groups lead to curves with a finite number of loops, directed to the same "interior" side of the curve in the locally convex case and possessing symmetries concerning the shape of "interior" and "exterior" loops and bumps in the general case. This shape can be visualized very easily, and it obviously determines a fairly special structure for these curves.

If the tangential symmetry is assumed to preserve the normal lines in addition, then we get what has been introduced by H. Farran and S.A. Robertson [FR] as exterior self-parallelism in the general context of immersions. For closed curves this reduces to the notion of rosettes of constant width. In particular, these curves are locally strictly convex, only the cyclic group of order two can appear as a non-trivial group of self-parallelisms, and the rosettes can be generated in a very special way: A rigid line segment could be moved along the curve, connecting point and image point for the only non-trivial tangential symmetry, such that this segment always will be a normal to the curve at its endpoints. It is easy to imagine the shape of such curves, though they are more general than the classical rosettes, where a rotational symmetry can be observed. In this general case the rotational symmetry only refers to the loop structure. Details on rosettes of constant width can be found in the paper of W. Cieslak and W. Mozgawa [CM] and in [W1].

Finally, assuming as a stronger requirement, that any normal line of the curve can intersect the curve as a normal line only, we arrive at the notion of transnormality.

This has been introduced by S.A. Robertson [R1] for the more general situation of immersions into Euclidean spaces, but in the simple case of a closed curve in the plane it leads immediately to convex curves of constant width bounding planar convex domains of constant width. This characterization is a classical result. There is a vast amount of literature on ovals of constant width. For, example a comprehensive survey on classical results already could be found in the book of T. Bonnesen and W. Fenchel [BF]. Later surveys are given in handbooks on convexity. These curves may be considered as the trivial case of rosettes of constant width where no loops occur. Hence the kinematic interpretation is simple. They are frequently used in applied geometry. Most famous is their application to the construction of the cylinder and the piston for the Wankel engine.

3. TANGENTIAL SYMMETRIES FOR SPACE CURVES

For closed smooth curves in Euclidean 3-space the most general type of tangential symmetry described above will be too general to lead to conclusions on the shape of the curve. Hence we start immediately with the notions of parallelism and self-parallelism [FR]:

The *exterior parallelism* of two smooth closed curves $c_1, c_2 : S^1 \to E^3$ is defined by the following condition: For every parameter $t \in S^1$ the affine spaces normal to c_1 at $c_1(t)$ and c_2 at $c_2(t)$ coincide. This condition has been shown to be equivalent to the condition that both curves are connected by a parallel section of their normal bundles (see [W2]), i.e., there is a smooth normal vector field e_1 along c_1 such that

$$c_2(t) = c_1(t) + \lambda e_1(t) \quad \text{and} \quad prn(\nabla_{\dot{c}_1(t)} e_1) = 0 \tag{1}$$

for all $t \in S_1$, where $\dot{c}_1(t)$ denotes the tangent vector field of c_1 as usual and *prn* denotes the orthogonal projection to the corresponding normal (vector) space of c_1. Previous investigations of this notion for curves could be found in the paper [CR2] by F.J. Craveiro de Carvalho and S.A. Robertson and in [W3].

Hence the existence of a parallel mate for $c : S^1 \to E^3$ has been reduced to the search for a global parallel normal vector field along c. Generally, these vector fields only exist locally along c. Parallel transfer of the normal plane along one period of c with respect to the normal connection leads to a rotation of the normal plane, which is characterized (up to integer multiples of 2π) by an oriented angle $\alpha(c)$, which we call the *total normal twist* of c. For Frenet curves this quantity is given by their total torsion up to integer multiples of 2π. Looking at general orthonormal frame fields $\{T, e_1, e_2\}$ along c, where T denotes the unit tangent field of c, and setting

$$\omega_{12} = < \nabla_T e_1, e_2 > = - < \nabla_T e_2, e_1 > = - \omega 21, \tag{2}$$

we get for the total normal twist of c (up to integer multiples of 2π)

$$\alpha(\mathrm{c}) = \int_{S^1} \omega_{21}(t)dt \tag{3}$$

where there is no need to parametrize c with arclength.

A *self-parallelism* of c is given by a diffeomorphism $\delta : S_1 \rightarrow S_1$ such that c and $c \cdot \delta$ are parallel in the exterior sense. This is the notion of tangential symmetry to be discussed now. Clearly only closed curves with vanishing total twist possibly will admit such a tangential symmetry. The variety of these curves has been studied in much detail in a joint paper with T.F. Mersal [MW]. But already in a previous paper [W3] curves with non-trivial tangential symmetries have been related to a center curve with total normal twist being a rational multiple of 2π in the following way: Take the non-vanishing normal vector to the center curve, connecting the center curve and the original curve at some point, and apply normal parallel transfer to this vector along several periods of the center curve, until the trace of the endpoint of this vector will lead to a closed curve. This will restore the original curve. Moreover, starting the same procedure with any curve, where the total normal twist is a rational multiple of 2π, and with a suitable normal vector such that the construction will avoid singularities, we get a curve exhibiting Z_n as its group of tangential symmetries, where n is the number of periods until the trace of the end point of the vector will provide a closed curve.

This gives a fairly clear picture of space curves admitting this kind of symmetries. But there are other advantages related to the visual perception of these curves. There is an obvious one, coming from the kinematic interpretation of the parallel transfer in the normal bundle of a curve. Consider a normal frame as a rigid two-dimensional figure, moving without acceleration freely along the curve, with the constraint to remain in the normal plane forever. Then the motion will be described by the parallel transfer in the normal bundle of the curve. Hence, considering the intersection of a tangentially symmetric curve, having Z_n as its symmetry group, with its normal plane as the vertex set of a regular n-gon, this motion of the frame along the central curve will preserve this figure after one period, though a permutation of the vertices may have happened.

Interpreting the original curve as a figure located on a tube around the central curve, it carries information for the visualization of the center curve as follows: The tube may be taken as a more solid image of the center curve. The twist of the center curve may be visualized by drawing families of curves on the surface of this tube. The most appropriate curves for this will be those obtained by parallel transfer in the normal bundle of the center curve, and they will be closed curves only, if they exhibit a tangential symmetry. Clearly, then the surface of the torus bounding the tube can be foliated by curves with this kind of tangential symmetry. Furthermore, if another profile than a disk should be taken to thicken the center curve to a solid body, this only will be possible if the tangential symmetry of the original curve is respected by this profile. This can be observed in many images where closed space curves are displayed.

The more restricted form of tangential symmetry which is given by the notion of transnormality has been studied by M.C. Irwin [I], and several geometric results have been obtained for them in [W4]. Within the current context the only interesting result is, that in this case the symmetry group can be Z_2 only, and that overmore the central curve cannot avoid singularities. There are no further conclusions than those of the preceding paragraph resp. preceding section for this case.

There also are a lot of considerations concerning tangential symmetries of curves in higher-dimensional spaces. In particular, the case of Minkowski 4-space is of special interest for Relativity. But this is beyond the goal of this presentation.

4. THE CASE OF SPATIAL POLYGONS

For simple explicit constructions and examples the analytic techniques behind the theory presented above will be an essential obstruction. The theory may be reduced to C^1-curves which are piecewise C^2, and then there is a gateway for considering curves composed of pieces of circles with C^1-matchings for these pieces. Here everything can be reduced to the consideration of a finite set of data given by the finte number of circles in space. But the whole theory even can be broken down to the level of spatial polygons, and this opens a wide field for nice constructions which everybody can pursue on his own. Here just this concept should be presented in analogy to the preceding section, leaving explicit constructions to the reader.

Two simply closed polygons in space are called *parallel*, if there is a bijection between their vertices, mapping consecutive vertices to consecutive ones again, such that the angle-bisecting planes at corresponding vertices coincide. (It should be noted that this is more restrictive, than assuming that corresponding line segments are parallel.) A system of normal vectors to the edges of a polygon is called a *parallel field of normals*, if every two vectors belonging to edges with a common vertex are symmetric with respect to the reflection at the angle-bisecting plane of the polygon at this vertex. Then parallel polygons are related by a parallel normal vector field (of constant length) again. A *self-parallelism* or tangential symmetry is just a combinatorial automorphism of the combinatorial structure behind the polygon, such that the original and the relabelled polygon are parallel.

Furthermore, if a closed polygon possesses a non-trivial self-parallelism, then its total normal twist (which is defined in obvious analogy to the smooth case) vanishes mod 2π. We have the same kind of classification for tangentially symmetric polygons like in the smooth case: i) Assume that the polygon P in E^3 admits a parallel section in the normal bundle of its k-fold covering, then the obvious construction of a parallel at suitable distance has Z_k as its group of self-parallelisms. ii) Viceversa, every polygon possessing more than one self-parallelism can be obtained by the preceding construction from a suitable central polygon.

Details for these considerations can be found in [W5]. Very explicit calculations in the case of spatial quadrilaterals as center polygons have been obtained in [W6]. Everything can be reduced to the framework of elementary Euclidean geometry in 3-space. Examples of self-parallel polygons, which are not restricted to a plane, start with 8 vertices at least. Those with 8 vertices bound a PL-Moebius strip, and the central quadrilateral has to satisfy very special constraints for its angles and side lengths. Constructions with pentagons as central polygons are easier to visualize. They provide very simple families of imbedded PL-Moebius strips having a bounding polygon with 10 vertices, such that the strip may be composed from planar pieces of constant breadth. This is the case of symmetry group Z_2.

For symmetry group Z_4 and a center quadrilateral we receive recipes for composing four bars with the same square as their profile, such that the result will be a skew frame and such that the edges resulting from the bars form one connected closed polygon (with 16 vertices). The common picture frames have a planar rectangle as their basis, and the edges coming from the bars decompose into four rectangles of different sizes. It is easy to produce wooden models for these skew frames also by cutting the bars into pieces of appropriate lengths, such that planes for the sawing have the right angle with respect to the center line of the bar. Comparing this model with other models of compositions of four bars of variable rectangular profile to a skew framework, the symmetric construction will really appear as the most symmetric (and appealing) solution. Anyway, there will be no other solutions having a square as their costant profile.

The same will be true for other groups of tangential symmetries, only the regular polygon for the profile will change. Hence there will be a more simple solution for Z_3 than for Z_4, but for building a model, it will be easier to get bars with a square profile than with an equilateral triangle. For rectangular profiles which are not squares, the Z_2 models will be good solutions, but then the boundary polygon will decompose into two linked octogons then. Finally it should be observed, that these construction also may be of interest for non-closed polygons: They will provide a solution of the problem to connect two planes in space with bars of the same (regular or semi-regular) profile, such that the starting point and the end point of the connection may be prescribed, the bars start resp. arrive in normal direction at the planes, and the resulting framework fits properly to a polyhedron, i.e., the matchings are done face to face without any prominent pieces.

Conclusion: The preceding considerations show, that a rather general concept of symmetry for curves in the plane and in 3-space leads to interesting versions of visible symmetries for them. These symmetries motivate in many cases why these curves are preferred for applications and for geometric constructions.

REFERENCES

[BF] Bonnesen, T. and Fenchel, W. (1974) *Theorie der konvexen Korper*, Ber Reprint, Springer-Verlag.

[CM] Cieslak, W. and Mozgawa, W. (1987) On rosettes and almost rosettes, *Geom. Dedicata* 24, 221-228.

[CR1] Craveiro de Carvalho, F. J. and Robertson, S A. *The parallel group of a plane curve* (to appear).

[CR2] Craveiro de Carvalho, F. J. and Robertson, S A. (1989) Self-parallel curves, *Math. Scand.* 65, 67-74.

[FR] Farran, H. R. and Robertson S.A. (1987) Parallel immersions in Euclidean space, *J. London Math. Soc.* (2) 35, 527-538.

[I] Irwin, M. C. (1967) Transnormal circles. *J. London Math. Soc.* 42, 545-552.

[MW] Mersal, T F and Wegner, B. (1997) Variation of the total normal twist of closed curves in Euclidean spaces, *Proceedings Conf. on Geometry and Topology*, Braga (to appear)

[R1] Robertson, S A. (1964) Generalized constant width for manifolds. *Michigan Math. J.* 11, 97-105.

[W1] Wegner, B. (1989) Some global properties and constructions for closed curves in the plane, *Geom. Dedicata* 29, 317-326.

[W2] Wegner, B. (1989/1991) Some remarks on parallel immersions, *Coll. Math Soc. J. Bolyai* 56, 707-717.

[W3] Wegner, B. (1991) Self-parallel and transnormal curves, *Geom. Dedicata* 38, 175-191.

[W4] Wegner, B. (1972) Globale Sätze uber Raumkurven konstanter Breite I, II, *Math. Nachr.* 53, 337-344, ibid. 67 (1975), 213-223.

[W5] Wegner, B. (1991) Exterior parallelism for polyhedra, *Math. Pannonica* 2, 95-106.

[W6] Wegner, B. (1991) *Self-parallel polygons and polyhedra*, Geometry, Proc. Conf. Thessaloniki / Greece, 444-451.

There are many disciplinary periodicals and symposia in various fields of art, science, and technology, but broad interdisciplinary forums for the connections between distant fields are very rare. Consequently, the interdisciplinary papers are dispersed in very different journals and proceedings. This fact makes the cooperation of the authors difficult, and even affects the ability to locate their papers

In our 'split culture', there is an obvious need for interdisciplinary journals that have the basic goal of building bridges ('symmetries') between various fields of the arts and sciences Because of the variety of topics available, the concrete, but general, concept of symmetry was selected as the focus of the journal, since it has roots in both science and art.

SYMMETRY: CULTURE AND SCIENCE is the quarterly of the INTERNATIONAL SOCIETY FOR THE INTERDISCIPLINARY STUDY OF SYMMETRY (abbreviation: ISIS-Symmetry, shorter name: Symmetry Society). ISIS-Symmetry was founded during the symposium *Symmetry of Structure* (*First Interdisciplinary Symmetry Symposium and Exhibition*), Budapest, August 13-19, 1989. The focus of ISIS-Symmetry is not only on the concept of symmetry, but also its associates (asymmetry, dissymmetry, antisymmetry, etc) and related concepts (proportion, rhythm, invariance, etc.) in an interdisciplinary and intercultural context. We may refer to this broad approach to the concept as *symmetrology*. The suffix *-logy* can be associated not only with knowledge of concrete fields (cf., biology, geology, philology, psychology, sociology, etc.) and discourse or treatise (cf., methodology, chronology, etc), but also with the Greek terminology of proportion (cf., *logos*, *analogia*, and their Latin translations *ratio*, *proportio*)

The basic goals of the *Society* are

(1) to bring together artists and scientists, educators and students devoted to, or interested in, the research and understanding of the concept and application of symmetry (asymmetry, dissymmetry);

(2) to provide regular information to the general public about events in symmetrology;

(3) to ensure a regular forum (including the organization of symposia, congresses, and the publication of a periodical) for all those interested in symmetrology.

The Society organizes the triennial Interdisciplinary Symmetry Congress and Exhibition (starting with the symposium of 1989) and other workshops, meetings, and exhibitions. The forums of the Society are *informal* ones, which do not substitute for the disciplinary conferences, only supplement them with a broader perspective.

The Quarterly - a non-commercial scholarly journal, as well as the forum of ISIS-Symmetry - publishes original papers on symmetry and related questions which present new results or new connections between known results. The papers are addressed to a broad non-specialist public, without becoming too general, and have an interdisciplinary character in one of the following senses:

(1) they describe concrete interdisciplinary 'bridges' between different fields of art, science, and technology using the concept of symmetry;

(2) they survey the importance of symmetry in a concrete field with an emphasis on possible 'bridges' to other fields.

The Quarterly also has a special interest in historic and educational questions, as well as in symmetry-related recreations, games, and computer programs.

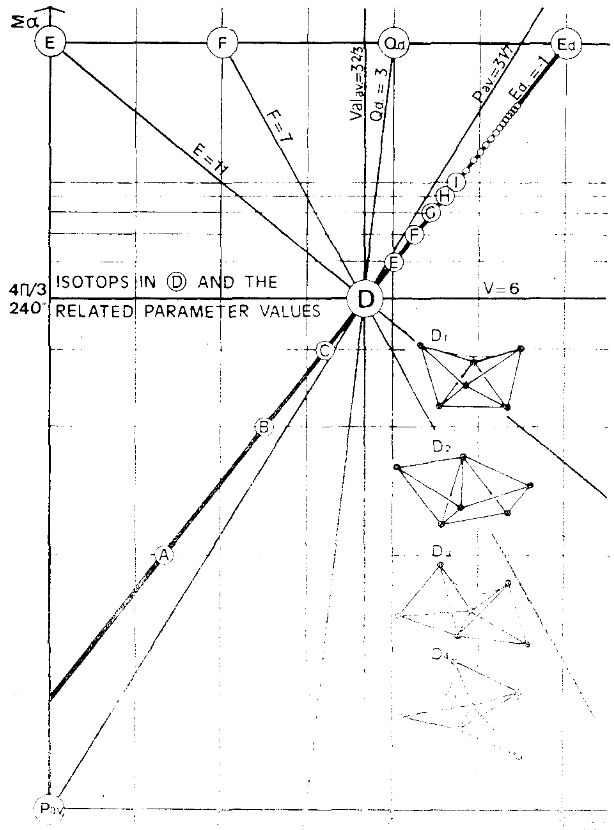
The regular sections of the Quarterly.

- **Symmetry: Art & Science** (papers classified as humanities, but also connected with scientific questions)
- **Symmetry: Science & Art** (papers classified as science, but also connected with the humanities)
- **Symmetry in Education** (articles on the theory and practice of education, reports on interdisciplinary projects)

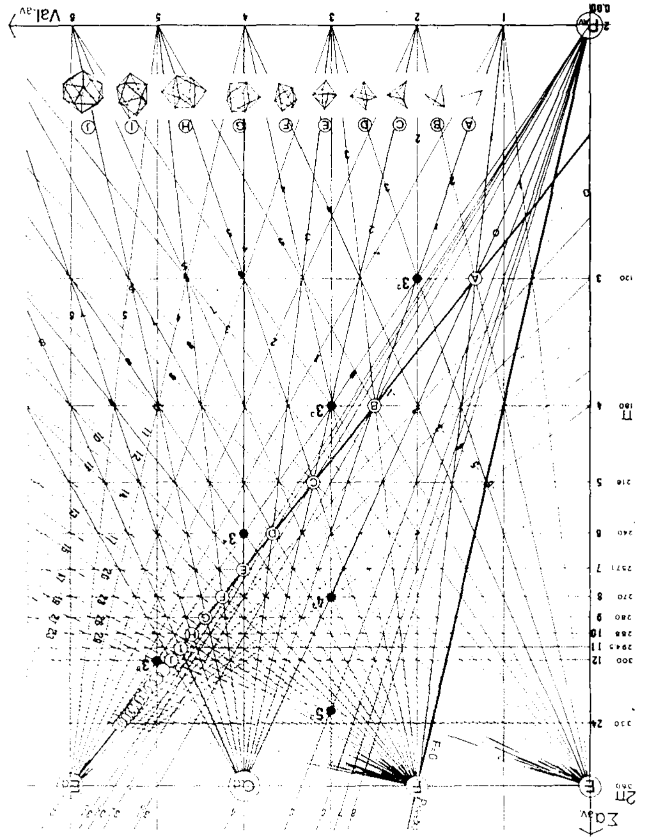
There are also *additional, non-regular* sections.

Both the lack of seasonal references and the centrosymmetric spine design emphasise the international character of the Society; to accept one or another convention would be a 'symmetry violation'. In the first part of the abbreviation *ISIS-Symmetry* all the letters are capitalized, while the centrosymmetric image ¡SIS! on the spine is flanked by 'Symmetry' from both directions. This convention emphasises that ISIS-Symmetry and its quarterly have no direct connection with other organizations or journals which also use the word *Isis* or *ISIS*. There are more than twenty identical acronyms and more than ten such periodicals, many of which have already ceased to exist, representing various fields, including the history of science, mythology, natural philosophy, and oriental studies ISIS-Symmetry has, however, some interest in the symmetry-related questions of many of these fields.

ISIS-Symmetry
Budapest, P.O Box 994
H-1245 Hungary



Symmetry



ISIS-Symmetry
Budapest, P.O Box 994
H-1245 Hungary